



# Наука о зрении

*От фотонов к феноменологии*

Стивен Пальмер  
Брайан Ванделл  
Стивен Шварц

ЧЕРНОВИК  
Главы с 1 по 8-ую.  
Перевод: А. Шадрин  
Декабрь 2011



Часть первая

**ОСНОВЫ НАУКИ О ЗРЕНИИ**



Большинству из нас дарована способность видеть. И как это нам удается, вроде бы не секрет: просто открываем глаза и смотрим! Мы глядим и воспринимаем при этом сложный массив жизненно важных предметов, расположенных в окружающем нас трехмерном пространстве. К примеру, на рис. 1.1.1 показана типичная сцена Беркли-кампус Калифорнийского университета: на переднем плане студенты семенят через сатеровские ворота, кругом деревья и колоколенка на заднем плане. Мы схватываем все это моментально и без усилий и не полагаем наше восприятие, чем-то особенным. Однако когда начинаешь разбираться в механизмах зрения, то приходит понимание того, что процесс зрительного восприятия невероятно сложен и то, что у нас есть способность видеть, кажется уже почти чудом.

Богатая зрительная картина, возникающая при восприятии естественных сцен (к примеру аналогичных тем, что показаны на рис. 1.1.1) образуется тогда, когда нервные ткани, выстилающие заднюю поверхность глаз, стимулируются двухмерным световым паттерном, включающим в себя лишь фрагменты окружающих предметов: большая часть колокольни, как видим, скрыта деревьями, те в свою очередь прикрыты башнями ворот. Однако мы не воспринимаем колокольню, как парящую в воздухе, а деревья как продырявленные башнями. Но даже те объекты сцены, что, казалось бы, присутствуют в кадре целиком, в частности башенки ворот и студенты, на самом деле видны лишь частично, поскольку их обратные стороны закрыты передними. Как же мы можем так легко и быстро воспринимать трехмерную сцену во всей ее полноте и единстве на основе очевидно неполного двухмерного светового паттерна, действующего на наши глаза?

Сие есть фундаментальный вопрос науки о зрении — *визиологии* — и все, о чем пойдет речь в этой книге — это развернутый ответ на него с научной точки зрения. Нет ничего страшного в том, что мы начинаем книгу с вопроса: на первом этапе изучения любой научной дисциплины задаются вопросы, ответы на которые на первый взгляд кажутся очевидными и простыми. В наших лекциях будет поднято множество очень серьезных вопросов и часть из них мы сразу перечислим:

- почему предметы выглядят окрашенными?
- каким образом мы отличаем крупный удаленный предмет от маленького близкого?
- за счет чего мы можем определить, какой участок в зрительном образе соответствует фрагменту объекта сцены?
- как мы определяем назначение видимых объектов?
- за счет чего мы отличаем то, что в данный момент мы движемся относительно тех или иных предметов, а после предметы относительно нас?
- правда ли, что новорожденные видят мир также как мы?
- могут ли люди «видеть» не имея исходного представления о том, что видят?

Однако, начав с вопросов, мы, в итоге, должны постараться найти на них ответы. Поэтому большая часть нашей книги посвящена рассказу о том, как ученые исследуют зрение и о том, что им удалось обнаружить.





**Рис. 1.1.1** Беркли-кампус Калифорнийского университета. Некоторые объекты этой сцены не видны полностью, и мы лишь догадываемся о том, каков низ колокольни или деревьев, прикрытых колоннами ворот, а также о том, каковы обратные стороны всех видимых предметов.

Очевидно, что слагаемые ответов на поставленные вопросы являются нам из разных наук — биологии, психологии, математики, нейрофизиологии, лингвистики и когнитивной антропологии; все перечисленное — часть единой науки — *когнитологии* (науки о мышлении).

Базовым посылом когнитологии является то, что проблемы понимания механизмов мышления могут быть решены тем быстрее и тем полнее, чем с большего числа точек зрения подходить к решению этих проблем. Современная визиология целиком основана на таком подходе, получившем название *междисциплинарного*. Наука о зрении лежит на стыке знаний из разных областей и являет собой плотную мозаику, каждый элемент которой предоставлен самостоятельной научной дисциплиной. Итак, мы показали, что наука о зрении междисциплинарна и является составной частью когнитологии — удивительной и увлекательной науки. В нашей книге мы помимо всего попытались передать то чувство азарта, которое владело и владеет всеми исследователями зрения и без которого мы не можем двигаться дальше в понимании его механизмов.

Итак, задача текущей главы, построить вводный понятийный каркас визиологии, основанный на трех направлениях исследования:

- исследовании феноменов зрительного восприятия;
- исследовании характера оптической информации;
- изучении физиологии нервных процессов в зрительной системе.

Разбор как каждого направления в отдельности, так и взаимоотношений между ними — залог понимания нами зрительных механизмов.

В первом разделе текущей главы природу зрительного восприятия мы рассмотрим в эволюционном ключе, отвечая при этом на вопрос «зачем?». Дадим определение и поговорим о наиболее важных свойствах зрительного восприятия, рассматривая его как инструмент *выживания* организмов.

Затем мы разберем характер оптической информации, поскольку любая зрительная система зиждется на структуре света, отраженного от окружающих предметов и попадающего в глаза.

В третьем разделе главы мы достаточно подробно коснемся анатомии зрительной системы и детально разберем механизм преобразования световой энергии в отклик фоторецептора — энергии фотона в нервный импульс.

Наконец, мы в общих чертах опишем физиологию той части нервной системы, что лежит в основе нашей способности видеть: глаза — неотъемлемая часть зрительного аппарата, но при этом суть огромные выросты головного мозга, функцию которого ученые только-только начинают понимать.

Отметим, что по некоторым из упомянутых направлений материал данной главы примитивен и неполон, но в начале пути важно понять суть самих направлений и характер взаимоотношений между ними, чтобы уже в дальнейшем получить полное научное объяснение механизмов зрения. Короче говоря, в текущей главе мы рассмотрим общую схему того, что в последующих главах будет разобрано подробно.

## 1.1 ЗРИТЕЛЬНОЕ ВОСПРИЯТИЕ

До сих пор мы полагали, что смысл словосочетания «зрительное восприятие» очевиден, и что читатель воспринимает данный текст через орган зрения (а не на слух, к примеру) и, следовательно, знает, каково зрительное ощущение. Однако прежде чем двигаться дальше, мы должны дать точное определение зрительного восприятия.

### 1.1.1 Определение зрительного восприятия

В контексте данной книги зрительному восприятию будет дано следующее определение:

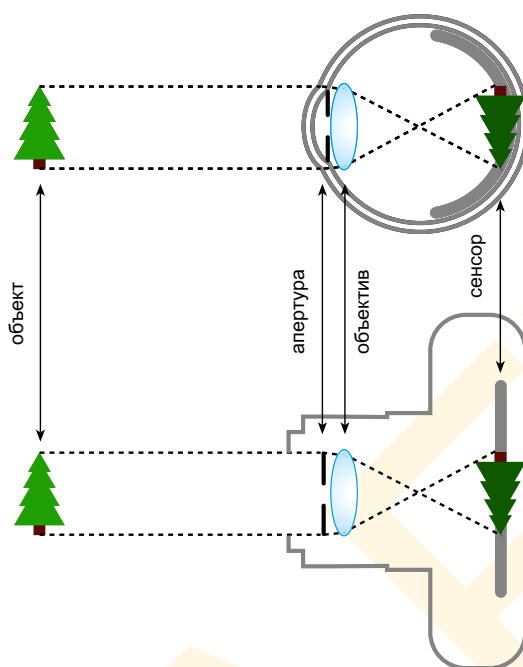
⇒ **Зрительное восприятие** — это процесс познания особью окружающей обстановки и происходящих в ней событий, реализуемый через извлечение информации из световых потоков (отраженных от чего-либо или испущенных чем-либо).

Некоторые аспекты этого определения стоит рассмотреть детально:

1. Зрительное восприятие нацелено на *получение знания*, или, иными словами, *достижение осведомленности* о происходящем в окружающем мире. Сие значит, что зрение в основе своей есть когнитивный акт (от латинского «*cognoscere*», означающего «знать» или «изучать»), и в корне отлично от чисто оптических процессов, скажем фотографических.

Однако механическое подобие фотокамеры глазу свидетельствует о том, что физический компонент восприятия аналогичен показанному на рис. 1.1.2: в самом деле, в логике оптических феноменов фотокамера и глазное яблоко очень схожи между собой (что мы покажем в разделе 1.2), но нет совершенно никакого сходства в логике феноменов *восприятия* — фотокамера не может ничего воспринимать и она ничего *не знает* о фиксируемой ею сцене. Фотоизображения лишь несут *информацию*, тогда как зрячие





**Рис. 1.1.2** Очевидно, что по характеру оптических процессов глаз более всего походит на обычный фотоаппарат. Однако во всем остальном он принципиально отличен.

люди и животные из этой информации извлекают *знание*. Это знание позволяет особи действовать адекватно той или иной жизненной ситуации.

2. Знание, полученное путем зрительного восприятия, касается предметов и событий окружающей среды. Однако восприятие — это не один лишь субъективный зрительный опыт наблюдателя: мы не можем сказать, например, что возникновение даже высокодетализированной галлюцинации или зрительной фантазии можно полагать актом зрительного восприятия. Мы очень подробно будем рассматривать природу человеческого субъективного опыта (в частности, в 13-й главе, где пойдет подробный разговор о зрительной осведомленности), который, безусловно, является частью зрительного восприятия, но только тогда, когда он свидетельствует о свойствах *реальных* объектов.

3. Визуальное знание об окружающем мире формируется путем обработки оптической информации. Данный аспект нашего определения восприятия предполагает «метатеоретический» подход к пониманию механизмов зрительной перцепции и мышления, один из которых базируется на концепции информации и ее обработки. Подробно об этом подходе мы будем говорить во второй главе нашей книги, а сейчас достаточно сказать, что это подход, позволяющий исследователям рассуждать о зрении человека в логике «зрения» компьютерного. Опять же: о компьютерном зрении мы поговорим подробнее в Главе 13, когда будем обсуждать проблему визуального распознавания.

4. Информация, которая обрабатывается в процессе зрительного восприятия, поступает из света, испущенного, либо *отраженного* объектами действительности (оптическая информация). Оптическая информация — это основа зрения, и появляется она в результате *взаимодействия* света с объектами окружающего нас мира. Поскольку возникающая в результате такого взаимодействия реструктуризация световых потоков определяет то, какая информация об объекте будет доступна зрительной системе в

первую очередь, сие является отправной точкой любого систематического анализа механизмов зрения (Gibson, 1950). В разделе 1.2 мы покажем, что первые проблемы, с которыми сталкивается исследователь в понимании механизмов зрения, возникают из-за неизбежных потерь, случающихся, когда свет, отраженный от трехмерных объектов, проектируется на двухмерную заднюю внутреннюю светочувствительную выстилку глазного яблока — сетчатку. Следовательно, изучение того, какую информацию несут двухмерные проекционные изображения — это важнейший аспект исследования в науке о зрении. Отметим, что он активно прорабатывается в том числе математиками, пытающимися открыть новые источники визуальной информации, которыми активно пользуется зрение, но которые нам пока неизвестны.

### 1.1.2 Зрение в контексте эволюции

Итак, мы рассмотрели то, что есть зрительное восприятие, теперь же нам предстоит выяснить, *зачем* оно нужно? Учитывая высокую биологическую значимость зрения в жизни огромного числа видов животных, ответ может звучать так: **зрение развилось как гарант выживания и успешного размножения организмов**. Желаемые объекты и ситуации, в частности, такие как полноценная еда, надежное убежище и спаривание — должны быть увидены и достигнуты. Опасных ситуаций и объектов, таких как отвесные скалы, падающие предметы, а также голодные и злые хищники нужно избегать. Иными словами, чтобы вести себя в эволюционно адаптивном ключе, мы должны тем или иным образом получать информацию о том, какие предметы находятся вокруг, где они расположены и что они могут нам дать; все чувства: зрение, слух, осязание, вкус и обоняние — участвуют в этом действии.

Существуют некоторые животные, у которых доминируют неоптические сенсоры, к примеру: летучие мыши ориентируются по слуху, но у *homo sapiens* (равно как и у многих других видов теплокровных) зрение доминирует. Так произошло потому, что зрение предоставляет организму пространственно точные сведения об объектах окружающего мира на определенном расстоянии от этих объектов. Иными словами, оно дает высокореалистичную информацию о расположении и свойствах объектов окружающего мира с безопасного для особи расстояния.

Однако слух и обоняние порой предоставляют данные даже с больших расстояний, но при этом (по меньшей мере у людей) редко когда обеспечивают должную точность в плане идентификации и месторасположения объектов.

Осязание и вкус безусловно предоставляют особи более точную информацию о физических свойствах предметов, поскольку эти органы чувств включаются только при непосредственном контакте с объектом, но при этом не в состоянии дать информации на расстоянии.

В эволюционном контексте зрительное восприятие оправданно только тем, что оно высокоточно. Если бы информация, которую несет свет, была недостаточной для того, чтобы отличить один объект от другого или узнать, каково пространственное расположение объекта, зрение никогда не развилось бы до того исключительного уровня, каков он у человека. Свет — это невероятно эффективный носитель информации об окружающем мире, и человеческое зрение пользуется этим носителем, что называется «по полной программе». Действительно, зрение потому и высокоточно, что эффективно утилизирует оптическую информацию. Одним словом: *WYSIWYG* — *What You See Is What You Get* — что увидел, то и получил. Когда сей принцип выполняется, мы говорим о т.н. *веридикальном восприятии* (от латинского «*veridicus*», означающем «говорить



только правду») — восприятию, отражающем истинное положение дел в окружающей особь обстановке.

Мы теперь видим, что зрительное восприятие целиком подпадает под определение веридикального, то есть гарантирующего всю полноту информации об окружающем мире, и, казалось бы, зрение — суть настешь распахнутое окно в реальность! Не тут-то было...

В заключение данного раздела придется доказать то, что зрительное восприятие — это отнюдь не «распахнутое окно в реальность», но активно действующая, значимая *модель окружающего мира*, позволяющая особи эффективно *прогнозировать* то, что произойдет в ближайшем будущем (если особь эта предпримет те или иные действия) и тем самым повышать свои шансы на выживание. Сие основополагающий феномен зрительного восприятия, к которому мы не раз еще вернемся на страницах нашей книги.

### 1.1.3 Зрительное восприятие как конструктивный акт

Первая проблема, которую нам предстоит обсудить, состоит в том, действительно ли работа зрения отвечает требованию «что вижу, то и получу»?; действительно ли зрение непогрешимо точно? Сии вопросы очень важны, и ответы на них покажут нам, можем ли мы всерьез говорить о том, что зрение — это действительно «окно в реальность».

#### 1.1.3.1 Адаптация и последствия

**Адаптация.** Одним из доказательств того, что зрительный опыт не является окном в реальность является тот факт, что зрительное восприятие меняется во времени, приспособляясь к специфике условий. Когда, к примеру, вы входим в кинотеатр с залитой солнцем улицы, то не видим ничего, за исключением изображения на экране. Спустя несколько минут, однако, мы уже можем различить людей вокруг, а спустя 20 мин (или около того) наше зрение на удивление хорошо распознает все детали обстановки зрительного зала. Сие повышение чувствительности к свету именуется «*темновой адаптацией*»: стены кинозала и зрители вокруг были изначально, но мы не могли их видеть первое время, поскольку наша зрительная система не обладала в тот момент достаточной чувствительностью, но со временем обрела ее.

Другим типичным примером темновой адаптации может служить «поведение» звездного неба: когда ясной ночью мы выходим из хорошо освещенной комнаты на улицу, то звезд на небе мало и они довольно тусклы. Однако спустя несколько минут, их становится больше и выглядят они начинают ощутимо ярче. Спустя 20-30 мин. мы видим, что небо усеяно тысячами звезд, которых мы напрочь не видели в первый момент. Понятно, что звезды не стали ярче от того, что мы присмотрелись к ним, но наша зрительная система повысила свою чувствительность к их свету.

Адаптация — важнейший феномен зрительного восприятия. Как мы покажем позднее, зрительное впечатление становится слабее<sup>1</sup> вследствие длительной экспозиции в широком диапазоне вариантов стимуляции: хроматической, пространственной,

<sup>1</sup> Тем, кто знаком с основами нормальной физиологии человека, может показаться странным, что при темновой адаптации зрительная система становится более чувствительной, а не менее. Однако сие лишь кажущееся отличие от других видов адаптации, которое легко устранить, если рассудить так, что в процессе темновой адаптации зрительная система становится *менее* чувствительной к *темноте*.

стимуляции движением и т.д. Изменения в зрительном впечатлении демонстрируют то, что зрительное восприятие далеко не всегда является «окном в реальность» — на разных стадиях адаптации наши зрительные впечатления разнятся при неизменном физическом воздействии на орган зрения. То есть, со временем что-то меняется в нашей зрительной системе, но не в окружающем нас мире. Нельзя забывать при этом, что хотя из-за адаптации некоторые объекты и явления могут оказаться невидимыми, то, что все-таки было воспринято, является точным отражением реальности.

**Последствия.** Иным следствием длительной (или очень сильной) стимуляции является существование т.н. *зрительных последствий*. Рассмотрим следующим пример: когда некто фотографирует нас со вспышкой, то наше первое впечатление — это слепящее бесформенное пятно вместо фотографа. Сие безусловно веридикальное восприятие, но после которого на том самом месте, где исходно была вспышка, долго висит темное «облако». Данный *постобраз* будет накладываться на все, куда бы мы ни поглядели в течение нескольких минут, искажая наше зрительное впечатление вплоть до того, что некоторые предметы окажутся невидимыми. Понятно, что сие не является веридикальным восприятием, поскольку постобраз длится много дольше, чем фактическая вспышка.

Однако далеко не все последствия приводят к тому, что предметы становятся неразличимы: некоторые из последствий могут исказить лишь отдельные свойства видимых объектов. На рис. 1.1.3 показан пример пространственного последствия: поглядим на правую часть рисунка и убедимся, что прутья верхней и нижней решеток вертикальны и одинаковы; затем на 30-60 сек. остановим взгляд на наклоненных решетках в левой части рисунка, сосредоточившись на горизонтальной полоске; после чего посмотрим на квадратик между верхней и нижней решетками правой части рисунка — верхняя решетка будет казаться наклоненной влево, нижняя — вправо. Сие искажение в восприятии еще раз доказывает, что то, что наше зрительное впечатление есть результат *воздействия* внешнего мира на нашу нервную систему в ее текущем состоянии.

### 1.1.3.2 Иллюзии и реальность

Существует множество проявлений систематического неверидикального восприятия, которые обычно именуют «зрительными иллюзиями». Одним очень эффектным примером, с которым, вполне вероятно, знаком читатель — является т.н. *лунная иллюзия*: ночное светило кажется крупнее, когда оно непосредственно над горизонтом, нежели в вышине ночного неба. Чем объяснить данное явление?

Многие думают, что лунный эффект — это следствие оптических искажений, вносимых атмосферой. Другие предполагают, что причина в определенной форме орбиты планеты. Однако оптический размер Луны остается постоянным на всем протяжении ее хода по небосклону: доказать сие очень легко, выполнив серию фотоснимков в период восхода нашего ночного светила и его хода по небосклону — диаметр Луны на фотографиях останется практически неизменным. Таким образом, меняется лишь результат нашего восприятия размера планеты. В данном контексте возникновение иллюзии — суть неверидикальное восприятие, поскольку изображение Луны на сетчатке нашего глаза не изменило своего пространственного размера (равно как и на фотографиях планеты). В 7-й главе мы детально рассмотрим механизм возникновения иллюзии большой Луны (Kaufman и Rock, 1962; Rock и Kaufman, 1962), тогда как сейчас



важно лишь показать, что воспринимаемое нами отличие в размерах планеты при разных ее высотах над горизонтом — сугубо иллюзорно.

Существует множество других зрительных иллюзий, наглядно демонстрирующих то, что зрительное восприятие нельзя назвать инструментально-точным. Сей тезис проиллюстрирует рис. 1.1.4:

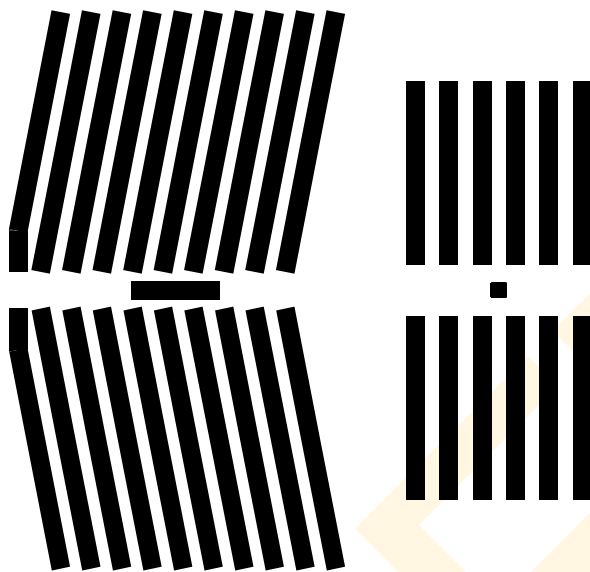
- А — метрически длина стрелок одинакова;
- В — то же у горизонтальных линий;
- С — длинные линии вертикальны и параллельны друг другу;
- D — диагональные линии фактически коллинеарны;
- Е — диаметры центральных кружков метрически одинаковы.

В каждом из приведенных примеров наша зрительная система несколько обманута сбоями восприятия относительно, казалось бы, очевидных свойств простейших объектов. Данные иллюзии утверждают нас в выводе о том, что зрению свойственно ошибаться и, о нем нельзя говорить как об окне в окружающий нас мир. Следовательно, в процессе обработки информации, несомой световыми потоками, зрительная система по меньшей мере отчасти *дистраивает* образ действительности. Как мы увидим далее, характер этой дстройки предполагает определенные скрытые допущения, которых мы не осознаем. Когда допущения эти оказываются неверными, возникает зрительная иллюзия (вопрос зрительных иллюзий еще не раз будет обсуждаться на страницах нашей книги, в частности в 7-й главе).

После разговора о зрительных иллюзиях легко впасть в крайность и начать думать о зрении как о сильно неточной и нереалистичной системе. Сие было бы ошибкой. Как уже было сказано ранее, эволюционный смысл зрения именно в его точности, или, если угодно, зрение столь точно, чтобы именно в нем была необходимость. Даже иллюзорное восприятие по большому счету весьма точно, к примеру: на рис. 1.1.4В действительно имеются две коротких горизонтальных линии и две длинных, которые не соприкасаются друг с другом; неточность восприятия в данном случае касается лишь одного параметра: относительной длины горизонтальных линий, и разница между результатом восприятия и реальностью весьма скромна. Более того, зрительные иллюзии почти не проявляют себя в повседневной жизни, возникая по большей части на страницах книг по физиологии зрения.

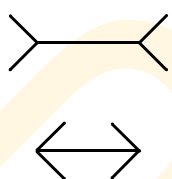
Итак, было бы ошибкой полагать, что несущественные погрешности, вносимые нашим зрением в картину мира, умаляют его эволюционную ценность. Позднее мы поговорим о том, что перцепционные ошибки, порожденные зрительными иллюзиями, фактически являются относительно безобидными сторонними эффектами тех же процессов, что обеспечивают веридикальное восприятие в привычной для особи обстановке (см. гл. 7). Важным пунктом текущего разговора является то, что факт существования зрительных иллюзий убедительно доказывает: восприятие — это не просто регистрация объективной реальности — это гораздо большее.

Итак, после первого знакомства с феноменом зрительных иллюзий, мы можем сказать, что нет особых оснований утверждать, будто зрительное восприятие обязано быть окном в реальность: предметы обстановки, которые мы с такой легкостью видим, не являются первопричиной нашего восприятия — верно говорить, что восприятие вызывается двухмерными световыми паттернами, стимулирующими наши глаза (чтобы убедиться в справедливости данного утверждения, можно их просто закрыть: объекты действительности по-прежнему будут существовать, но более уже не оставят зрительного впечатления). Таким образом, для предоставления нам информации о трехмерном окружающем мире, зрение должно являть собой т.н. *интерпретативный про-*



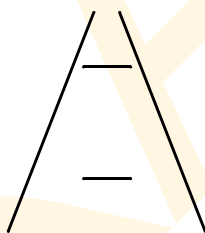
**Рис. 1.1.3** Пространственные последействия. Если на полминуты-минуту сосредоточить взгляд на горизонтальной полоске слева, а затем перевести его на квадратик справа, то верхняя решетка будет казаться слегка повернутой влево от вертикали, а нижняя — вправо.

Какая из  
горизонтальных  
линий длиннее?



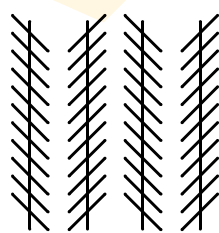
A

Какая из  
горизонтальных  
линий длиннее?



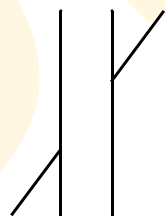
B

Вертикальные линии  
параллельны или нет?



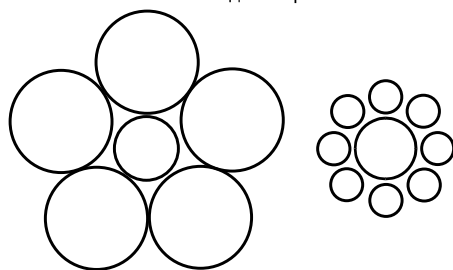
C

Диагональная линия  
ровная или нет?



D

Какой из центральных  
кружков больше в  
диаметре?



E

**Рис. 1.1.4** Зрительные иллюзии: А, В — горизонтальные отрезки одинаковы по длине, несмотря на то, что выглядят разными; С — вертикальные линии физически параллельны, несмотря на то, что кажутся сходящимися; D — диагональные линии коллинеарны, несмотря на то, что выглядят смещенными друг относительно друга; Е — внутренние кружки одинакового диаметра.



цесс, определенным образом преобразующий сложные двухмерные световые мозаики,двигающиеся по задней внутренней поверхности глазного яблока, в стабильные *перцепты* (образы) объемных объектов, расположенных в трехмерном пространстве. Итак, мы подходим к выводу, что восприятие объектов действительности зиждется именно на *интерпретации* структуры двухмерных изображений, образованных световыми потоками на задней поверхности глаз, нежели на непосредственной регистрации физической реальности.

### 1.1.3.3 Двойственные фигуры

Убедительной демонстрацией интерпретативного характера зрения являются т.н. *двойственные фигуры*: единые изображения, которые могут восприниматься как минимум двояко. Некоторые эффектные примеры двойственных фигур показаны на рис. 1.1.5. Фигура «ваза/лица» (рис. 1.1.5. А) может восприниматься как белая ваза на черном фоне (А1), либо как силуэты двух лиц на белом фоне (А2). Куб Неккера (1.1.5 В) — может восприниматься в двух ориентациях по отношению к наблюдателю: в первом варианте наблюдатель «находится» слева сверху (В1); во втором — справа снизу (В2). Когда восприятие «выворачивается», изменяется интерпретация глубины залегания линий: передние края уходят назад, задние — вперед.

Несколько иной вариант двойственности продемонстрирует рис. 1.1.5 С: он может восприниматься как утка, смотрящая влево (С1), либо же как заяц, глядящий вправо (С2). И вновь мы наблюдаем то, что интерпретация линий сдвигается от одного перцепта к другому, но в данном случае, акцент смещается с одного края картинка на другой: клюв утки превращается в заячьи уши, а хохолок на утином затылке — в заячий нос.

В контексте двойственных фигур мы должны отметить две очень важных вещи:

— во-первых, варианты интерпретации изображений взаимоисключающи и воспринимаются в данный момент только по какому-то одному из вариантов, к примеру — либо утка, либо заяц, но не оба одновременно. Сие согласуется с той идеей, что восприятие прибегает именно к интерпретативному моделированию, поскольку только одна из возможных моделей сможет быть применена к сенсорным данным в текущий момент времени;

— во-вторых, если мы хотя бы раз осознали оба варианта интерпретации, наше восприятие становится, как говорят, *мультиустойчивым* (*мультистабильным*), то есть переходящим, по мере рассматривания изображения, из одного варианта его интерпретации в другой и обратно. Сие свидетельствует о том, что обе модели конкурируют друг с другом и «победитель», что называется «берет измором», а «побежденный» уступает.

Описанные феномены можно смоделировать в рамках теории нейронных сетей, о которой речь пойдет в гл. 6.

### 1.1.4 Зрительное восприятие как процесс моделирования

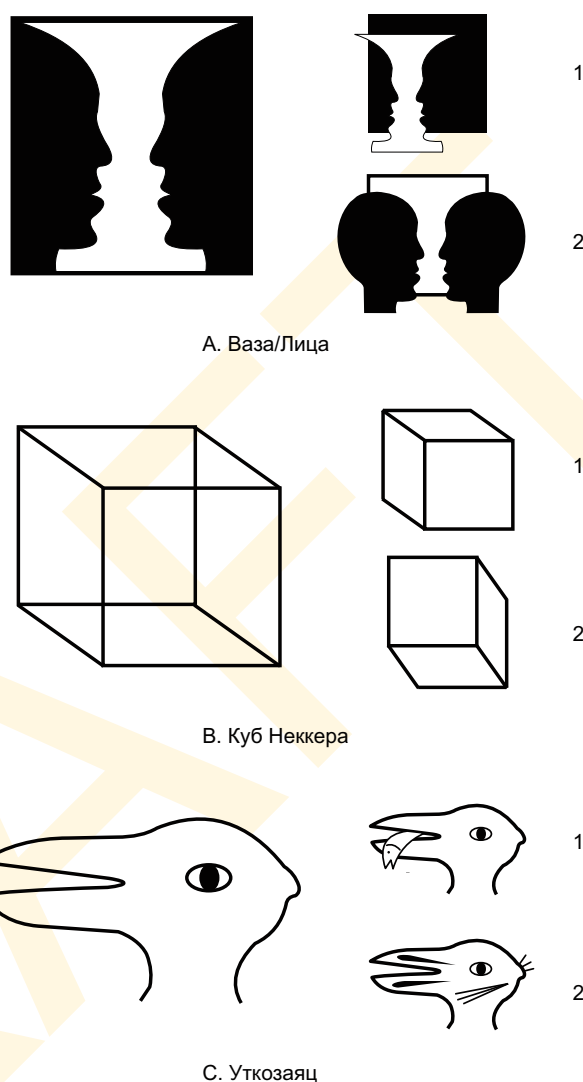
Двойственные фигуры свидетельствуют о конструктивном характере зрительного восприятия, поскольку мы видим, что зрительная система *интерпретирует* визуальную стимуляцию и что возможны более одного варианта этой интерпретации. Если бы восприятие целиком определялось световой стимуляцией глаза, то двойственных фигур не существовало бы, поскольку каждый стимулирующий световой паттерн инициировал бы появление уникального перцепта. Очевидно, что такого не происходит.

При зрительном акте осуществляется нечто более креативное и сложное, выходящее за пределы информации, предоставленной световым стимулом (Bruner, 1973).

Но каким образом зрение умудряется выйти за рамки оптической информации и, главное, зачем? На сей момент мы дадим такой ответ: зрительная система наблюдателя строит *модель* того, какая внешняя ситуация создала на сетчатке глаза паттерн сенсорной стимуляции. Важной и перспективной идеей здесь является то, что человеческое восприятие фактически соответствует *модели*, сконструированной зрительной системой на основе сенсорной стимуляции, но не собственно сенсорной стимуляции (или, по крайней мере, не ей одной). Именно поэтому восприятие может быть иллюзорным и двойственным, несмотря на неиллюзорность и однозначность исходного оптического материала. Отметим также, что иногда на основе оптической информации мы создаем ложные модели, а иногда — две и более равновелико истинных.

Точка зрения, что целью зрительной системы является создание моделей окружающего мира была впервые предложена во второй половине XIX века выдающимся немецким ученым Германом фон Гельмгольцем; который рассматривал зрительное восприятие как процесс построения гипотезы, наиболее адекватной паттерну зрительной стимуляции (Helmholtz, 1867/1925). Такая точка зрения более века была основополагающей в понимании механизмов зрения, хотя впоследствии и была расширена и усовершенствована такими исследователями как Ричард Грегори (1970), Дэвид Марр (1982) и Ирвин Рокк (1983). Направления, по которым эти ученые развивали нашу науку мы будем обсуждать на протяжении всей книги.

Тезис о том, что зрительное восприятие основано на моделировании, должен быть понят нами правильно и ни в коем случае нельзя допустить его неверного толкования. Однако принятие концепции моделирования не означает, что зрительное восприятие суть «чистая фикция». Если бы восприятие было фикцией, то оно не отвечало бы эво-



**Рис. 1.1.5** Двойственные фигуры. А — рисунок может быть истолкован как белая ваза на черном фоне, либо как пара черных лиц на белом. В — куб, который может выглядеть как куб сверху, либо как куб снизу. С — фигура, которую можно истолковать либо уткой, либо зайцем.

люционному требованию точной передачи сведений об окружающем мире. Чтобы удовлетворить данному требованию, перцепционные модели:

- должны быть тесно связаны с информацией, содержащейся в проекционном изображении окружающего мира на сетчатке глаза;

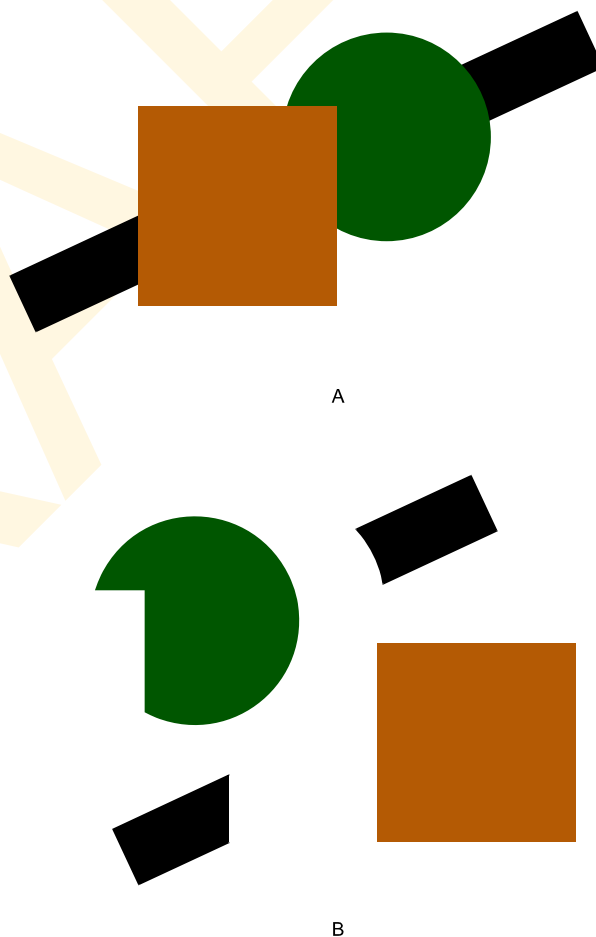
- должны обеспечивать оправданно точную интерпретацию этой информации.

Далее. Факт существования зрительных иллюзий свидетельствует о том, что наши зрительные модели подчас оказываются неточными, а двойственные фигуры, речь о которых шла ранее, доказывают то, что модели эти порой и неоднозначны. Однако и первое, и второе обычно возникает в нетипичных условиях, в частности на страницах научных трудов и в стенах лабораторий. Ежедневные зрительные впечатления свидетельствуют о том, что наши перцепционные модели обычно и точны, и однозначны (уникальны). Безусловно, если сенсорный сигнал достаточно богат и высокоорганизован, то зрительная система практически никогда не ошибается в интерпретации оптической информации об окружающем нас мире (Gibson, 1966).

#### 1.1.4.1 Зрительная достройка

Пожалуй, наиболее ясным и убедительным доказательством того, что зрительное восприятие конструирует модели окружающего мира является тот факт, что наше зрение достраивает невидимые фрагменты поверхностей. Давайте посмотрим на рис. 1.1.6 А: не вызывает сомнения, что на нем изображены три простейших фигуры — круг, квадрат и вытянутый прямоугольник. А теперь задумаемся, как данное нами словесное описание соотносится с тем, что фактически представлено в изображении?: круг частично прикрыт квадратом и его нижняя часть не видна; у прямоугольника видны лишь концы, а середина закрыта кругом и квадратом. Тем не менее, фрагмент круга мы воспринимаем как цельный круг, а концы прямоугольника как фрагменты цельной фигуры. Если же сказанное вызывает сомнение, то можно сравнить результат восприятия с таковым от рис. 1.1.6 В, на котором предъявлены в точности те же участки фигур, но во взаиморасположении, не позволяющем зрительной системе восполнить недостающее.

Перцепционное заполнение скрытых фрагментов именуется «*зрительной достройкой*». Сие происходит всегда, «автоматически» и без каких-ли-



**Рис. 1.1.6** Зрительная достройка частично закрытых предметов. Рис. А выглядит как композиция из полноценных квадрата, круга и прямоугольника, даже несмотря на то, что собственно видимыми являются лишь фрагменты этих фигур (на рис. В).



бо затруднений. Поглядим вокруг и убедимся в том, как много в поле зрения предметов, зрительное восприятие которых основано на достройке фактически невидимых участков: практически ни один предмет из окружающей обстановки не видим полностью, но при этом почти все они воспринимаются цельно.

Отметим, что зрительное восприятие умеет извлекать знание, о т.н. «самозакрытых поверхностях» — тех участках объекта, что полностью скрыты от наблюдателя его видимыми поверхностями. К примеру, на рис. 1.1.7А непосредственно видна лишь половина куба. Наше восприятие пытается тем или иным способом включить в результирующий перцепт невидимые грани этой фигуры, т.е. закрытые ее видимыми гранями. Мы сильно удивимся, если сменив точку наблюдения обнаружим, что на противоположной стороне куб таков, как показано на рис. 1.1.7В: действительно, существует бесконечное множество возможных физических ситуаций, подпадающих под ситуацию на рис. 1.1.7А и воспринимаемых единственно возможным образом — как полноценный куб.

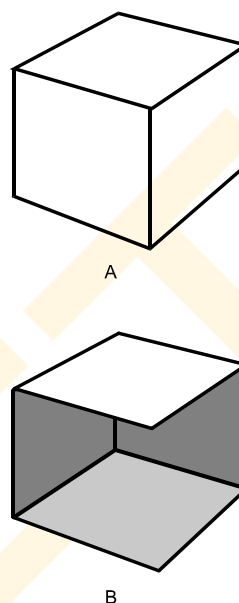
Достройка представляет собой пожалуй даже более эффектный пример креативного характера зрительного восприятия, чем зрительные иллюзии и двойственные фигуры. То есть, речь идет о чем-то, что выходит за рамки обычной оптической информации. Действительно, всегда высока вероятность того, что самозакрытые и частично закрытые поверхности объектов схожи с открытыми, но восприятие их должно строиться из нечто большего, чем свет, попавший в глаз, поскольку оптическое изображение на задней поверхности глаза не несет всей полноты визуальной информации об объекте.

#### 1.1.4.2 Невозможные фигуры

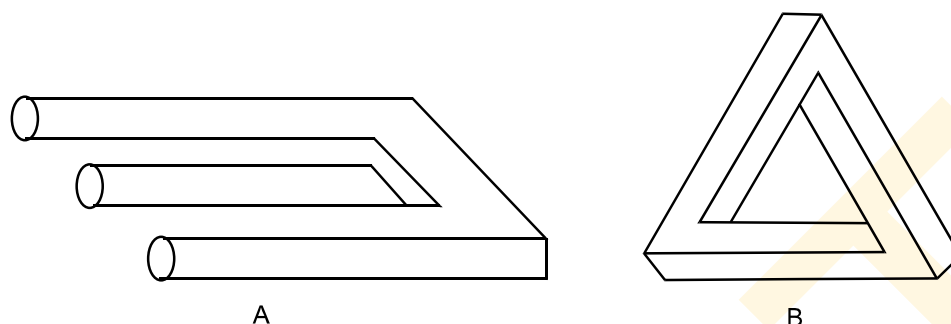
Существует еще один феномен, наглядно демонстрирующий то, что зрительное восприятие носит моделирующий характер — невозможные фигуры, представляющие собой двухмерные штриховые рисунки, воспринимаемые как изображения объемных трехмерных предметов, но которые при этом физически невозможны.

На рис. 1.1.8 даны несколько широко известных примеров: «посейдонка» (рис. 1.1.8А), которая на первый взгляд выглядит вполне реальным предметом, но стоит всмотреться, как становится ясно, что такой объект существовать не может, поскольку трезубец (слева) не согласуется со своим собственным основанием (справа). То же касается и непрерывного трехмерного треугольника (рис. 1.1.8В), который вначале воспринимается реальным, но на самом деле существовать не может, поскольку поверхности его сторон не согласуются друг с другом (Penrose & Penrose, 1958).

Невозможные фигуры интересны тем, что эффектно иллюстрируют тот факт, что наше зрительное восприятие — суть внутреннее конструирование *предполагаемой* внешней реальности. Если бы зрительное восприятие было безошибочным отражени-



**Рис. 1.1.7** Зрительная достройка самозакрытых поверхностей: фигура на рис. А однозначно воспринимается как куб, хотя физически на обратной стороне этот «куб» может оказаться таким, как на рис. В.



**Рис. 1.1.8** Невозможные фигуры. В первый момент обе фигуры кажутся вполне реальными, но через некоторое время становится ясно, что они физически невозможны. Сие лишний раз подтверждает ту мысль, что зрение не просто регистрирует действительность, но активно строит модели окружающего мира.

ем окружающего мира, физически невозможные фигуры просто не воспринимались бы, т.е. их не существовало бы не только физически, но и перцепционно. Однако же человеческое зрение с готовностью воспринимает (моделирует) такие объекты, когда мы смотрим на умело изготовленные обманки. Сей факт указывает на то, что зрительное восприятие организовано как 3D-интерпретация двухмерной оптической информации регистрируемой глазом. Более того, характер ошибок, которые возникают при восприятии изображений невозможных фигур, указывает как минимум на то, что некоторые зрительные процессы исходно работают на локальном уровне и только на более поздних этапах зрительного акта встраивают результат в общую структуру: отдельные части предметов, показанных на рис. 1.1.8, вполне реальны; но стоит только попытаться собрать их в единую картину, как несовместимость и абсурдность станут очевидными.

#### 1.1.4.3 Зрительное предвидение

Итак, мы начинаем понимать, что зрительная система человека нацелена скорее на конструирование гипотетических моделей окружающего мира, нежели на прямое извлечение информации из сенсорной стимуляции. Соответственно возникает, вопрос: почему так устроено? В первом приближении ответ таков: с эволюционной точки зрения модели практичнее, нежели изображения, с которых эти модели были построены. Почему так происходит, не совсем ясно. Однако практическая польза зрительной достройки вполне вероятно объясняется тем, что 3D-модели объектов, включающие в себя в том числе и скрытые поверхности, несут много более полную информацию об окружающем мире, нежели чисто стимульная перцепция. Дополнительная информация, которую создает модель, весьма полезна, поскольку помогает особи *предвидеть будущее*: вспомним рис. 1.1.7 — восприятие фигуры на этом рисунке как трехмерного объекта — куба, закладывает основу нашего *предположения* о том, что мы увидим, если будем перемещаться в пространстве и когда будут открываться новые, ранее невидимые поверхности интересующего нас объекта. Сие критично в жизни подвижных существ, в частности, людей. Плюс к тому, стабильная трехмерная модель избавляет нас от необходимости повторного восприятия пустого места по мере нашего перемещения в пространстве.

Перцепционная модель трехмерного мира не требует серьезной модификации, по-

сколько при нашем движении меняется только один параметр — точка наблюдения (относительно стабильной картины объектов и поверхностей). Фактически, модель лишь один раз требует серьезной переработки — когда построенные ею прогнозы не подтвердились из-за того, что в поле зрения появились поверхности непредвиденного характера. Однако же житейский опыт сообщает нам, что такое случается ощутимо реже, чем устойчивое подтверждение зрительных прогнозов.

Таким образом, хотя создание трехмерной модели окружающей обстановки на первый взгляд может показаться весьма хилой эволюционной стратегией, ее краткосрочные издержки с лихвой компенсируются долгосрочными выгодами. Безусловно, требуется достаточно много времени и усилий, чтобы построить адекватную модель, но когда таковое выполнено, то уже не нужно особых затрат на ее поддержание. Итак, построение модели — это самое экономичное решение проблемы стабилизации процесса зрительного познания окружающей обстановки и обеспечения точности его результатов.

Способность предвидеть результаты будущего восприятия также имеет колоссальное эволюционное значение, поскольку мы живем в мире, наполненном движущимися объектами. Безусловно, знать текущую позицию движущегося предмета необходимо, но гораздо полезнее знать *скорость* и *направление* его движения, дабы иметь возможность предсказать будущую локализацию объекта, его потенциальную траекторию. Сие особо важно, когда кто-то приближается к нам, бо нужно оперативно принимать решение: подойти, отступить, удрать или просто проигнорировать. Без перцепционной модели, молниеносно выходящей за куцые рамки стимульной информации, зрение не могло бы адекватно руководить нашими действиями.

Точка зрения, что целью мозговой работы является динамический расчет т.н. *предиктивных (прогностических) моделей* окружающего мира, была высказана британским психологом Кеннетом Крайком в 1943 г: ученый блестяще показал, что организмы, способные быстро экстраполировать текущую ситуацию в будущее, имеют эволюционное преимущество над остальными. Особь, которая способна к точным прогнозам, может *планировать* свои действия, тогда как не способная — лишь *реагировать* на происходящее.

Стоит отметить, что польза от экстраполяции есть только тогда, когда последняя срабатывает быстрее, чем наступает прогнозируемое событие, и неудивительно, что большинство перцепционных прогнозов строятся зрительной системой с фантастической скоростью. Действительно, модели образуются столь стремительно, что в отношении них у нас нет т.н. *осознанного опыта* (если, конечно, не возникло прогностической ошибки). Но даже в случае сильной погрешности наш осознанный опыт отражает именно ошибку, нежели собственно прогноз.

### 1.1.5 Восприятие как выявление смысла

Перцепционные конструкторы внешнего мира идут даже дальше, чем просто трехмерная достройка невидимых поверхностей — они несут информацию о смысле объектов и ситуаций, т.е. об их *функциональной значимости*. Мы воспринимаем объект не только как имеющий специфические форму и локализацию, но персонифицированно: собака, дом и т.п. Иными словами, у нас имеется способность к *классификации (распознаванию, идентификации)* объектов окружающего мира, как принадлежащих к определенной категории. Сие открывает нам доступ к хранилищу огромного количества



информации, накопленной из опыта прежнего общения с подобными объектами (и что в конечном итоге позволяет нам реагировать на них оптимальным образом).

### 1.1.5.1 Классификация

Пожалуй, самым простым способом оценить важность классификации — это вообразить совершенно неизвестный предмет. Мы можем воспринять его физические характеристики, такие как окраска, текстура, размер, форма и положение в пространстве, но мы не знаем, что это такое и что с ним делать: он живой? съедобный? опасный? можно ли подойти к нему? не лучше ли сбежать от него? Ответы на эти вопросы редко когда можно получить из одних лишь физических характеристик предмета, и требуется отнести его к той или иной группе. Мы обнимаем любимых, убегаем от злых собак, прогуливаемся по паркам, едим яблоки и сидим в креслах. Все это столь очевидно, что едва ли осознается, но без перцепционной классификации этих вещей на категории будет очень трудно вести себя адекватно огромному количеству новых предметов, которые встречаются нам ежедневно и ежечасно. Мы спокойно и уверенно прогуливаемся по аллеям незнакомых парков потому, что прошлый опыт информирует нас: деревья как правило неподвижны. А вот злая собака подвижна и еще как!

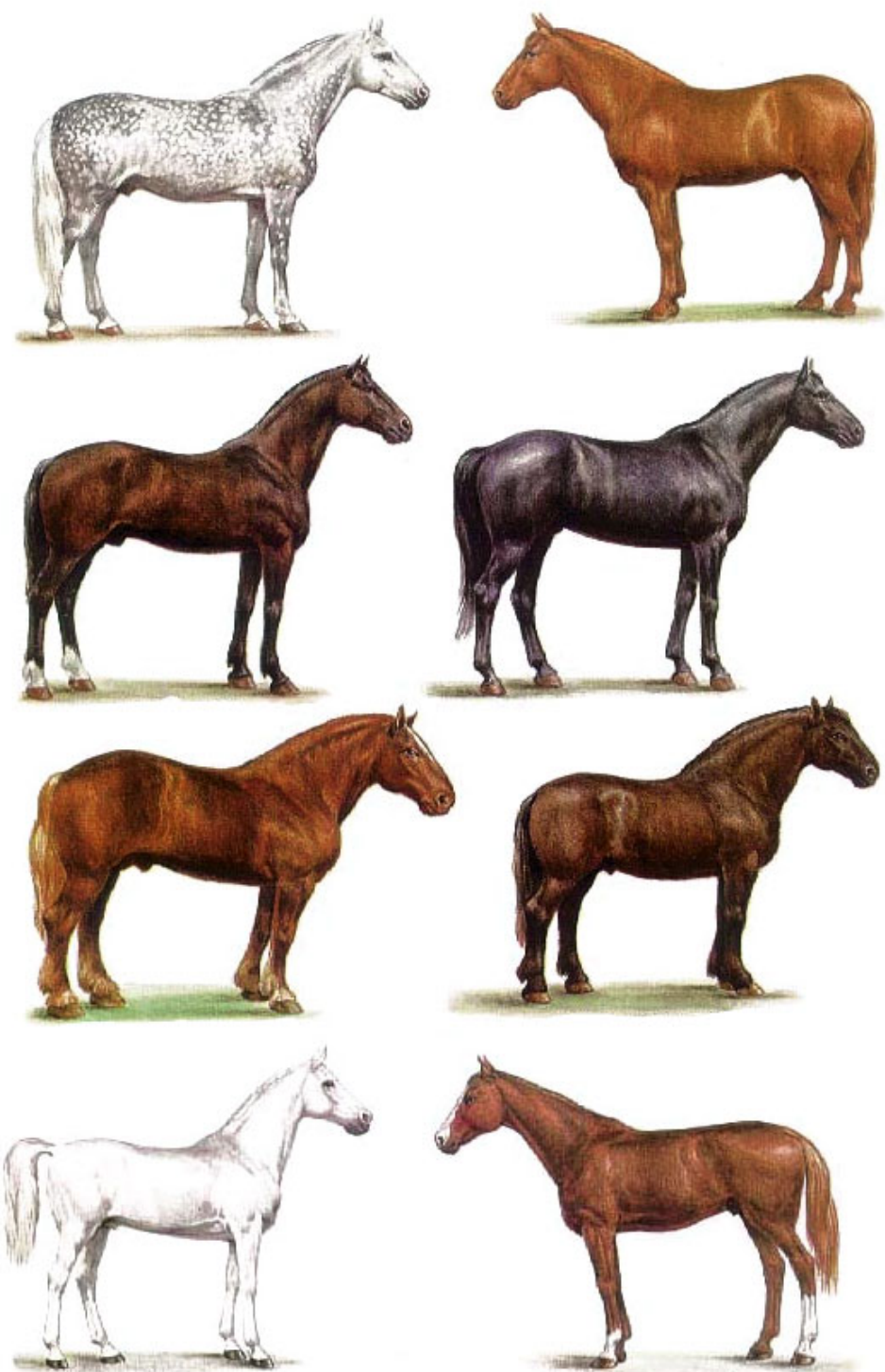
Итак, классификация полезна, поскольку объекты одной категории обладают множеством общих свойств и «повадок»: все кресла несколько разные, все яблоки немного разные, но одно кресло гораздо более походит на другое кресло, нежели на яблоко (и наоборот). Следовательно, прежний опыт общения с представителем той или иной категории предметов позволяет нам с необходимой и достаточной точностью прогнозировать, как поведет себя представитель данной категории. То есть, мы можем обходиться с большинством новых объектов на абстрактном уровне их категории, даже если ранее никак не встречались с данным конкретным предметом.

Зрительная сортировка объектов на известные категории, казалось бы, проста, но фактически это гигантское достижение эволюции. К примеру, взглянем на разнообразие пород лошадей, представленных на рис. 1.1.9: каким образом мы моментально распознаем то, что все это лошади? Сие есть проблема идентификации объектов — одна из наиболее сложных и до сих пор не решенных — сверхзагадка зрительного восприятия. В главах 8 и 9 мы рассмотрим несколько современных гипотез того, как происходит зрительная классификация.

### 1.1.5.2 Внимание и сознание

Очевидно, что внешний мир несет в себе гораздо больше информации, чем та, что может быть воспринята кем-либо. Следовательно, мы должны *выбирать*, на что обращать внимание, на что нет, и наш выбор будет зависеть по большей части от наших потребностей, целей, планов и желаний. К примеру, бутерброд так и останется бутербродом, но то, как мыотреагируем на него, зависит от нашего состояния в данный момент: если, к примеру, сутки голодали, то еда моментально привлечет наше внимание; если же после сытного обеда, то, вероятнее всего, не обратим на хлеб с маслом особого внимания (а если и обратим, то почувствуем легкую тошноту).

Приведенный пример демонстрирует то, что восприятие, зрительное в частности — это не только и не столько стимуловедомый процесс. Иными словами, восприятие определяется не только и не столько оптической информацией, зашитой в зрительный стимул, сколько на него воздействуют и т.н. *когнитивные факторы* — высокоуровневые цели, планы и ожидания. Было бы по меньшей мере странно, если бы про-



**Рис. 1.1.9** Породы лошадей. Дабы отнести тот или иной объект к известной категории, зрение выходит далеко за рамки физического описания этого объекта. Несмотря на резко выраженные физические отличия в облике, все животные на данном постере очевидно принадлежат к виду домашних лошадей.

исходило иначе, поскольку **конечной эволюционной целью восприятия является, как уже было показано, — создание связи между потребностями организма и соответствующими возможностями их удовлетворения за счет объектов внешнего мира.** Существует несметное множество путей, по которым высокоуровневые когнитивные факторы влияют на результаты нашего восприятия, и многие из которых задействуют селективный процесс зрительного внимания. Пример с бутербродом иллюстрирует то, что разные вещи в нашем окружении мы воспринимаем (или не воспринимаем) в зависимости от того, чего хотим достичь. Сие исключительно важно, хоть и, казалось бы, самоочевидно.

Одной из функций зрительного внимания является получение зрительной информации для последующего осознания, но точные свойства объектов не войдут в осознанный опыт, если не являются для особи приоритетными. Неприоритетные объекты чаще всего обрабатываются зрением вне сознания наблюдателя, пока они не привлекут его внимания. Примеров множество. Скажем, мы можем не обращать внимания на некий неподвижный объект, находящийся на периферии поля зрения, но если он вдруг неожиданно начнет приближаться, то мы немедленно посмотрим на него, сами не зная отчего, и лишь тогда осознаем факт его существования. Когда мы управляем автомобилем, мы иногда начинаем пристально всматриваться во впереди идущую машину лишь от того, что ее водитель взглянул на нас. В обоих случаях зрительная обработка происходит вне сознания, обращая наше внимание лишь на особо интересные или важные аспекты окружения: движущийся объект или человека, глядящего в нашу сторону. Если же объект привлек наше внимание, мы начинаем осознавать его свойства, получая тем самым возможность идентифицировать этот объект и оценить его значимость в текущей ситуации.

Устремляя зрительное внимание на объект, мы обычно перемещаем глазные яблоки и фиксируем взгляд на нем, но следует понимать, что зрительное внимание и зрительная фиксация — не одно и то же. Все мы наверняка сталкивались с тем, что иногда пристально смотрим на какой-нибудь предмет, не обращая при этом на него ни малейшего внимания: мысли наши бродят где-то далеко, но если вдруг наше внимание обратится к поступающей зрительной информации, мы можем неожиданно осознать, что понятия не имеем о предмете, присутствовавшем в поле зрения в период наших раздумий. И наоборот: мы можем сосредоточить свое внимание на объекте, не фиксируя взгляда на нем.

Многие высокоуровневые аспекты восприятия вероятно полностью подотчетны сознанию, к примеру: когда мы осматриваем комнату пытаясь найти ключи от квартиры, мы ясно осознаем свою цель (найти ключи), которая и направляет наше внимание на разные «ключеподобные» предметы.

Иные аспекты восприятия очевидно не осознаваемы, причем даже в сходных ситуациях, к примеру, похожий на ключ предмет в процессе поиска обязательно притянет к себе взгляд. В целом, низкоуровневое восприятие либо недоступно нашему осознанному знанию и опыту, либо поступает туда видоизмененным.

Роль сознания в восприятии до конца не изучена, и, к сожалению, мы очень мало знаем об эволюционных преимуществах осознанного восприятия. Однако имеется твердая уверенность в том, что такие преимущества есть, но пока еще никто не смог четко показать, каковы они. Иными словами, остается невыясненным вопрос: в чем же преимущества осознанно воспринимающей особи над особью, которая решает те же перцепционные задачи, но не имеет при этом осознанного зрительного опыта? Лишенный зрительного осознания «овощ» эволюционно может быть вполне успешен: эффек-

тивно находить еду, убежище, пару, избегать обрывов, уворачиваться от хищников и падающих предметов. Так что пока остается неясным, каков эволюционный мотив осознанного зрительного восприятия. Однако остается немалый шанс, что приведенные контраргументы некорректны или по крайней мере неполны и можно предположить, что гипотетический биоболван не может эффективно справляться со всеми без исключения задачами, с которыми успешно справляется осознанно воспринимающий организм. Возможно, сознание играет критичную, но пока еще не понятую роль в наших перцепционных возможностях. Мы вернемся к этому разговору в 13-й главе, когда будем рассуждать о том, что известно сегодня о взаимоотношениях между сознанием и восприятием.

## 1.2 ОПТИЧЕСКАЯ ИНФОРМАЦИЯ

Данное нами определение зрительного восприятия предполагает, что зрительный акт строится на взаимодействии трех сущностей: света, поверхностей (отражающих свет) и зрительной системы наблюдателя. Если удалить какой-либо один из этих компонентов — зрительный акт невозможен. Таким образом, имеет смысл начинать изучение зрения с рассмотрения основных положений по каждой из трех озвученных позиций.

В текущем разделе будет описано то, как свет взаимодействует с поверхностями, производя т.н. *оптические события*, являющиеся отправной точкой зрения.

В следующем разделе мы опишем общую структуру зрительной системы человека, которая обрабатывает информацию, содержащуюся в оптических событиях.

В оставшейся части книги мы будем детально разбирать то, как зрительная система извлекает полезную информацию из световых потоков, производя биологически ценные перцепты реальных сцен, объектов и ситуаций.

Итак, в предыдущих разделах было показано, что эволюционная роль зрительного восприятия состоит в обеспечении организма высокоточной информацией об окружающем мире, и свет, проникающий в глаза, должен нести какую-то информацию об окружающей обстановке. Однако в извлечении всей имеющейся оптической информации нет необходимости, но требуемый минимум должен быть достаточным для того, чтобы быть истолкованным зрительной системой с оправданной точностью.

В текущем разделе мы рассмотрим то, как свет формирует и переносит информацию о предметах окружающего мира, но прежде поговорим об элементарных понятиях из сферы радио- и фотометрии.

### 1.2.1 Радиометрия и фотометрия

*Радиометрия* — это методологическая система измерений свойств электромагнитных излучений, которая *не* учитывает факт и характер воздействия этих излучений на зрительную систему человека.

*Фотометрия* — такая же система, но построенная с учетом воздействия электромагнитного излучения на человеческое зрение.

Базисом фотометрии является т.н. *функция фотопической световой эффективности*,  $V(\lambda)$ , которая показывает, излучение каких длин волн видимого спектра больше воздействует на зрительную систему, каких меньше (рис. 1.2.1). Однако следует особо отметить то, что хотя кривая  $V(\lambda)$  внешне похожа на кривую фотопической спектральной чувствительности зрения (см. рис. 1.3.18 В) — она не тождественна ей по смыслу.



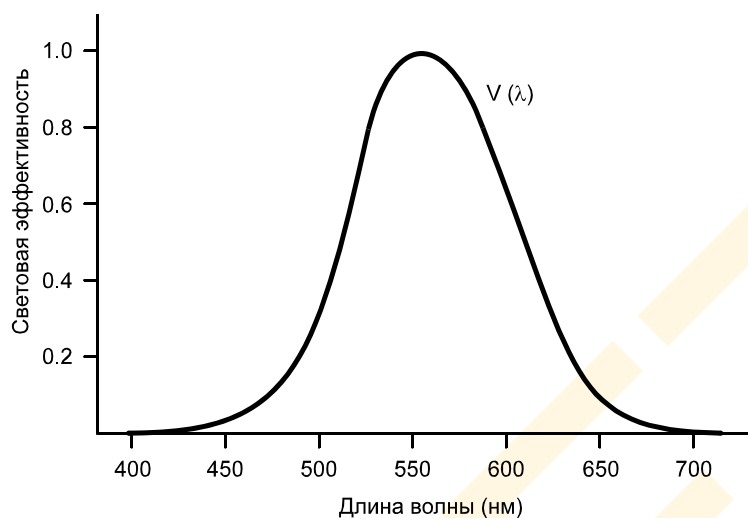


Рис. 1.2.1 Функция фотопической световой эффективности  $V(\lambda)$ .

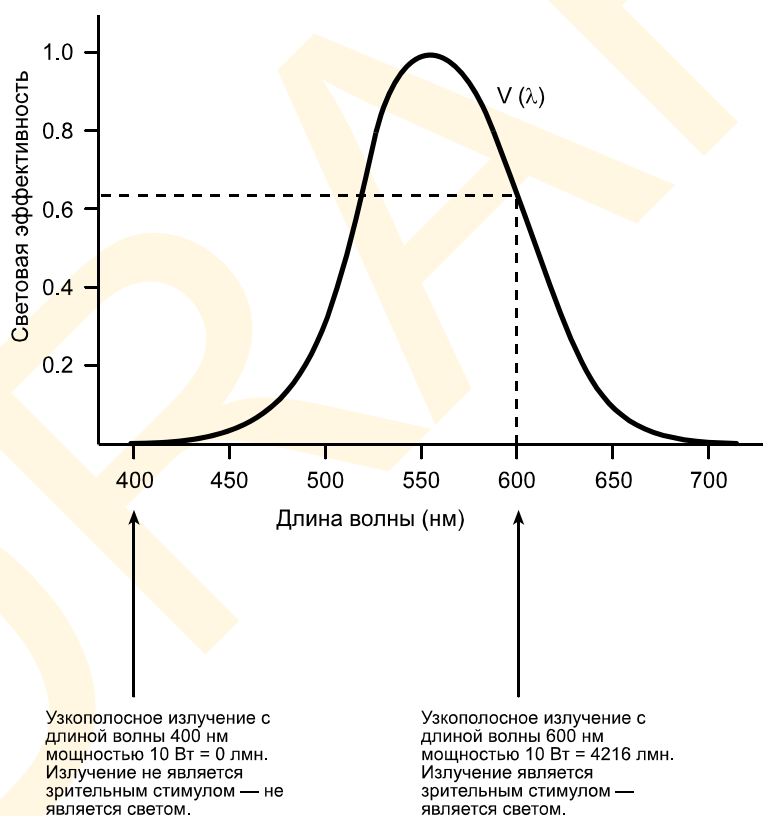


Рис. 1.2.2 Узкоспектральное излучение мощностью в 10 Вт при длине волны обладает нулевым воздействием на зрительную систему человека и не является светом; узкоспектральное излучение той же мощности, но другой длины волны (600 нм) является светом, потому что оказывает воздействие на зрительную систему человека (световая эффективность 0.62). То есть, световая эффективность — это мера общего (недифференцированного) воздействия излучения на зрение.

Рассмотрим пример. На рис. 1.2.2 показаны спектральные распределения энергии двух монохроматических излучений (монохроматических стимулов): одно с длиной волны 400 нм, другое — 600 нм. Мощность каждого 10 Вт. При этом по графику мы видим, что световая эффективность монохромата с длиной волны в 400 нм равна нулю, а монохромата в 600 нм — 0.62. Радиометрически стимулы эквивалентны друг другу по мощности, но фотометрически отличны: монохромат в 400 нм не оказывает на зрение практически никакого воздействия (не является *светом*), тогда как стимул в 600 нм — оказывает световое воздействие и воспринимается зрительной системой человека.

### 1.2.1.1 Световой поток

Основной фотометрической единицей, выражающей *световой поток*, является люмен (лмн). Пик кривой  $V(\lambda)$  приходится на 555 нм и соответствует 680 лмн/Вт. Прочие участки видимого спектра менее эффективны и дают меньше количество люмен на ватт, к примеру: световая эффективность монохромата с длиной волны в 650 нм (т.е. мера его общего воздействия на зрительную систему человека) составляет 0.1 или  $0.1 \times 680 = 68$  лмн/Вт.

Вернемся к первому примеру, в котором речь шла о двух монохроматических излучениях с длинами волн 400 нм и 600 нм. То, что световая эффективность монохромата в 400 нм равна нулю означает, что его световой поток тоже равен нулю, но при этом поток его излучения не равен нулю. Световой поток монохромата с длиной волны в 600 нм равен  $0.62 \times (680 \text{ лмн/Вт}) \times 10 \text{ Вт} = 4216$  лмн. Несмотря на то, что мощность излучения этого стимула ровно та же, что и у монохромата с длиной волны 400 нм, он представляет собой определенный световой поток, т.е. в определенной степени воздействует на зрительную систему человека. Иными словами, монохромат в 400 нм фактически не является стимулом, но лишь электромагнитным излучением, тогда как монохроматический стимул в 600 нм является и электромагнитным излучением, и световым (зрительным) стимулом с определенной степенью воздействия на зрительную систему.

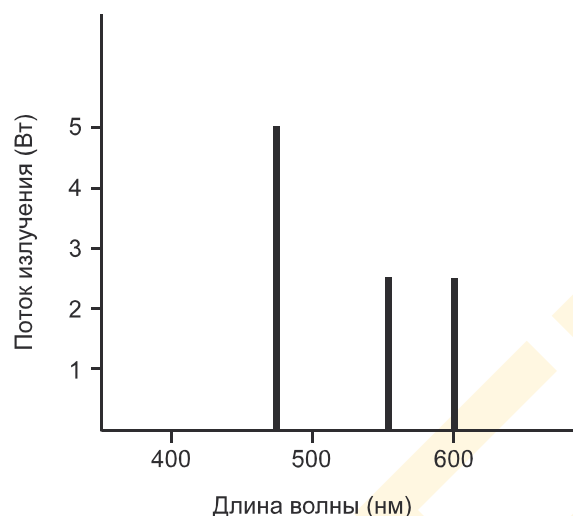
Параметр светового потока характеризует общее количество света, испускаемое источником по всем направлениям.

### 1.2.1.2 Закон аддитивности Эбнея

Большинство объектов отражают (или испускают) электромагнитную энергию, представляющую собой смесь монохроматических стимулов с разными длинами волн. Такую смесь принято называть *полихроматической*. Результирующую силу света полихроматических стимулов можно вычислить суммированием величин светового потока каждой из монохроматических составляющих. К примеру, по стимулу, показанному на рис. 1.2.3 процедура будет выглядеть следующим образом:

5 Вт на 475 нм	$0.1 \times (680 \text{ лмн/Вт}) \times 5 \text{ Вт} = 340 \text{ лмн}$
2.5 Вт на 555 нм	$1.0 \times (680 \text{ лмн/Вт}) \times 2.5 \text{ Вт} = 1700 \text{ лмн}$
2.5 Вт на 600 нм	$1.0 \times (680 \text{ лмн/Вт}) \times 2.5 \text{ Вт} = 1054 \text{ лмн}$
Итого:	$340 \text{ лмн} + 1700 \text{ лмн} + 1054 \text{ лмн} = 3094 \text{ лмн}$

Отметим, что мы имеем полное право складывать значения по монохроматам, пото-



**Рис. 1.2.3** Спектральное распределение энергии искусственно созданного стимула, состоящего из трех спектральных полос с длиной волны 475 нм и мощности излучения в 5 Вт, длиной волны 555 нм и мощности излучения в 2.5 Вт, длиной волны 600 нм и мощности излучения в 2.5 Вт.

му что фотометрические системы обладают свойством *аддитивности* (слагаемости), получившим название *закона аддитивности Эбнея* или *эбнеевского закона аддитивности* (Abney, 1913).

### 1.2.1.3 Сила света

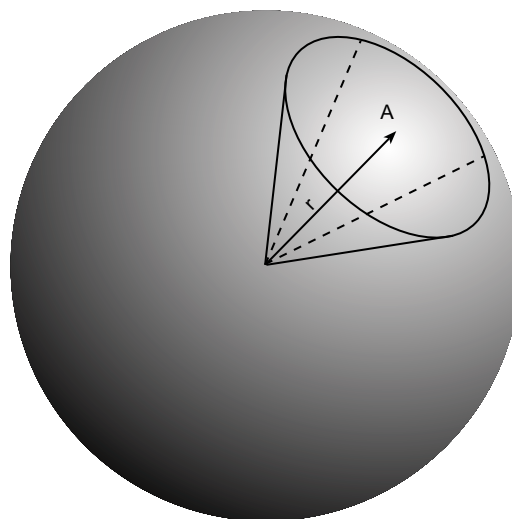
Если световой поток — это мера ненаправленного света, то такой показатель как *сила света* характеризует направленный световой поток, то есть количество люмен в выбранном направлении от точечного источника (см. рис. 1.2.4).

Единицей силы света является кандела (cd), которая определена как люмен на стерadian, то есть световой поток в единице телесного угла ( $\omega$  или Sr) — конической секции сферы (рис. 1.2.4):

$$\omega = A/r^2,$$

где:  $r$  — радиус сферы;  $A$  — площадь поверхности, вырезаемой конусом из этой сферы.

Исходя из геометрии сферы, силу света в 1 cd даст источник, световой поток которого по всем направлениям равен  $4\pi$  люмен. Следовательно, равномерный точечный источник света в 10 cd дает общий световой поток  $(10 \text{ cd})(4\pi \text{ лмн/сд}) = 127 \text{ лмн}$ .



**Рис. 1.2.4.** Стерadian (телесный угол) определен как  $A/r^2$ .

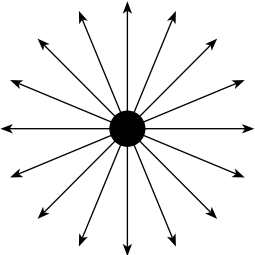
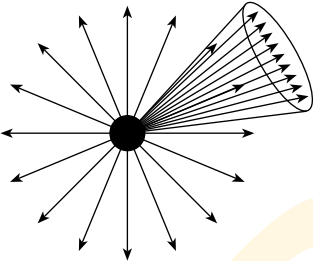
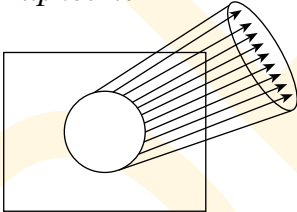
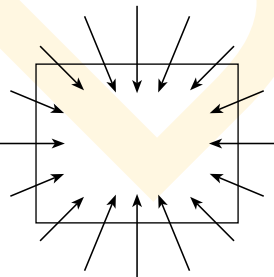
	Определение	Фотометрические единицы	Радиометрические единицы
<p><i>Световой поток</i></p> 	Общее количество световой энергии, испускаемой источником.	лмн	Поток излучения (Дж/сек или Вт)
<p><i>Сила света</i></p> 	Световая энергия от точечного источника на единицу телесного угла	лмн/Sr или cd (1 лмн/1 Sr = 1cd)	Сила излучения (Вт/Sr)
<p><i>Фотометрическая яркость</i></p> 	Сила света на единицу площади поверхности	cd/м <sup>2</sup> или fL	Энергетическая яркость (Вт/Sr/м <sup>2</sup> )
<p><i>Освещенность</i></p> 	Световая энергия, падающая на поверхность	лмн/м <sup>2</sup> или лмн/ft <sup>2</sup>	Облученность (Вт/м <sup>2</sup> )

Рис. 1.2.5 Четыре оптических понятия в фотометрическом и радиометрическом толкованиях.



### 1.2.1.4 Фотометрическая яркость

Фотометрическая яркость — это величина, характеризующая количество света, исходящего от поверхности (к примеру, листа бумаги) в определенном направлении (см. рис. 1.2.5). Фотометрическая яркость измеряется в канделах на площадь проекции источника, а общепринятыми единицами фотометрической яркости являются кандела на метр квадратный ( $\text{cd}/\text{m}^2$ ), либо фут-ламберт. Перцепционным (чувственным) коррелятом фотометрической яркости стимула является его *яркость субъективная*.

Будут ли фотометрическая яркость (физическая величина) и яркость субъективная (перцепционный показатель) оставаться неизменными по мере удаления от источника света? По мере увеличения расстояния от наблюдателя до источника света площадь оптического изображения на задней поверхности глаза будет уменьшаться пропорционально силе света; однако поскольку соотношение силы света к площади изображения на задней поверхности глаза будет оставаться постоянным, яркость фотометрическая и яркость субъективная не изменятся.

### 1.2.1.5 Освещенность

Понятие освещенность характеризует световой поток, падающий *на* поверхность. Общепринятыми единицами являются люмен на метр квадратный — *люкс*, а также люмен на квадратный фут — *фут-кандела*.

Освещенность можно уподобить дождю: чем выше плотность капель дождя (люмены), тем сильнее дождь (тем выше освещенность). Сила дождя, как известно, не зависит от свойств и размеров поверхности на которую он изливается, что в той же мере касается и освещенности — она никак не зависит от свойств и площади поверхности, на которую падает свет.

Ниже мы приведем несколько типичных цифр фотометрической яркости источников света и освещенности поверхностей, имеющих место в наиболее типичных ситуациях (см. табл. 1.2.1).

ТАБЛИЦА 1.2.1 Некоторые значения освещенности и соответствующие им величины фотометрической яркости

	ОСВЕЩЕННОСТЬ (lux)			Ф/М ЯРКОСТЬ* ( $\text{cd}/\text{m}^2$ )		
Яркое солнце	50 000	—	100 000	3000	—	6000
Солнце в дымке	25 000	—	50 000	1500	—	3000
Высокие облака	10 000	—	25 000	600	—	1500
Пасмурно	2000	—	10 000	120	—	600
Небо в густых низких тучах	100	—	2000	6	—	120
Рассвет/закат	1	—	100	0.06	—	6
Полная луна	0.01	—	0.1	0.0006	—	0.006
Звезды	0.0001	—	0.001	$6 \times 10^{-6}$	—	0.000 06
Операционная	5000	—	10 000	300	—	600

Витрины	1000	—	5000	60	—	300
Конструкторское бюро	300	—	500	18	—	30
Офис	200	—	300	12	—	18
Жилое помещение	50	—	200	3	—	12
Коридоры	50	—	100	3	—	6
Хорошо освещенная улица		20			1.2	
Плохо освещенная улица		0.1			0.006	

\*При коэффициенте яркости 20% (средний коэффициент отражения типичных сцен).

Отметим, что довольно легко перепутать фотометрическую яркость с освещенностью и наоборот, поэтому давайте еще раз проговорим то, что фотометрическая яркость — это свойство определенным образом освещенной (или самосветящейся) поверхности и характеризуется это свойство количеством *исходящего* от нее света; освещенность же характеризует количество света, *падающего* на поверхность. При определенной освещенности стола разные предметы на нем будут, отражая падающий свет, создавать зрительные стимулы разной фотометрической яркости; к примеру, вероятнее всего лист бумаги на столе даст стимул наибольшей фотометрической яркости (и, следовательно, наибольшей яркости субъективной).

### 1.2.1.6 Радиометрические единицы

Радиометрическими эквивалентами светового потока, силы света, фотометрической яркости и освещенности являются соответственно *поток излучения, сила излучения, энергетическая яркость и облученность* (см. рис. 1.2.5).

Фундаментально важным отличием радиометрии от фотометрии является то, что радиометрия оценивает излучение сугубо как излучение, т.е. без учета его воздействия на зрительную систему человека. Фотометрия же — это система измерений свойств излучения в контексте его воздействия на зрительную систему человека: если некий поток лучистой энергии не оказывает воздействия на зрение, то он не попадает в сферу интересов фотометрии и не существует для нее.

### 1.2.2 Свойства света

Наука, изучающая свойства света, именуется оптикой и является разделом физики. Согласно принятой ныне физической теории электромагнитное излучение состоит из мелких пакетов энергии — *квантов*, которые с одной стороны ведут себя как волны, с другой — как частицы. Кванты, которые способны оказать воздействие на зрительную систему человека (и вызвать ее отклик), именуются *фотонами*. Поток фотонов — суть свет. В нашей книге мы сделаем акцент на корпускулярном характере фотонов, однако разговор о цветовом зрении (гл. 3) потребует рассуждений в контексте их волновых свойств.

Фотоны, испущенные источником, например, нагретым телом (таким как Солнце, огонь или вольфрамовая нить лампы накаливания) — это бесконечно мелкие частицы, пронизывающие пространство по идеально прямым траекториям и со скоростью по-

рядка 300000 км/сек. Когда фотоны бомбардируют поверхность объекта, мы говорим, что эти объекты освещены. Количество видимого света (то есть количество фотонов), падающего на данную поверхность в единицу времени именуется *фотометрической яркостью*, определение которой мы дали выше и которая по определенной зависимости коррелирует с яркостью воспринимаемой, *субъективной*; однако взаимоотношения эти довольно сложны, о чем подробнее речь пойдет в гл. 3.

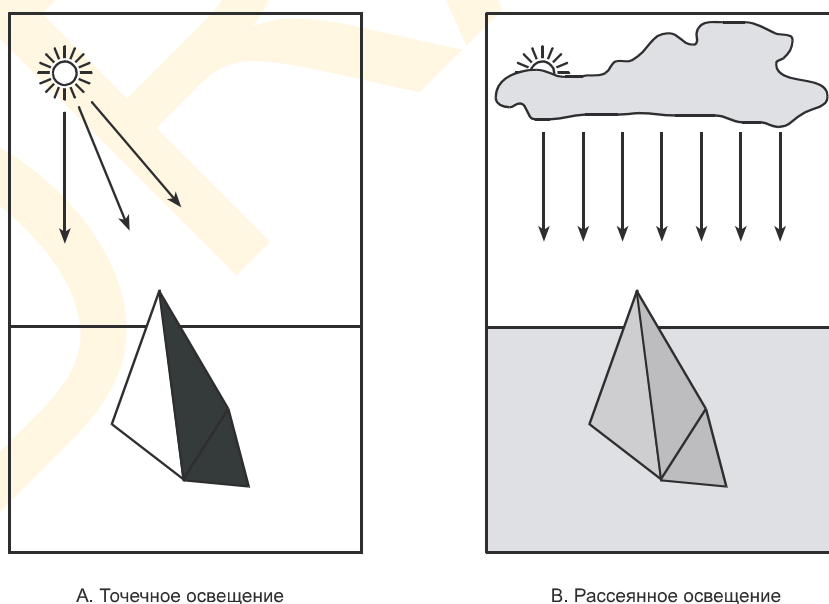
### 1.2.2.1 Освещение

Понятие «освещение» относится к световой обстановке в сцене. С аналитической точки зрения простейшей является ситуация, именуемая «освещением точечным источником». Сие идеализированная ситуация, в которой весь свет сцены берется из одной точки в пространстве. Одиночная лампа накаливания в темном помещении или солнце в ясный день — это весьма близкие аппроксимации освещения сцены точечным источником.

При точечном освещении предметы отбрасывают темные, четко обозначенные тени, а освещенные поверхности имеют жесткие полутона (см. рис. 1.2.6А). Оба эффекта являются следствием того факта, что прямой (неотраженный) свет приходит из одного места. Такие ситуации крайне редки в реальном мире. Точечное освещение интересно преимущественно исследователям зрения, поскольку дает возможность математических упрощений, к примеру, точечное освещение используется как упрощающее допущение при определении формы объекта затененного освещенным предметом.

Если имеется несколько точечных источников света, как, скажем, в помещении с двумя или несколькими лампами, то объекты начинают отбрасывать соответственно две и более тени и создавать сложную картину затенения. Мы говорим при этом, что каждый дополнительный источник усложняет *оптическую структуру* сцены.

В большинстве реальных ситуаций освещение является рассеянным, при котором свет излучается какой-либо пространственно крупной областью. К примеру, солнеч-



**Рис. 1.2.6** Точечное освещение (А) и освещение рассеянным светом (В). Точечные источники дают резко выраженные тени от объектов. При рассеянном освещении такого не происходит.

ный свет в пасмурный день равномерно и полностью рассеян, поэтому примерно равные количества света падают со всех точек верхней части поля зрения. В таких условиях тени, отбрасываемые освещенными предметами и тени, отбрасываемые на эти предметы, ощутимо слабее и менее жесткие, чем при точечном освещении (рис. 1.2.1В). Горным лыжникам хорошо знаком один из эффектов рассеянного освещения в облачную погоду: оно настолько обедняет оптическую информацию, что снег утрачивает рельефность, и поэтому в пасмурные дни у лыжников на склонах гораздо больше всяких неприятностей. Они даже выдумали подходящий термин — «плоский свет».

### 1.2.2. 2 Взаимодействие с поверхностями

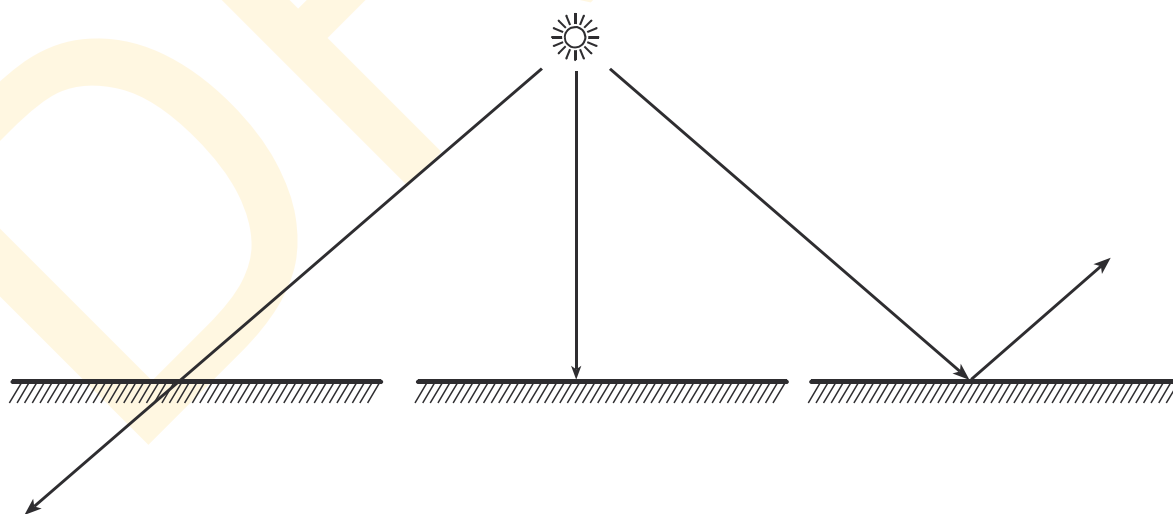
Мы уже говорили о том, что фотоны движутся строго по прямой, но только до того момента, пока не встретятся с поверхностью объекта. В каждом без исключения случае поверхность вносит радикальные изменения в траекторию ударившегося об нее фотона. Эти, как говорят, *поверхностно-индуцированные изменения* в движении фотонов, в конечном счете, и поставляют зрению необходимую информацию о предметах окружающего мира. Только те поверхности, что совсем не меняют траектории фотона, являются полностью прозрачными и буквально невидимыми. Все реальные поверхности в большинстве ситуаций достаточно активно взаимодействуют со светом и видимы наблюдателем.

Когда фотон сталкивается с поверхностью объекта, то происходит одно из трех возможных событий (см. рис. 1.2.7):

- проход фотона через толщу поверхности;
- поглощение поверхностью;
- отражение от поверхности.

Пропущенный свет может не изменить своего направления, а может отклониться (т.н. *рефракция*) — см. рис. 1.2.8.

Из перечисленных вариантов взаимодействия света с поверхностями зрению важнее всего отражение: во-первых, потому что отраженный свет изменен этим взаимодействием (что, собственно и кодирует визуальную информацию об этой поверхности);



**Рис. 1.2.7** Варианты взаимодействия света с поверхностью: фотон может быть пропущен, поглощен, либо отражен ею.





**Рис. 1.2.8** Рефракция света. Когда свет проходит сквозь объекты, он может преломиться и тем исказить восприятие, как, скажем в данном примере.

рактные концы некоей понятийной шкалы, на которой все реальные поверхности занимают промежуточное положение. На рис. 1.2.9 также показано, как выглядят высокоглянцевая и высокоматовая сферические поверхности.

Безусловно, наше описание оптических свойств поверхностей сильно упрощенно, тогда как в реальности их отражающие свойства много сложнее, к примеру: *коэффициент зеркальности* поверхности может меняться как функция от угла падения света (полуглянцевая поверхность). Если рассматривать такую поверхность под углом близким к параллели, то поверхность эта будет гораздо зеркальнее, чем если смотреть на нее строго перпендикулярно.

Чтобы лучше понимать суть процессов, происходящих со светом в реальном мире, давайте попробуем обсудить все то, что мы говорили о его излучении источниками и все то, что мы говорили о его отражении поверхностями:

- все поверхности в той или иной мере отражают свет за исключением абсолютно черных тел (полностью поглощающих его) и абсолютно прозрачных тел (полностью пропускающих);<sup>1</sup>

- большинство поверхностей скорее матовые, нежели зеркальные.

Оба этих факта предполагают, что часть света отражается по всем направлениям от

во-вторых, потому что отраженный свет впоследствии оказывает воздействие на т.н. *рецептивную* (светочувствительную) *поверхность* глаза наблюдателя, что передает информацию о поверхности зрительной системе.

Неотраженный свет тоже важен для понимания отдельных аспектов зрительного восприятия; поглощенный — особенностей цветового зрения; пропущенный — восприятия прозрачности. Однако в текущем разговоре мы сосредоточимся на оптической информации, передаваемой отраженным светом.

Изменение, которое вносит поверхность в поведение фотона, влияет на его траекторию: направление, в котором фотон «отскочит» от поверхности, зависит от направления, в котором он упал на эту поверхность, а также от микроскопической структуры последней.

Если поверхность тщательно отполирована (*зеркальна*), свет отражается от нее в направлении, симметричном тому, в котором упал (см. рис. 1.2.9В), то есть: угол падения будет равен углу отражения.

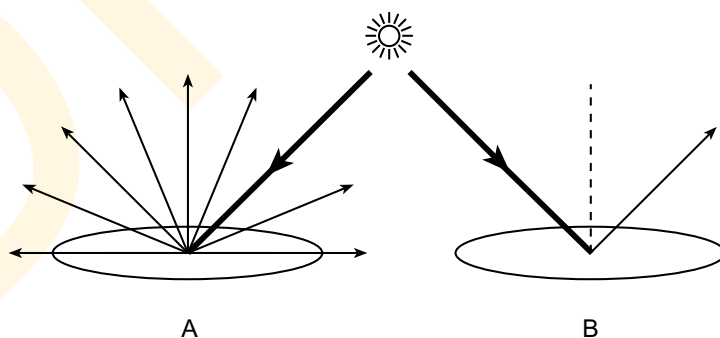
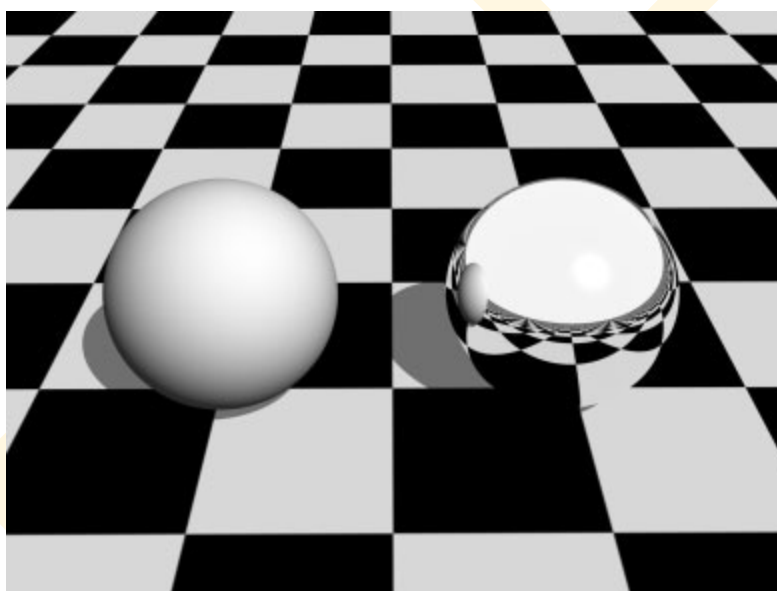
Если поверхность *матовая*, как скажем у листа бумаги, свет равномерно рассеется по полусфере (рис. 1.2.9А). Понятно, что идеально зеркальная и идеально матовая поверхности — это абстрактные

<sup>1</sup> В реальности не существует ни абсолютно черных поверхностей, ни абсолютно прозрачных. Даже угольно-черные предметы отражают от себя какую-то часть света и даже кристально чистые стекла отражают и поглощают некоторое количество упавших на них фотонов.

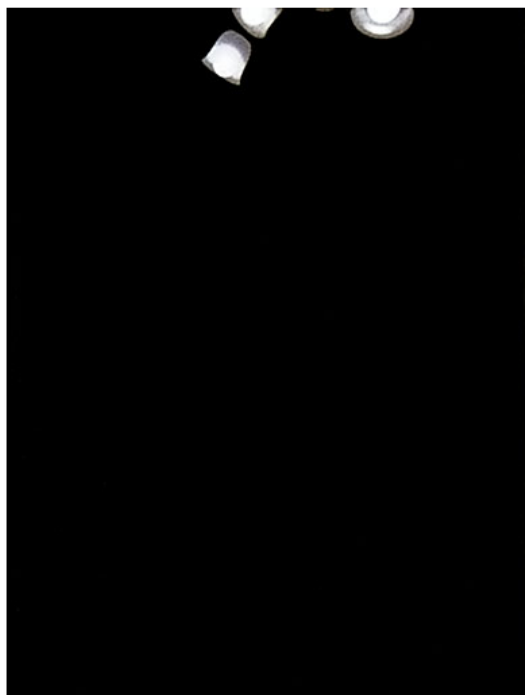
всех поверхностей в окружении. В результате, свет падает на предметы не только непосредственно от источника, но и будучи отраженным от поверхностей предметов сцены (т.н. *световые рефлекс*ы). Поверхности таким образом выступают в роли вторичных источников света, освещая прочие поверхности светом, отраженным от себя. Более того, все фотоны «прыгают» от поверхности к поверхности, т.е. многократно *переотражаются*, прежде чем поглотятся чем-либо.

На рис. 1.2.10 показано то, как воспринималась бы фотоизображение сцены, если бы удалось зафиксировать его на разных этапах хода световых потоков: «А» — свет вышел от источника, но не достиг ни одной поверхности; «В» — первичное отражение от несамосветящихся поверхностей; «С» — это изображение «В», к которому добавлены вторичные отражения; «D» — это изображение «С», к которому добавлены переотражения вплоть до пятого порядка.

Отметим, что несамосветящиеся поверхности не появятся в изображении до тех пор, пока не пройдут первичные отражения; отражение в зеркале не появится, пока не



**Рис. 1.2.9** Матовая (А) и зеркальная (В) поверхности. Два шара на изображении отличаются только структурой поверхности: шар с матовой поверхностью (слева) равномерно отражает падающий на него свет по всем направлениям, тогда как зеркальный шар (справа) отражает свет когерентно, то есть так, что угол отражения равен углу падения. (3D-графика — Д. Табанин).



A



B



C



D

**Рис. 1.2.10** «Фотографии» ванной комнаты, «выполненные» на разных этапах ее заполнения светом: А — свет вышел от источника, но не достиг ни одного из объектов сцены; В — момент времени после первичного отражения света от предметов; С — момент времени после вторичного отражения света от предметов; D — после многократных переотражений.

С любезного разрешения Greg Ward Larson (Silicon Graphics).

<http://radsite.lbl.gov/radiance/HOME.html>

пройдут вторичные отражения, а блики на двери душевой кабины — до тех пор, не пройдут многократные переотражения. Итак, мы видим, что воспринимаемая картина сильно зависит от сложных взаимодействий света с объектами сцены.

Эффект от всех переотражений таков, что свет распространяется среди предметов и заполняет пространство по всем направлениям. Сей факт первостепенно важен, потому как он являет собой сложную оптическую структуру, которая и позволяет собственноручно видеть. По одним направлениям света приходит больше, чем по другим и благодаря этому мы видим поверхности по-разному окрашенными.

Поле зрения, в котором количества света равновелики по всем направлениям, именуется бесструктурным полем зрения или *ганцфельдом*, воспринимаемым нами как безликий серый туман.

### 1.2.2.3 Внешний оптический массив

Американский исследователь-новатор физиологии зрительного восприятия Джеймс Джером Гибсон (1904–1979) предложил именовать оптическую информацию, доступную зрительной системе в данной сцене — *внешним оптическим массивом (ВОМ)*. Внешний оптический массив относится ко всем световым потокам, приходящим в точку наблюдения со всех направлений. Мы говорим «внешний», потому что точка наблюдения буквально окружена лучами света, сходящимися на ней отовсюду. Сие значит, что если наш глаз находится в этой точке, то свет, либо отраженный от окружающих поверхностей, либо испущенный непосредственно источником, будет доступен по всем направлениям. Зрительный акт в точке наблюдения возможен благодаря тому, что хоть и очень сложными, но целиком подчиняющимися законам природы путями, поверхности объектов организуют свет во внешний оптический массив. Сия «законность» оптической структуры ВОМ рождает информацию, позволяющую иметь место зрительному акту.

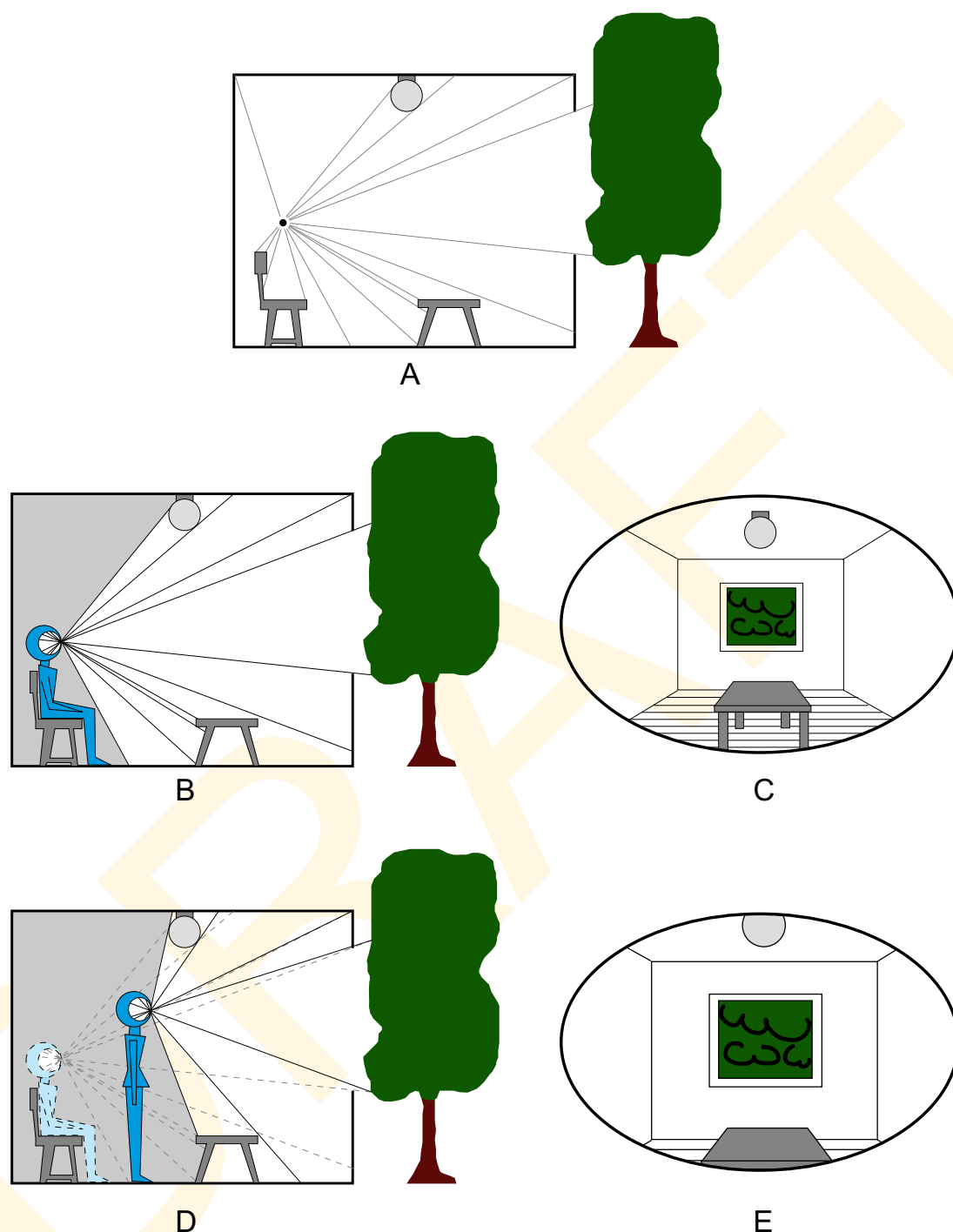
Рассуждая в таком ключе, зрительный акт можно рассматривать как решение паззла: поверхности объектов, меняя характер отражаемого света, в данной точке наблюдения формируют специфический сложный ВОМ. Зрительная система схватывает его структуру и пытается затем выполнить обратное действие — определить характер и взаиморасположение поверхностей, которые предположительно существуют в сцене и которые образовали тот самый ВОМ. Система срабатывает так быстро, что мы не осознаем этот процесс.

Чтобы лучше разобраться в концепции внешних оптических массивов, давайте рассмотрим несколько примеров.

На рис. 1.2.11А схематично показана структура ВОМ в определенной точке наблюдения в помещении, где перед окном с видом на дерево, стоит стул. Отметим, что в точку наблюдения свет приходит со всех направлений и что внешний оптический массив в данной точке существует независимо от присутствия зрячего наблюдателя.

Рис. 1.2.11В демонстрирует взаимосвязь между полем зрения наблюдателя и внешним оптическим массивом: глаз схватывает его некое направленное подмножество, а оставшаяся часть внешнего оптического массива остается невидимой, поскольку глаз принимает свет лишь строго фронтально. На рис. 1.2.11С показан результирующий световой паттерн, который будет воздействовать на глаз в данной точке наблюдения и который моментально отрисует на задней светочувствительной поверхности глазного яблока оптическое изображение. Отметим, что сие есть отправная точка зрительного акта.





**Рис. 1.2.11** Внешний оптический массив. А — образование внешнего оптического массива в данной точке пространства; В — когда глаз наблюдателя находится в этой точке, он «вырезает» (или как еще говорят — сэмплирует) фрагменты внешнего оптического массива по данной точке; С — оптическое изображение на сетчатке наблюдателя, глаз которого находится в обсуждаемой точке; D — характер изменения внешнего оптического массива при изменении местоположения наблюдателя; E — оптическое изображение на сетчатке наблюдателя, глаз которого находится в новой позиции. (А, В, и D — Gibson, 1979).

Важно понимать, что у каждой точки пространства имеется свой, специфический **внешний оптический массив**.

Каждый ВОМ неповторим и несет информацию об окружающем мире слегка отличную от того, какую несет сосед. Чтобы проиллюстрировать сей факт, рассмотрим, что произойдет, когда наблюдатель начнет вставать со стула: его глаза начнут смещаться по некой траектории, на которой лежат иные точки наблюдения и зрительная стимуляция будет определяться своим неповторимым ВОМ в каждой точке.

На рис. 1.2.11D сплошными линиями показан финальный вариант внешнего оптического массива, а пунктирными — исходного. Хорошо видно как отдельные фрагменты поверхностей, которые вначале были видимыми скрываются от наблюдателя, а те, что не были видны — появляются. На рис. 1.2.11Е показано оптическое изображение, воздействующее в этот момент на светочувствительный слой глаза наблюдателя и хорошо видно, как оно изменилось в сравнении с исходным.

Изменения в оптической информации, вызванные сменой точки наблюдения, проясняют исключительно важное отличие между моментальным, *статическим* ВОМ и протяженным во времени — *динамическим*.

Статический внешний оптический массив характеризуется паттерном света, сходящегося в точке наблюдения; динамический массив может быть охарактеризован *оптическим потоком* — т.е. последовательностью световых паттернов. Таким образом динамический ВОМ поставляет наблюдателю информацию в рамках дополнительной размерности — временной. Сие оказывается невероятно важным в контексте многих феноменов зрительного восприятия, в частности таких как наша способность воспринимать третью пространственную размерность внешнего мира — его глубину (расстояние от наблюдателя), определять траекторию движущихся объектов и воспринимать нашу собственную траекторию движения в окружающей обстановке.

О каждом из перечисленных выше явлений мы будем говорить подробнее в главах 5 и 10, когда будем детально разбирать вопросы восприятия глубины и движения.

### 1.2.3 Формирование изображений

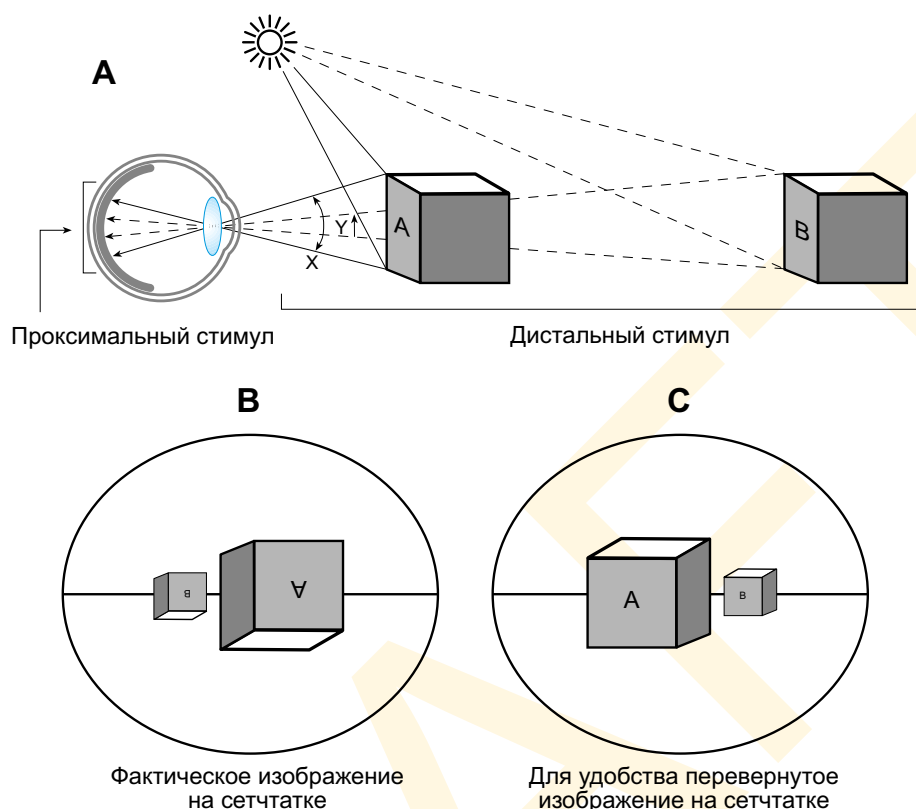
Если зрение существует для того, чтобы поставлять организму высокоточную информацию о внешнем мире, то должна существовать тесная взаимосвязь между геометрией поверхностей объектов сцены и светом, проникающим в глаз в данной конкретной точке наблюдения. На рис. 1.2.10 был показан характер такой взаимосвязи. В данном разделе мы подробно рассмотрим то, каким образом на задней поверхности глаза формируется изображение и как законы т.н. *перспективной проекции* описывают этот процесс.

#### 1.2.3.1 Оптические изображения

Образование изображения сцены показано на рис. 1.2.12: внешний мир трехмерен; свет, исходящий от источника и заполняющий трехмерное пространство отражается от объектов и по строго прямым траекториям летит в глаз наблюдателя (рис 1.2.12А); фотоны проникают сквозь оптические среды глаза, формируя на его задней поверхности перевернутое двухмерное изображение (рис. 1.2.12В и С).

Объекты внешнего мира создают, как говорят, *дистальный зрительный стимул* (т.е. удаленный от наблюдателя), а оптическое изображение на задней поверхности глаза представляет собой *стимул проксимальный*, т.е. ближайший к наблюдателю.

Размер оптического изображения обычно оценивают в градусах визуального угла:



**Рис. 1.2.12** Формирование на сетчатке оптического изображения сцены. Свет, отраженный от объемных (трехмерных) объектов (дистальный стимул), попадает в глаз и формирует двухмерное изображение на сетчатке (проксимальный стимул). Пространство между объектами сцены в ее двухмерной проекции на сетчатку напрямую не отражено, но разноудаленные от наблюдателя объекты обретают разный размер на этой проекции.

угла, образованного прямыми, идущими из т.н. *фокальной точки* глаза к краям оптического изображения. Важно понимать, что визуальный угол характеризует пространственный размер проксимального стимула, но не дистального. Внешний объект, создающий дистальный стимул, будет иметь меньший угол, когда удаляется от наблюдателя и больший — когда приближается. Понимание взаимосвязи между размером объекта, его удаленностью и размером оптического изображения очень важно для понимания того, как мы воспринимаем размер и пространственное положение объектов (см. гл. 5).

Пожалуй, наиболее важным в понимании механизмов зрения моментом является то, что изображение на задней поверхности глаза *двухмерно* (но не трех-), и напоминает проекцию слайда на искривленный экран. В результате объективная трехмерная информация о предметах реального мира теряется: утрачивается размерность *глубины* — расстояния от фокальной точки глаза до поверхности объекта.

Чтобы воспринимать мир трехмерным (что, собственно, и происходит) утраченная размерность должна быть каким-то образом восстановлена из двухмерной информации оптического изображения (о чем мы будем говорить позднее).

### 1.2.3. 2 Проективная геометрия

Процесс т.н. *мэппинга* трехмерной оптической информации в двумерное изображение идет по строго определенным законам и может быть описан математически. Наиболее подходящим методом такого описания является т.н. *проективная геометрия*: дисциплина, описывающая то, как многомерное пространство отображается в пространство маломерное. В случае статического зрения проективное отображение (мэппинг) происходит из трехмерного пространства реальности в двумерное пространство плоскости изображения. Следовательно, проективная геометрия может точно описать то, как трехмерные объекты сцены поточечно отображаются на двумерной плоскости изображения этой сцены и какие свойства этого изображения остаются, как говорят, *инвариантны* (неизменны) в разных проекциях.

При динамическом зрении проективное отображение происходит из четырехмерной пространственно-временной структуры в трехмерное пространство оптического потока, разворачивающегося во времени на двумерной плоскости задней поверхности глаза.

Есть все основания полагать, что проективная геометрия это идеальный математический инструмент понимания процесса образования изображения. Однако проблема состоит в том, что проективная геометрия сама по себе не может моделировать весь комплекс оптических феноменов, поскольку не содержит соответствующего аппарата моделирования отражений, поглощений, либо преломлений световых потоков. К примеру, в мире, наполненном непрозрачными поверхностями, в глаз попадает свет, отраженный по данному направлению только от *ближайших* поверхностей. Фотоны, летящие от удаленных точек, либо поглощаются, либо отражаются другими ближайшими к ним поверхностями, что, разумеется, препятствует воздействию этих фотонов на орган зрения наблюдателя.

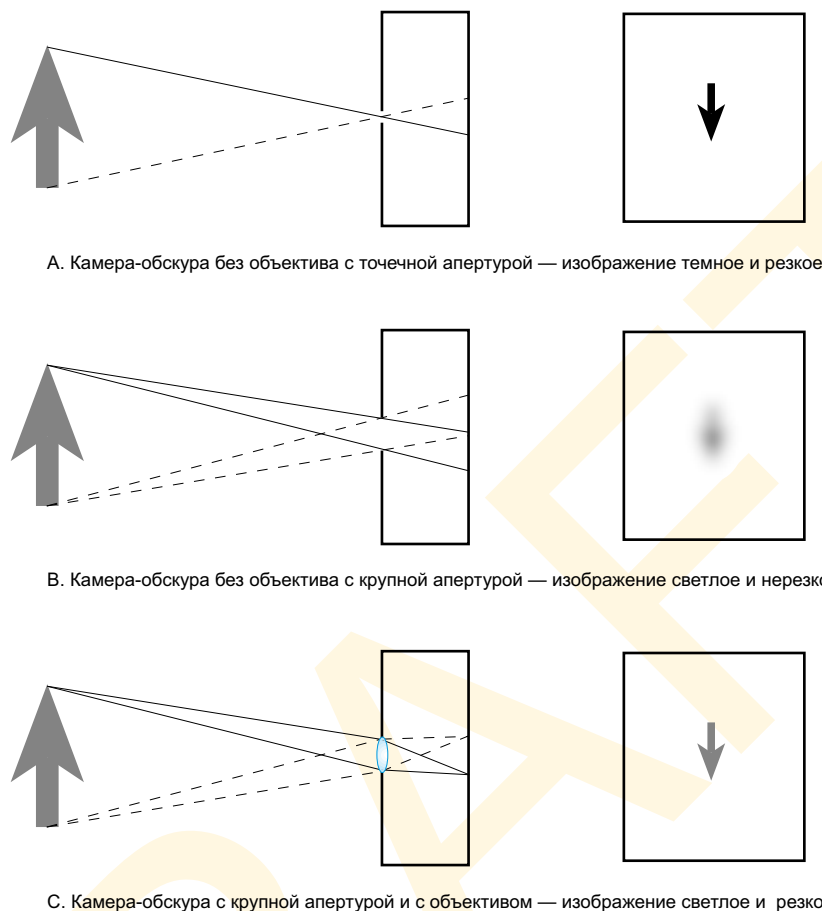
Сложности, которые возникают в результате взаимодействия света с поверхностями, несколько умаляют ценность проективной геометрии как инструмента, позволяющего строить модели возникновения двумерных изображений из трехмерной оптической информации реального мира, и поэтому приходится детально разбираться в том, как именно свет, отраженный трехмерными объектами, формирует оптическое изображение, когда падает на плоскость.

### 1.2.3. 3 Перспективные и ортографические проекции

Одним из простейших путей получения двумерного изображения трехмерных объектов является установка точечного отверстия на определенном расстоянии от плоскости изображения или какой-либо иной плоской поверхности, расположенной с внутренней стороны непрозрачной коробки. Такое приспособление именуется *камера-обскура* (рис. 1.2.13А).

Поскольку фотоны летят только по прямым линиям, свет, падающий на каждую точку плоскости изображения (плоскость проекции), приходит туда будучи отраженным (или испущенным) той или иной конкретной точкой внешнего пространства. Такая точка лежит на отрезке, берущем начало в соответствующей точке изображения и проходящим через отверстие в камере. Описанное является основой геометрии т.н. *перспективных проекций (полярных проекций)* — при которых световые лучи, формирующие изображение, сходятся в направлении единой фокальной точки (либо полюса).

С помощью камеры обскуры не так то просто получить добротное резкое изображение: диаметр отверстия должен быть очень маленьким, порядка 0.4 мм. Однако по-



**Рис. 1.2.13** Оптика камеры-обску́ры. А — камера-обску́ра с точечной апертурой без всяких объективов дает вполне резкое (но темное) проекционное изображение; В — камера-обску́ра с крупной апертурой без объектива дает расфокусированное проекционное изображение; С — камера-обску́ра с крупной апертурой, но оснащенная объективом, может дать четкое проекционное изображение, если фокусное расстояние объектива будет соответствовать расстоянию до проекционной поверхности.

сколько через столь мелкое отверстие на плоскость изображения попадает весьма малое количество света, то рассматривать полученную картинку нужно в хорошо затемненных условиях.<sup>1</sup> Если отверстие сделать больше, то результирующее оптическое изображение станет светлее, но потеряет резкость, поскольку свет будет проходить уже не через одну точку, но через множество (рис. 1.2.13В). Проблема решается установкой в отверстии прозрачной выпуклой линзы — объектива, отклоняющего входящий световой поток к точке, находящейся несколько кзади (т.н. *фокальная точка* объектива) (см. рис. 1.2.13С). Объектив в своей фокальной точке организует, если угодно, «виртуальное отверстие», создающее на задней поверхности камеры проекционное изображение — контрастное и резкое, но субъективно уже более яркое, поскольку

<sup>1</sup> Самый лучший способ — изготовить непрозрачный ящик с единичным отверстием на одной стороне и полупрозрачной стенкой на другой. Если (как фотограф) накрыться, черной тканью, то на задней стенке можно будет наблюдать перевернутое изображение сцены.



ку через большее отверстие проходит больше света. Как мы увидим при разборе материала в разделе 1.3, глаз имеет в своем устройстве такой объектив, функция которого состоит в фокусировании изображения на задней внутренней поверхности глазного яблока.

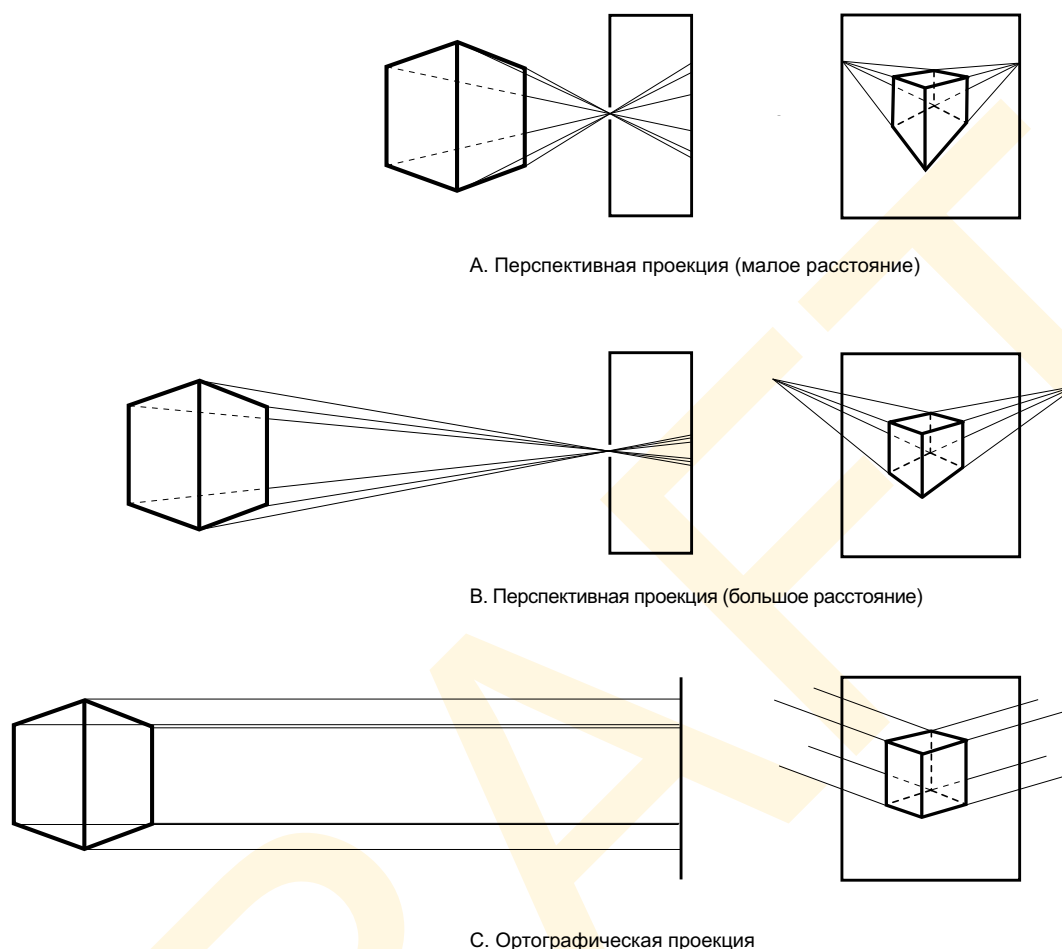
Далее. Коль скоро сложные эффекты свет-поверхностных взаимодействий моделируемы, математика перспективной (или полярной) проекции вполне подходит для моделирования процесса образования оптического изображения в глазу человека (см. рис. 1.2.14А и В). Однако, все намного сложнее, чем может показаться на первый взгляд, и чтобы упростить ситуацию, ученые, изучающие зрение, часто прибегают к т.н. *ортографическим (параллельным) проекциям* (см. рис. 1.2.14С). В этом случае изображение рассматривается как образованное параллельными друг другу световыми потоками, перпендикулярными плоскости оптического изображения (в отличие от того, что происходит с лучами, проходящими через отверстие в камере-обскуры).

Математическое упрощение при использовании ортографических проекций состоит в том, что размерность глубины пространства (расстояния от изображения до объекта) просто игнорируется, в то время как вся пространственная информация, перпендикулярная направлению просмотра полностью сохраняется. Сие свидетельствует о том, что когда расстояние от изображения до объекта относительно велико в сравнении с глубиной самого объекта (рис. 1.2.8В), ортографическая проекция является добротной аппроксимацией проекции перспективной. Однако на близких расстояниях различия между видами проекций могут оказаться весьма существенными (см. рис. 1.2.8А).

Чтобы лучше понимать взаимосвязь между перспективной и ортографической проекциями рассмотрим то, что происходит по мере удаления объекта от фокальной точки (см. рис. 1.2.14А и 1.2.14В): лучи, проходящие через фокальное отверстие становятся все более и более параллельными друг другу, а на бесконечном расстоянии они оказываются строго параллельными — т.е. как при ортографической проекции (рис. 1.2.14С). Стало быть, ортографическую проекцию можно рассматривать как частный случай перспективной проекции, в котором расстояние от объекта до фокальной точки бесконечно. Принципиальное же отличие состоит в том, что перспективное изображение объекта, находящегося на бесконечном удалении — это просто точка, тогда как при ортографической проекции изображение объекта будет иметь некий пространственный размер.

### 1.2.4 Зрение как процесс «инверсии»

Давеча мы описали то, как свет отраженный от трехмерных объектов образует на светочувствительной поверхности глаза двухмерное оптическое изображение. Процесс формирования изображения идет по законам оптики, поэтому в любой сцене с четко известными условиями освещения и точкой наблюдения мы с высокой точностью можем определить, какое оптическое изображение получится. В сфере *компьютерной графики* как раз и занимаются моделированием оптических изображений сцен. Многие вопросы в этой области сегодня хорошо изучены и мы можем по достоинству оценить реалистичность современных компьютерных изображений, синтезированных без какой-либо фактической оптической подоплеки, в частности, изображения на рис. 1.2.10, отрисованные по алгоритму т.н. *рэйтрейсинга* (метод бегущего луча), имитирующего образование изображения из модели поверхностей в помещении и модели поведения света, падающего на эти поверхности. Фактически, чтобы сконст-



**Рис. 1.2.14** Перспективная и ортографическая проекции. Перспективная проекция через точечное отверстие (А и В) сводит (конвергирует) параллельные линии в точке схода (при ближней перспективе схождение больше — А, при дальней — меньше — В). Ортографические проекции предполагают полную параллельность световых потоков и не дают конвергенции линий в оптическом изображении.

руировать изображение виртуальной реальности компьютерные программы имитируют оптические события эмиссии фотонов, их отражения, пропускания и поглощения. Такие программы позволяют воспроизводить эффекты последовательного отражения света от разных поверхностей (см. рис. 1.2.10А-Д): после каждой стадии имитации отражения программу можно остановить и посмотреть, что получается. Понятно, что сие невозможно в реальной оптической обстановке.

На ранних стадиях зрительного акта зрение пытается решить т.н. *задачу инверсии*, т.е. пройти путь от оптических изображений сцен на задней поверхности глаз (в связке с освещением породившим эти оптические изображения) к зрительному осознанию сцены. Иными словами, зрение стремится *инвертировать* процесс формирования оптического изображения за счет «оптических» преобразований в обратном направлении.

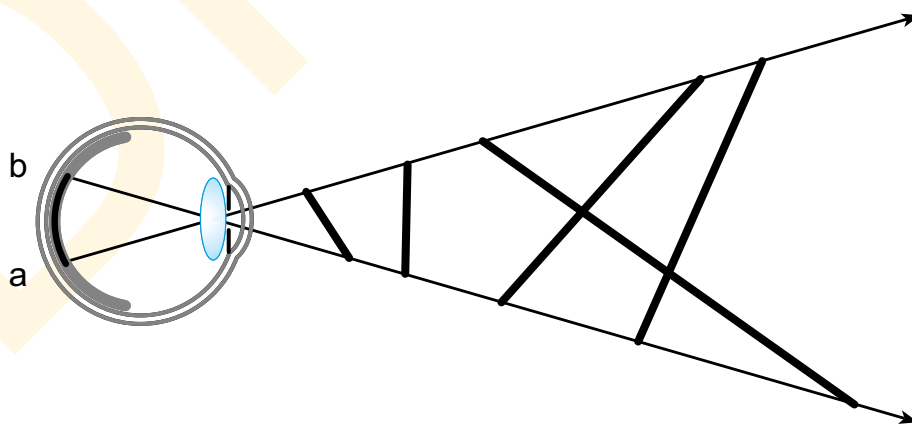
Простого способа инверсии, к сожалению, не существует. Трудность состоит в том, что математическая связь между реальной сценой и ее проективным изображением,

как говорят, *невозвратно*. Дело в том, что мэппинг внешнего оптического массива в проективное изображение идет от трех размерностей в две и *есть* т.н. *вполне определенная функция*: каждая точка сцены проецируется на уникальную точку оптического изображения. Обратный мэппинг, т.е. переход от оптического двумерного изображения к трехмерной реальности не является вполне определенной функцией — каждая точка оптического изображения может быть спроецирована на бесконечное число точек во внешнем мире, и, следовательно, на каждое двумерное оптическое изображение на задней поверхности глаза приходится бесконечное количество отличных друг от друга сцен. Рис. 1.2.15 иллюстрирует неопределенность обратного проектирования: отрезок линии в оптическом изображении может быть проекцией бесконечного числа линий сцены.

Итак, мы говорим, что задача инверсии — это т.н. *недоопределенная задача*, т.е. цель, которая недоопределена сенсорными данными в изображении.

Как уже было сказано выше, простых путей инверсии не существует, и далеко не факт, что наш мозг всегда успешен в поиске решения. Велик соблазн предположить, что трехмерное восприятие вообще невозможно, но мы знаем, что оно-таки имеет место, более того — с очень высокой точностью, поскольку зрительная система человека в большинстве ситуаций весьма эффективно справляется с задачей инверсии.

Итак, как же решается эта, казалось бы, неразрешимая задача? У разных ученых разный подход к объяснению данного феномена (см. гл. 5), однако ведущая гипотеза сегодня строится на той идее, что трехмерное восприятие является следствием большого числа, как говорят, *высокоправдоподобных гипотез* — допущений об окружающем мире, которые делает зрительная система, его объектах и условий их наблюдения. Эти гипотезы в большинстве случаев и позволяют справиться с задачей инверсии. Если допущения оказываются истинными, результирующее решение будет веридикальным. Таким образом зрение — суть *эвристический процесс* — процесс, в котором строятся предположения (версии) о наиболее вероятных внешних условиях, давших оптическое изображение на сетчатке глаза. Такой процесс мы называем эвристическим в том числе потому, что базируется он на выполнении т.н. *дедуктивных схем*, основанных на дополнительных (и не всегда правомерных) версиях, и которые ведут порой к ошибочным выводам (как, скажем, в случае зрительных иллюзий). Однако в повседневности эти предположения истинны и нормальное зрительное восприятие высоковеридикально.



**Рис. 1.2.15** Иллюстрация обратной проекции. Единичная линия на сетчатке может являться проекцией бесконечного числа линий окружающего мира.

На страницах нашей книги, в частности в гл. 5, мы будем подробно говорить о дополнительных версиях, где речь пойдет о восприятии деталей: мы покажем, каким образом дедуктивные правила позволяют зрительной системе выходить за рамки имеющейся оптической информации и решать, казалось бы неразрешимую задачу инверсии.

### 1.3 ЗРИТЕЛЬНАЯ СИСТЕМА И ЕЕ СОСТАВЛЯЮЩИЕ

Итак, мы понимаем, что задача инверсии в целом разрешима. Может быть не всегда идеально — но непрерывно и с точностью, достаточной для того, чтобы поставлять нам необходимое знание о внешнем мире. Вся наша книга, в сущности, посвящена тому, что известно сегодня о механизмах инверсии.

Мы начнем с беглого обзора общей структуры того отдела нервной системы, о котором известно, что он отвечает за обработку зрительной информации. Описание, которое мы дадим здесь во всех отношениях краткое и поверхностное. Сие намеренно, поскольку наша цель на данном этапе — дать общее представление о биологической структуре зрительной системы и подготовить читателя к дальнейшему, уже углубленному разговору. Обсудив некоторые основные положения касающиеся «аппаратной части» зрительной системы, то есть ее *анатомии* (физического строения), мы сможем уже с достаточным пониманием говорить о ее биологической функции, то есть — *физиологии*.

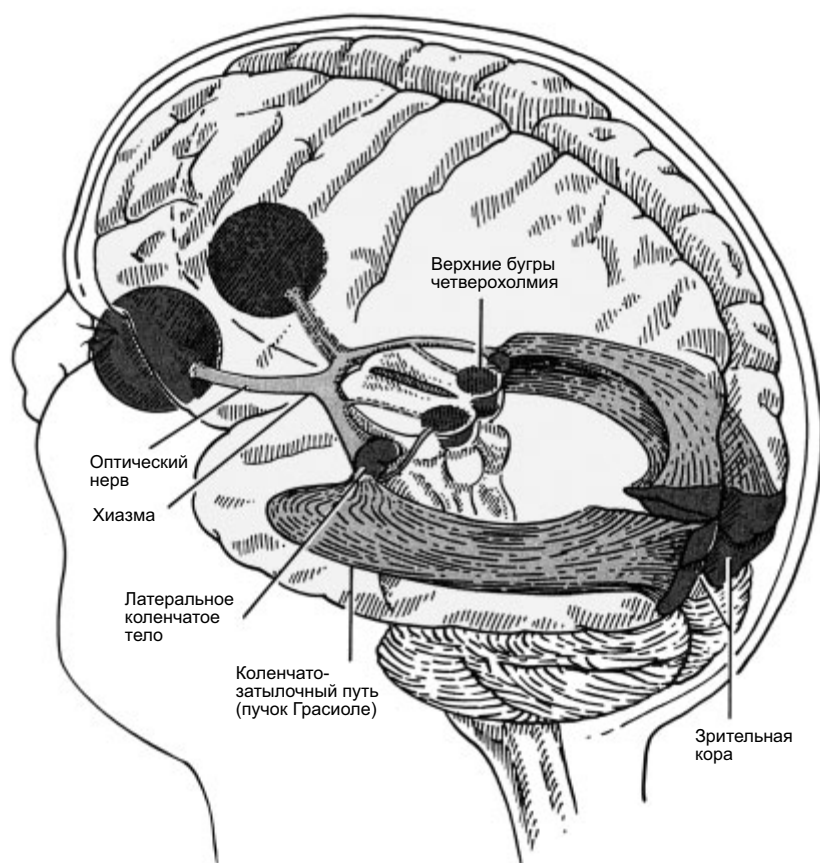
#### 1.3.1 Глаз

Хотя с античных времен известно, что глаза — это сенсорный орган зрения, четкое представление об их функции пришло совсем недавно. Древнегреческий философ Платон (427 — 347 гг. до н.э.) был уверен, что «чревный огонь» рождает некие лучи, эмануруемые (испускаемые) глазами на окружающие предметы. Эпикур (341 — 270 гг. до н.э.) опроверг платоновскую теорию эманации, но взамен уверовал в то, что некие тончайшие реплики (копии) предметов каким-то образом мгновенно проникают в мысли через глаза. Гален (130 — 200 н.э.) впоследствии детально развил обе этих идеи, украсив их «физиологическими» подробностями и предположив, что после того как лучи испущены глазом, они взаимодействуют с объектами и возвращаются назад в глаз. Он полагал, что в глазных средах эти лучи сочетаются со «зрительным духом», постоянно дефилирующим между мозгом и глазом, и перетаскивающим реплики воспринятых объектов.

Новая эра в физиологической оптике не начиналась до тех пор, пока известный арабский философ Альгазен (ибн аль-Хайсам) (965 — 1039 гг. н.э.) не подал идею, что глаз — суть камер-обскура, о которой мы говорили в разделе 1.2.2. Великим прозрением Альгазена явилось то, что зрительный акт обязан отражению света, испущенного его источником, от предметов с последующим проникновением этого света в глаза наблюдателя. Однако в полной мере оптика глаза стала понятной, когда была изобретена линза: известный астроном Иоганн Кеплер (1571 — 1630) поставил точку в этом вопросе, понятийно объединив оптическую систему глаза с линзой и заложив тем самым основу современной теории *физиологической оптики*, которую мы и будем обсуждать ниже.

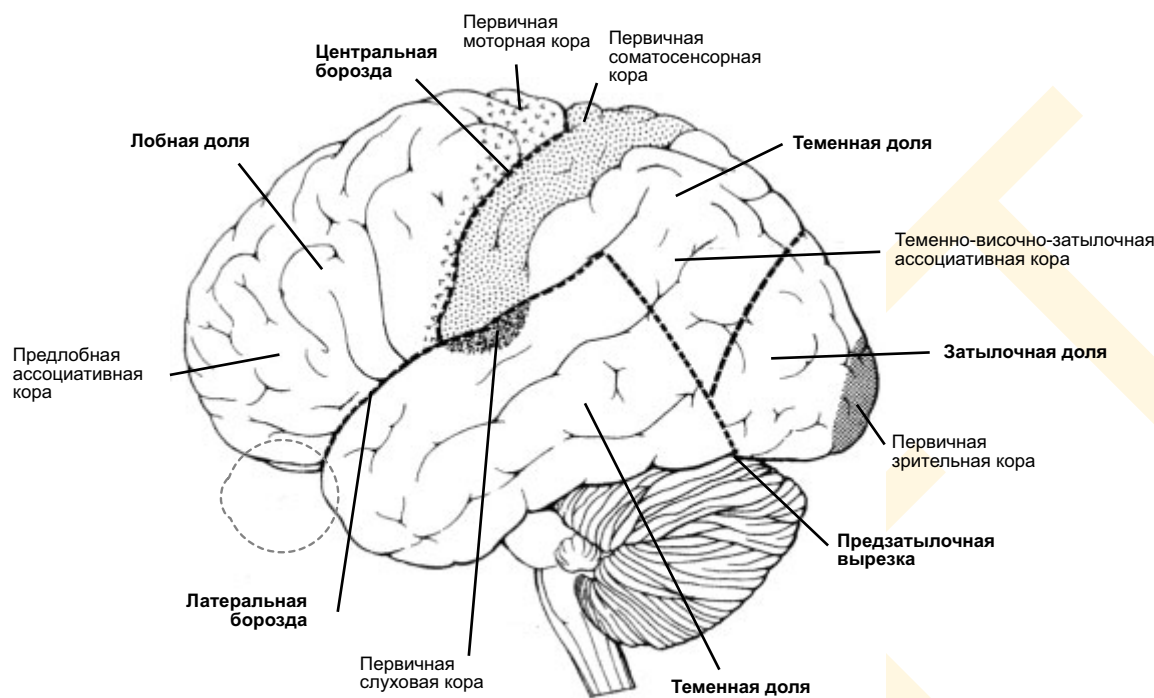
### 1.3.1.1 Глаз и мозг

Хотя Гален почти ничего не знал об анатомии и физиологии зрения, он безусловно был прав в том, что глаз и мозг неразрывно связаны друг с другом (см. рис. 1.3.1). Сегодня мы знаем, что оптическая информация передается из глаза в т.н. *первичную зрительную кору* затылочных долей головного мозга (см. рис. 1.3.1), а затем рассылается во многие зрительные центры, расположенные в задней *височной* и *теменной* коре (рис. 1.3.2). По некоторым оценкам более 50% коры головного мозга макаки отвечает за зрительную функцию (DeYoe & Van Essen, 1988; Van Essen et al., 1990), хотя у людей, вероятно, слегка поменьше. То есть, зрительная система состоит из мозга и глаз, и связка «глаз-мозг» должна работать очень четко, чтобы извлекать из внешнего оптического массива достоверную информацию об окружающем мире, жизненно необходимую особи: глаза должны собирать и регистрировать информацию, которую несет свет, а мозг должен обрабатывать эту информацию так, чтобы она была максимально полезна для жизнедеятельности.



**Рис. 1.3.1** Общая схема зрительной системы человека. Зрительная обработка начинается в глазах, после чего данные передаются по оптическому нерву в подкорковые отделы головного мозга. Из подкорковых отделов в зрительную кору информация идет по т.н. *первичному зрительному пути* (колленчато-затылочному пути, пучку Грасиоле). Из зрительной коры информация расходуется в разные отделы мозга. *Вторичный зрительный путь* исходит от оптического нерва и направляется к верхним буграм четверохолмия, а затем в прочие центры мозга (Rosenzweig & Leiman, 1982).





**Рис. 1.3.2** Зрительные области коры головного мозга человека. Из первичной зрительной коры в затылочной доле зрительная информация расходится по двум крупным путям: нижнему (вентральному) в нижние отделы височной коры и верхнему (дорзальному) — в теменную долю. Оба пути заканчиваются в лобной коре (Kandel, Schwartz & Jessel).

То, что зрительная функция — суть сочетанная функция глаз и мозга, объясняет феномен слепоты у людей со здоровыми глазами, но поврежденной зрительной корой и наоборот. Действительно, существуют оба варианта слепоты. Травмы (либо заболевания) глаз иногда приводят к нарушению их функции (состоящей, напомним, в регистрации оптической информации и/или отправке их в зрительные отделы мозга). Такие ситуации вызывают тот вид слепоты, о котором все слышаны: утрате зрения, в результате нарушения передачи зрительной информации.

Довольно редко, но встречаются также и люди, кто не может видеть (или как минимум не имеет зрительного опыта), но глаза у которых работают нормально. Такой вид слепоты называют «слепозрением», которое является следствием повреждения критично важных участков зрительной коры, возникшего в результате какого-то патологического процесса, хирургического вмешательства, травмы и т.п. (см. раздел 13.2.2). Повреждение других отделов зрительной коры не приводит к слепоте, но может вызывать т.н. *селективную (избирательную) недостаточность* зрительного восприятия. К примеру, некоторые больные легко распознают лица, но не могут зрительно идентифицировать даже членов своей семьи (см. раздел 9.2.5). Другие могут описать и порой нарисовать тот или иной бытовой предмет, но не могут назвать его или пользоваться им. Замечено также, что некоторые больные великолепно видят статические сцены, но совсем не могут воспринимать движение, ощущая происходящее как серию застывших кадров (см. раздел 10.1.4).

Подробно об этих проблемах мы будем говорить в соответствующих разделах, в частности, в конце книги, когда коснемся темы осознанного зрительного опыта.

### 1.3.1.2 Анатомия глаза

Многое о строении глаз общеизвестно. К примеру то, что у человека два более или менее круглых глаза с выпуклостью на передней поверхности и что расположены они примерно по горизонтальной срединной линии головы и посажены в глазницы — полукруглые отверстия в черепе, стенки которых защищают глазные яблоки от внешних механических воздействий, но позволяющие при этом им вращаться по разным направлениям.

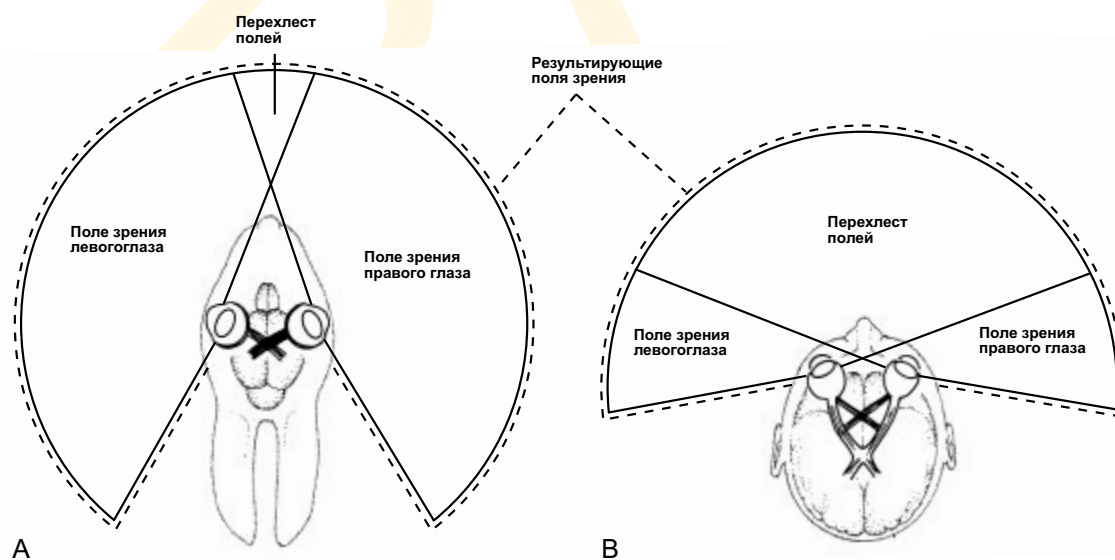
Глазное яблоко движется в пространстве за счет шести сильных мышц (четырех прямых и двух косых), именуемых *экстраокулярными* или *глазодвигательными* и управляемыми определенной областью головного мозга (см. раздел 11.1.2).

*Глазные движения* позволяют сканировать поле зрения без участия моторики головы, а также фиксировать взгляд на разноудаленных объектах.

*Веко и ресницы* механически защищают глаз, а также увлажняют и очищают его.

Перечисленное касается не только глаз человека, но как минимум всех млекопитающих. У представителей некоторых видов птиц, например у голубя и совы глаза неподвижны и этим пернатым приходится целиком полагаться на моторику головы. Сие снижает частоту смены полей зрения потому, что голова много тяжелее и инертнее глаз.

У разных видов животных глазные яблоки посажены принципиально по-разному: у многих (в отличие от человека) они разнесены по сторонам; фронтальное расположение глаз (человек, обезьяна, волк и пр.) обеспечивает большой перехлест полей зрения (см. рис. 1.3.3) и при этом, соответственно, небольшое общее поле. Биологическая вы-



**Рис. 1.3.3** Наложение монокулярных полей зрения. У зайца (А) латеральная посадка глаз дает огромное результирующее поле зрения, но малый бинокулярный перехлест (бинокулярное поле зрения). У человека (В) фронтальная посадка глаз дает ощутимо меньшее результирующее поле зрения, но радикально большее бинокулярное (Sekuler & Blake, 1985).

года такого наложения в том, что оно обеспечивает т.н. *бинокулярное* («двуглазое») зрение, необходимое для точного определения глубины пространства (см. раздел 5.3). Способность точно оценить расстояние от себя до объекта — это колоссальное эволюционное достижение хищников (точная оценка расстояния критична во время охоты), глаза которых расположены обычно фронтально и с сильным перекрытием полей зрения каждого.

В противоположность сему панорамные поля зрения — это эволюционное преимущество потенциальной добычи, позволяющее эффективно отслеживать опасность и требующее того, чтобы глаза были расположены латерально (по сторонам головы). Впрочем, у многих животных фронтальное, либо латеральное расположение глаз обусловлено иными причинами, каждая из которых — суть компромисс между точностью в восприятии глубины пространства и углом обзора.

### 1.3.1.3 Физиологическая оптика

Физиологическая оптика изучает процессы, происходящие на «переднем крае» зрительной системы. Глаз (как и фотокамера) решает две важных оптических задачи: захват световых потоков, отраженных от поверхностей объектов внешнего мира, и сбор этих потоков в оптическое изображение на светочувствительной поверхности глаза — сетчатке.

**Анатомия глаза.** Глаз состоит из трех концентрических слоев: внешнего слоя, именуемого *склерой*; промежуточного, именуемого *сосудистой оболочкой*, (*uveальным трактом*, *uvea*); и внутреннего, именуемого *сетчаткой* (по латыни — *retina*).

Склера — это белый непрозрачный фрагмент глазного яблока, которую видно невооруженным глазом, и состоящая в основном из коллагеновых волокон. Функция склеры состоит в организации вместилища и механической защите внутренних функциональных элементов глаза. Склера является продолжением *роговицы* — прозрачной ткани на самой передней поверхности глазного яблока (см. ниже).

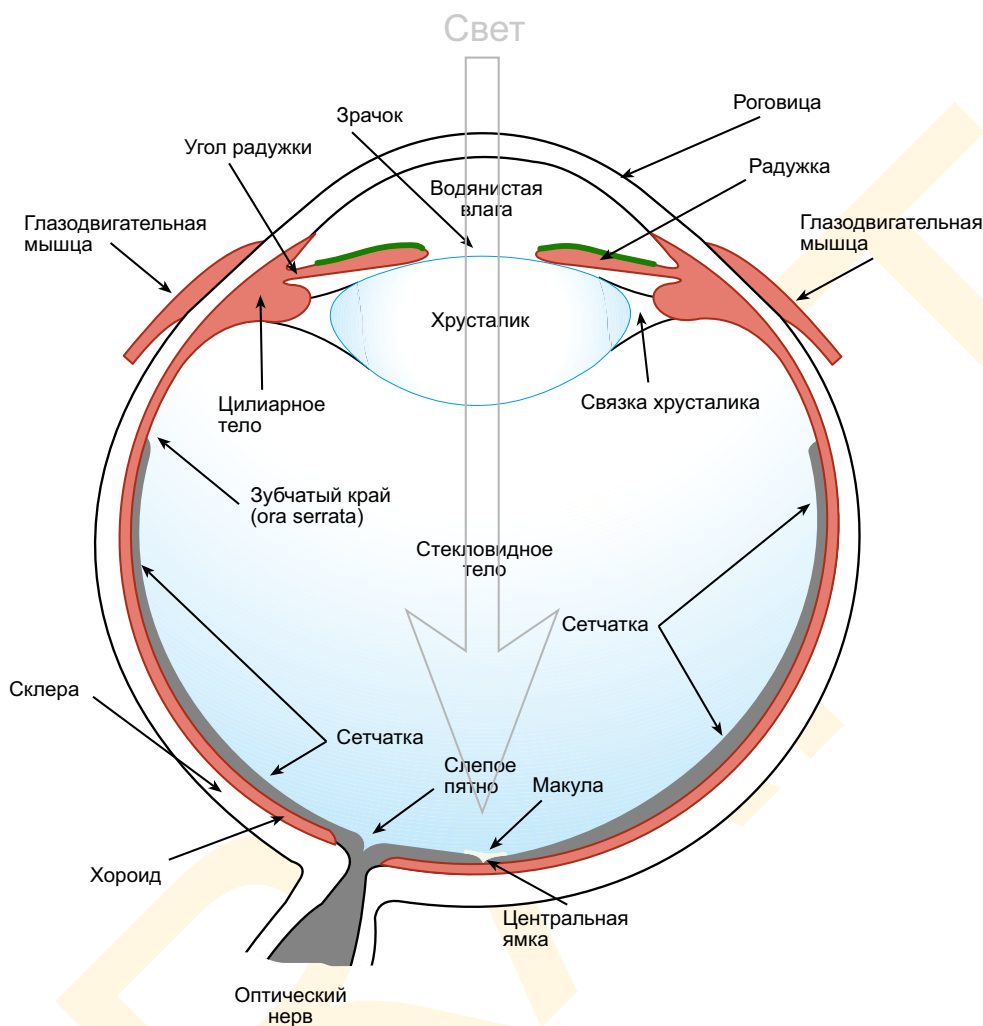
У глаза множество отделов, выполняющих разные оптические функции (см. рис. 1.3.4), и, чтобы получить четкое представление о том, как глаз пропускает свет, давайте пройдем вместе с фотоном через оптические среды глазного яблока.

Итак, вначале фотон проходит сквозь роговицу. Позади роговицы находится полость — *передняя камера глаза*, заполненная бесцветной прозрачной жидкостью — *водянистой влагой*. Водянистая влага создает давление (порядка 16 мм рт. ст.), именуемое *внутриглазным*. Заметим, что водянистая влага непрерывно вырабатывается т.н. *цилиарным телом* (см. рис. 1.3.4) и дренируется *шлеммовым каналом* (*канал Шлемма*), расположенным в углу, образованном радужкой и роговицей.

Миновав водянистую влагу, фотон пролетает *зрачок* — отверстие переменного диаметра в непрозрачной и специфически окрашенной ткани — *радужке* (*радужной оболочке*). Окраска радужки — это, соответственно, и есть «цвет глаз». Сразу позади радужной оболочки расположена линза — *хрусталик*, форма которого меняется под воздействием т.н. *цилиарных мышц*, прикрепленных к его краю по всему периметру.

Пройдя через хрусталик, фотон попадает в заднюю камеру глаза, заполненную *стекловидным телом*. Стекловидное тело, задающее собственно глазную массу, состоит в основном из коллагена и гиалуроновой кислоты и имеет гелеподобную консистенцию.

Наконец, фотон достигает «цели» и сталкивается с *сетчаткой* — светочувствительной выстилкой задней поверхности глаза, толщина которой составляет всего лишь 0.2 мм.



**Рис. 1.3.4** Горизонтальный разрез правого глаза. Прежде чем достичь светочувствительных рецепторов сетчатки свет проходит оптические среды глаза: роговицу, водянистую влагу, хрусталик, стекловидное тело, клетки различных слоев сетчатки. В сетчатке энергия фотонов конвертируется в электрохимические сигналы, которые идут по зрительному нерву в мозг.

Сетчатка содержит в себе более чем 100 млн. светочувствительных клеток — *фоторецепторов*, конвертирующих *изменения* световой энергии, воздействующей на сетчатку, в *нервные импульсы*. Иными словами, нервные импульсы кодируют информацию о *модуляциях* света, упавшего на сетчатку. Эта информация отправляется затем в зрительные центры головного мозга.

Позади сетчатки находится слой *пигментного эпителия*. Пигментный эпителий не реагирует на свет и не участвует в кодировании оптической информации, но он выполняет очень важную *метаболическую* (т.е. обменную) функцию, поставляя светочувствительным клеткам необходимые для работы вещества, а также осуществляет т.н. *фагоцитоз* — ликвидацию погибших клеток и их фрагментов. Плюс к тому, пигментный эпителий поглощает свет, не захваченный фоторецепторами, препятствуя тем самым его рассеянию и возникновению паразитной подсветки сетчатки.

Между пигментным эпителием и склерой расположен *хороид (хориоидея)* — сосудистая оболочка глаза, реализующая т.н. *трофику* его тканей, т.е. поставляющая им из крови питательные вещества и кислород.

**Адаптационный диапазон зрения.** Каждый из отделов глаза играет критично важную роль в обеспечении т.н. *сенсорной способности* органа, т.е. его чувствительности к свету.

Общее количество света, воздействующего на сетчатку, регулируется зрачком: когда освещение слабое, зрачок расширен; когда сильное — сужен. Важно знать, что зрительная система человека действует в огромном диапазоне уровней фотометрических яркостей стимулов (см. табл. 1.3.1): мы способны различить и мелкую звездочку на ночном небосклоне и птицу, парящую в вышине сверкающего полуденного неба. Иными словами наш *адаптационный диапазон* составляет порядка 10 логарифмических единиц (Boynton, 1979).

Любопытно выяснить, какая доля адаптации приходится на изменение диаметра зрачка? Предположим, что у некоей сцены понизилась освещенность и диаметр зрачка наблюдателя увеличился с 3 мм до 9 мм; применив уравнение площади круга мы легко рассчитаем прирост количества света, падающего на сетчатку — когда диаметр зрачка равен 3 мм, мы имеем:

$$\begin{aligned}A &= \pi r^2 \\A &= \pi (1.5)^2 \\A &= 2.25\pi\end{aligned}$$

Когда диаметр зрачка стал равным 9 мм:

$$A = 20.25\pi$$

Соотношение площади зрачка при тусклом освещении к площади зрачка при ярком составит в нашем случае:

$$20.25\pi / 2.25\pi = 9.0$$

Сей простейший расчет показывает, что на зрачковый рефлекс приходится малая толика адаптации, составляющая едва лишь 1 логарифмическую единицу от 10D-диапазона. Оставшаяся часть адаптации обязана свойствам светочувствительных клеток двух классов — *палочкам* и *колбочкам*, о которых речь пойдет ниже, а также когнитивным механизмам.

Любопытно также, что изменение диаметра зрачка может являться откликом на ряд физиологических факторов, к примеру: положительная эмоциональная реакция расширяет зрачок; скажем, у большинства мужчин зрачок расширяется, когда они видят изображение обнаженной девушки, а у большинства женщин, когда они видят обнаженного мужчину (Hess & Polt, 1960). Размер зрачка также отражает умственные усилия — он расширяется, когда мы сосредоточены (Hess & Polt, 1964; Kahneman & Beatty, 1967). Отметим, что зрачковый рефлекс и, следовательно, диаметр зрачка не контролируются нашим сознанием.

**Фокусировка и аккомодация.** Свет, проникший в глаз ценен только в том случае,



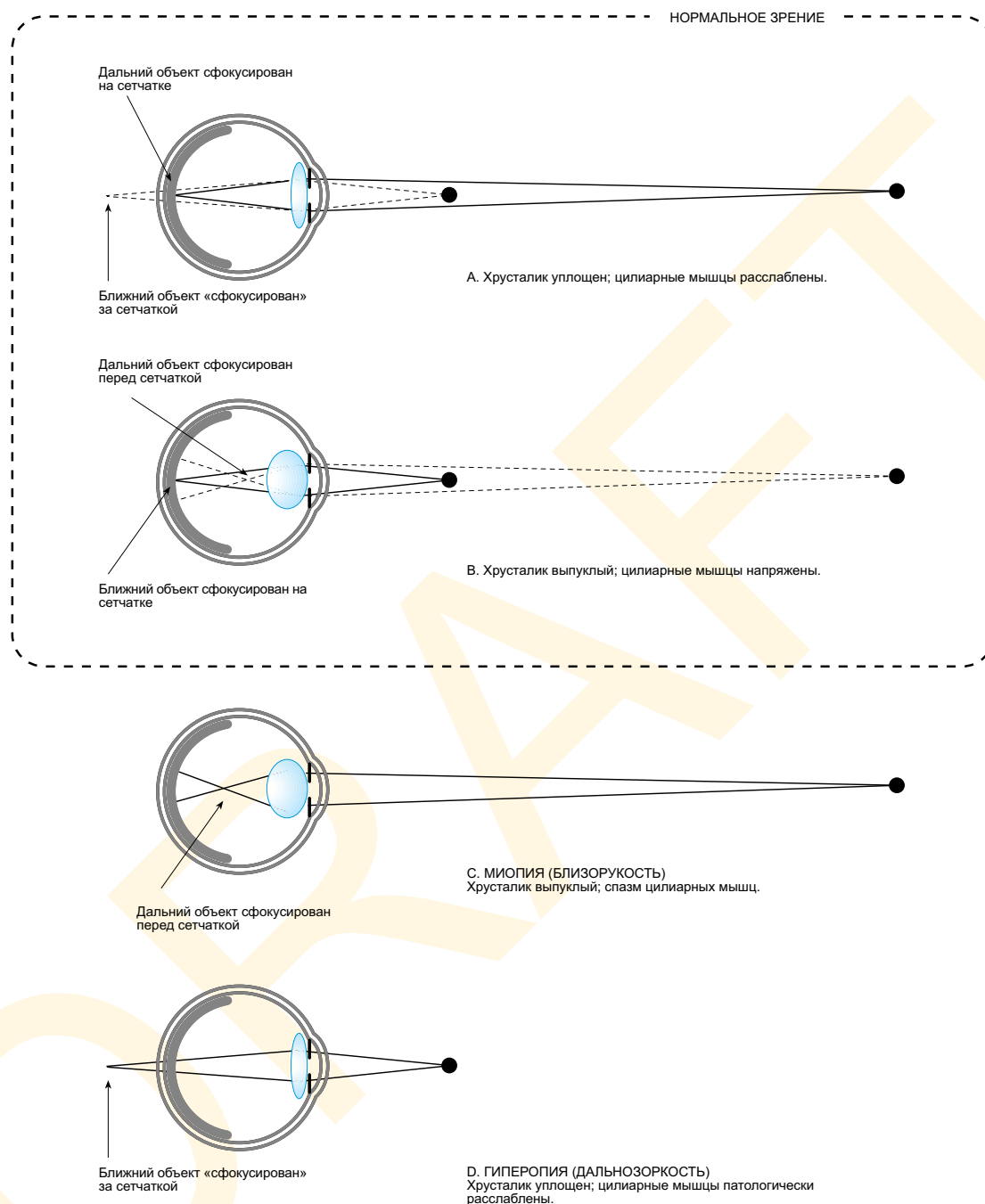
ТАБЛИЦА 1.3.1 Диапазон воспринимаемых яркостей

Стимул	Фотометрическая яркость (cd/m <sup>2</sup> )			
Солнце	10 <sup>10</sup>	Возможное повреждение тканей	Оптимальная острота зрения Палочковое насыщение	
	10 <sup>9</sup>			
	10 <sup>8</sup>			
	10 <sup>7</sup>			
Лампа накаливания 100 Вт	10 <sup>6</sup>	Фотопическое зрение		
Освещенная солнцем бумага	10 <sup>5</sup>			
	10 <sup>4</sup>			
	10 <sup>3</sup>			
Лист бумаги в обычных условиях	10 <sup>2</sup>	Мезопическое зрение		
	10			
	1			Скотопическое зрение
Лист бумаги на рассвете	10 <sup>−1</sup>			
	10 <sup>−2</sup>			
Лист бумаги при свете звезд	10 <sup>−3</sup>			
	10 <sup>−4</sup>			
	10 <sup>−5</sup>			
	10 <sup>−6</sup>			
Порог чувствительности зрительной системы				

если организовал на сетчатке достаточно резкое оптическое изображение, но поскольку диаметр зрачка много больше такового у отверстия камеры-обскуры, необходима принудительная фокусировка световых лучей (см. рис. 1.3.5). Миссию фокусировки по большей части выполняет роговица, а хрусталик доводит дело до конца. Однако функция хрусталика тоже очень важна и состоит в обеспечении *способности переменной фокусировки*.

Оптические свойства хрусталика меняются при изменении его кривизны — сей процесс именуется *аккомодацией*. Чтобы оптическое изображение удаленных объектов на сетчатке было резким, хрусталик должен уменьшить свою кривизну, т.е. стать тоньше (рис. 1.3.5А), что достигается за счет расслабления цилиарных мышц. Чтобы резким стало оптическое изображение близких объектов — кривизна хрусталика должна увеличиться, т.е. он должен стать толще (см. рис. 1.3.5В), что достигается за счет сокращения цилиарных мышц.

Близорукие люди (*миопы*) отлично видят вблизи, но не могут сфокусироваться на удаленных объектах, поскольку их хрусталик слишком выпуклый (при данном размере глазного яблока) (рис. 1.3.5С).



**Рис. 1.3.5** Фокусировка световых лучей хрусталиком. В здоровом глазу (А) уплощенный физиологическим расслаблением цилиарных мышц хрусталик фокусирует на сетчатке световые лучи, поступающие от отдаленных объектов сцены (но не от ближних). Также в здоровом глазу утолщенный физиологическим напряжением цилиарных мышц хрусталик (В) фокусирует на сетчатке световые лучи от ближних объектов (но не от дальних). В глазу больного, страдающего миопией (близорукостью), утолщенный хроническим спазмом цилиарных мышц хрусталик фокусирует световые лучи от дальних объектов перед сетчаткой (С). В глазу больного, страдающего гиперопией (дальнозоркостью), уплощенный хроническим расслаблением цилиарных мышц хрусталик, фокусирует световые лучи от ближних объектов позади сетчатки.

Дальнозоркие (*гиперопы*) хорошо видят вдаль, но не могут сфокусироваться на близких объектах, поскольку кривизна их хрусталика недостаточна (см. рис. 1.3.5D). С возрастом эластичность хрусталика падает и он теряет способность усиливать кривизну, что необходимо при фокусировке на близкие объекты (*пресбиопия* или *старческая дальнозоркость*). Те, кто всю жизнь не носил очков, к 40-50 годам все-таки надевают их, чтобы читать или выполнять какую-то мелкую работу. Те, кто к этому возрасту уже был в очках, часто переходят на бифокальные линзы.

Большая доля света, прошедшего через зрачок, поглощается фоторецепторами сетчатки, что вызывает их *нейрохимический отклик*. Однако некоторая часть света поглощается оптическими средами глаза (роговицей, водянистой влагой, хрусталиком, стекловидным телом), кровеносными сосудами и различными нерелевантными клетками сетчатки (сие вполне закономерно, поскольку кровеносные сосуды и пигментный эпителий достаточно темные, т.е. обладают высоким коэффициентом поглощения). Однако в основном фотоны беспрепятственно проходят сквозь среды глаза, успешно достигая сетчатки и *воздействуя* на нее.

### 1.3.2 Сетчатка

После того, как оптические среды глаза пропустили и направили свет, глаз приступает к исполнению своей второй и критично важной функции — *конверсии энергии фотонов в нервный импульс*: благодаря такой конверсии мозг получает возможность обрабатывать оптическую информацию. Чтобы понять, как сие происходит мы должны коротко пояснить базовое строение мозга и функцию его основных анатомических элементов.

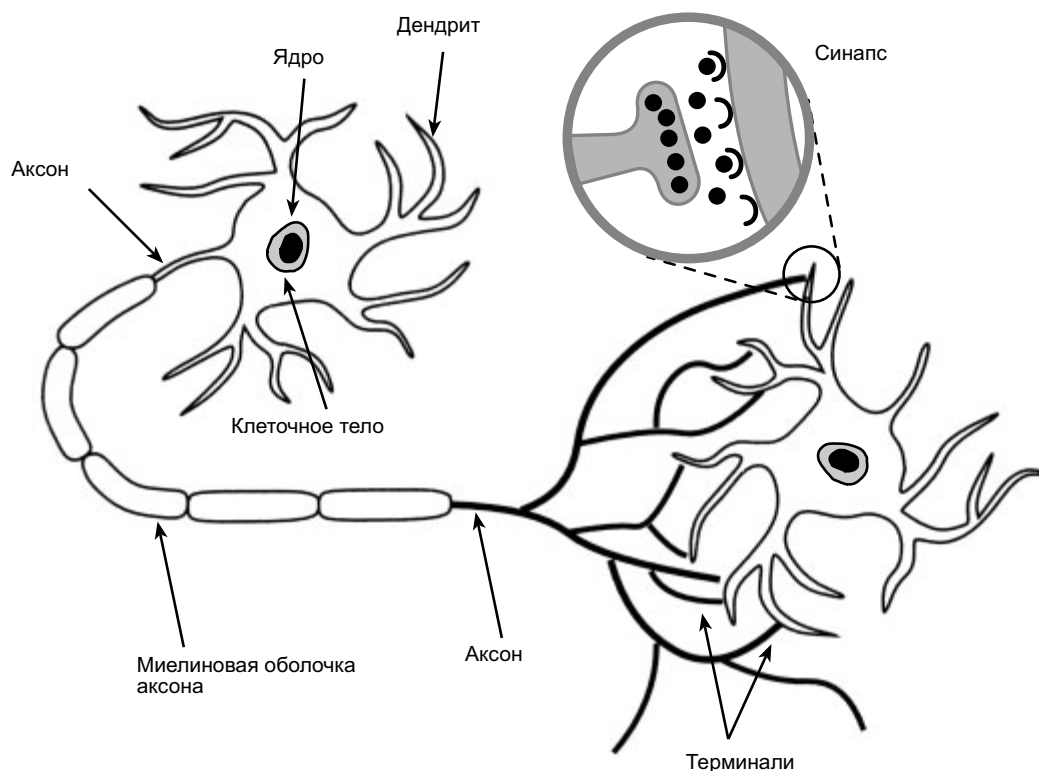
#### 1.3.2.1 Нейроны

Полагают, что главным функциональным элементом мозга<sup>1</sup> является нейрон — специализированная клетка, собирающая сигналы от соседних клеток — нейронов, с которыми она соединена (вход), и раздающая их другим нейронам (выход). Сей процесс интеграции и передачи сигналов реализуется за счет серии сложных биохимических событий внутри нейрона. Фрагмент нейрона показан на рис. 1.3.6., а его функция может быть описана следующим образом:

1. Т.н. *дендриты* — тонкие выросты из тела нервной клетки — собирают *химические* сигналы от других нейронов и конвертируют эти сигналы в электрический потенциал на поверхности тонкой мембраны, в которую заключена клетка. Сей потенциал именуется *градуальным* и представляет собой разность зарядов между внутренней и внешней поверхностью дендрита. Разность эта может плавно меняться в определенном диапазоне, и величина этой разности будет зависеть от того, насколько сильно дендрит стимулирован другими нейронами.

2. *Клеточное тело* состоит из цитоплазмы, покрытой оболочкой. Внутри цитоплазмы находится ядро и разного рода клеточные включения. Мембрана, окружающая клеточное тело, собирает электрические сигналы, поступающие от всех дендритов клетки, кодирует их в виде градуального потенциала и конвертирует его в серию сиг-

<sup>1</sup> Напомним, что глаза — это огромные выросты головного мозга, следовательно и сетчатка — тоже элемент мозга.



**Рис. 1.3.6** Типичный нейрон. Нейрон — это клетка, собирающая градуальные электрические сигналы от своих дендритов и передающая результирующий сигнал другим нейронам в виде дискретных *потенциалов действия*. Потенциалы действия движутся по аксону, покрытому миелиновой оболочкой, и доходят до терминалей, на концах которых высвобождаются нейротрансмиттеры, стимулирующие дендриты следующих нейронов.

налов класса «все или ничего» («да/нет»). Последние обычно называют *потенциалами действия*, *нервными импульсами* или *спайками* (от spike — Прим. пер.).

Далее потенциалы действия движутся по *аксону* нейрона.

3. *Аксон* — длинный (подчас очень длинный), тонкий вырост, по которому биопотенциалы достигают других нейронов. Аксоны большинства нейронов покрыты миелиновой оболочкой, что повышает проводимость. Мощность суммарного сигнала, передаваемого аксоном, обычно выражают в показателях *частоты пульсации* — количестве электрических импульсов в единицу времени (к примеру, пик/сек).

4. *Терминали аксона* — ветвящиеся концы аксона, на которых электрические потенциалы конвертируются обратно в химический сигнал, стимулирующий другой нейрон. Последнее достигается путем высвобождения т.н. веществ — *нейротрансмиттеров* в тонкий зазор (т.н. *синапс*) между терминалью и дендритом следующего нейрона. Нейротрансмиттеры — это химические соединения, способные возбуждать дендриты других нейронов. Мощность сигнала, передаваемого через терминаль определяется количеством высвободившегося нейротрансмиттера.

5. *Синапс* — тонкий зазор, между терминалью аксона одного нейрона и дендритом другого. Нейротрансмиттер, высвободившийся в синапс, очень быстро преодолевает просвет последнего и воздействует на дендрит следующего нейрона, заполняя собой

специальные участки на мембране последнего. В этом месте происходит конверсия химического сигнала, пришедшего от одного нейрона, в электрический сигнал следующего нейрона (см. позицию 1).

Таким образом, нейроны получают входной сигнал от одних нейронов и передают его другим нейронам. Но прежде чем возникнет какая-либо нейронная активность, что-то должно преобразовать световую энергию, поступившую из внешнего мира в форму, потребную нейронам. В зрительной системе такую функцию выполняют *фоторецепторы* — палочки и колбочки — специальные клетки сетчатки, сложным образом стимулируемые световой энергией и выдающие на эту стимуляцию нервный отклик. Позднее мы опишем схему этого процесса.

После того, как оптическая информация была кодирована нейронными откликами, в самой сетчатке происходит начальная обработка этой информации нейронами следующих видов: *горизонтальными*, *биполярными*, *амакриновыми* и *ганглиозными клетками*, каждая из которых собирает сигналы от множества соседних клеток (см. рис. 1.3.7). Детально функцию этих нейронов мы будем обсуждать в последующих главах (в частности в гл. 3 и 4).

Аксоны ганглиозных клеток отправляют информацию из глаза по оптическому нерву в зрительные центры головного мозга. Поскольку головной мозг — это главный отдел *центральной нервной системы* человека, мы говорим, что проводящие пути сетчатки организованы по *центростремительному принципу*. Однако же в ней существуют и *центробежные пути*, организованные за счет т.н. *интерплексиформных клеток* и передающие информацию от ганглиозных клеток назад к фоторецепторам (Lindberg & Fisher, 1986).

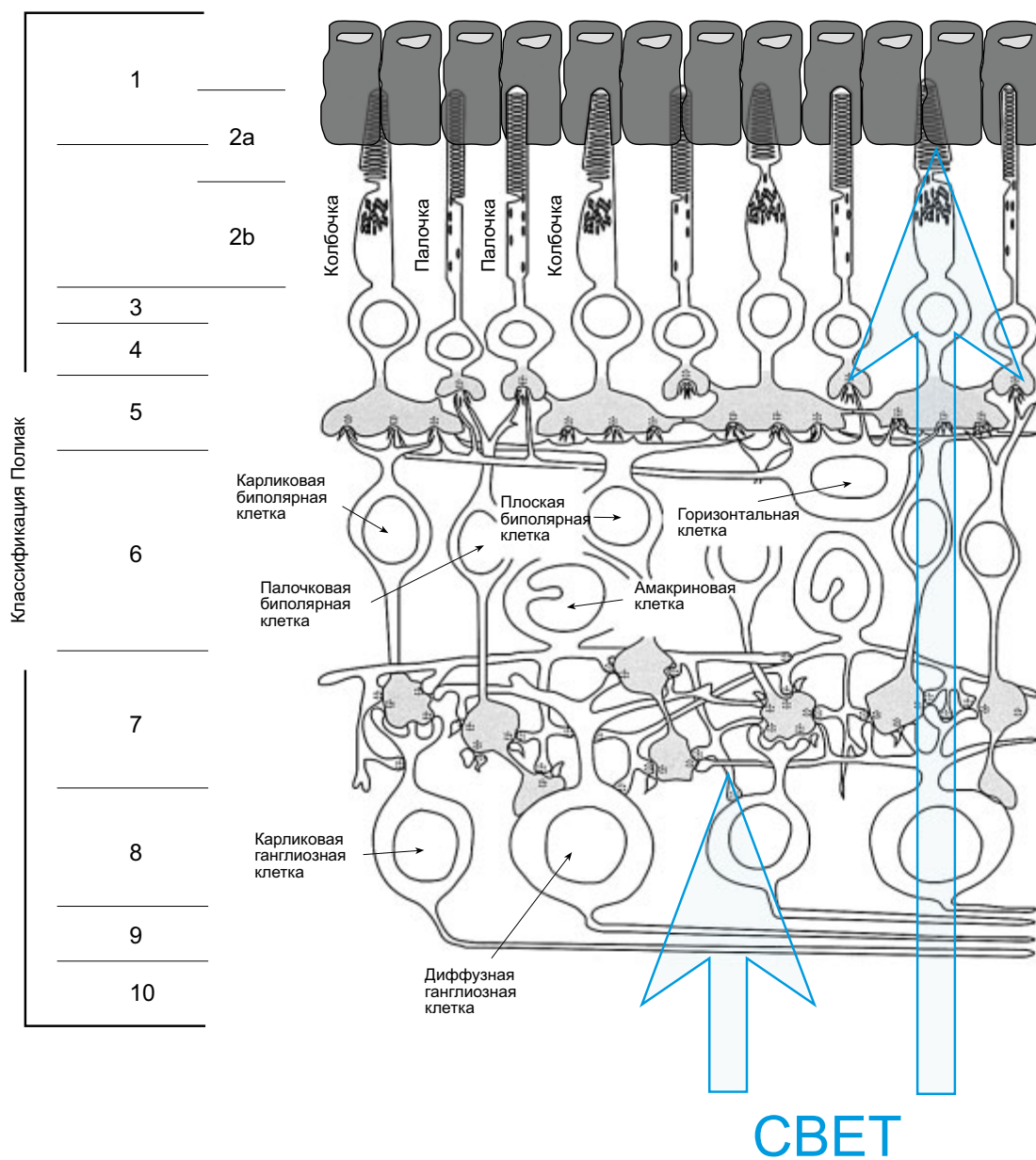
### 1.3.2.2 Гистология сетчатки<sup>1</sup>

Несмотря на то, что человеческий глаз — это совершенный орган детекции света, некоторые аспекты анатомии и гистологии глаза требуют отдельного рассмотрения. К примеру, было бы логично предположить, что рецепторные клетки должны располагаться в *первом* слое сетчатки, дабы первыми встречать входящий свет, однако фактически они расположены в последнем слое (рис. 1.3.7 и 1.3.8). Более того, и у палочек, и у колбочек светочувствительный внешний сегмент обращен назад, то есть является наиболее удаленным от входящего света фрагментом клетки. Причина такого «неудобного» расположения вероятно в том, что ферменты, участвующие в регенерации фотопигмента, находятся в пигментном эпителии, который непрозрачен. Поскольку рецепторные диски должны находиться в тесном контакте с этим критично важным биохимическим ресурсом, они вынуждены располагаться в нижнем (заднем) слое сетчатки.

Еще одной причиной того, что фоторецепторы лежат позади всех остальных клеток, является их высочайший *метаболизм*, т.е. колоссальная интенсивность обменных процессов в этих клетках, что требует их максимальной близости к хороиду (сосудистой оболочке), поставляющему кислород и питание. Продукты обмена и естественного распада фоторецепторов поглощаются т.н. *фагоцитами* пигментного эпителия — клетками, специализированными на утилизации *метаболитов* (т.е. продуктов обмена).

<sup>1</sup> Гистология (от греч. histos — ткань + logos — знание, слово, наука) — раздел биологии, изучающий строение тканей живых организмов. — *Прим. пер.*





**Рис. 1.3.7** Сетчатка человека. От внешнего слоя к внутреннему (сверху вниз): 1 — ретиальный пигментный эпителий; 2a — внешние сегменты фоторецепторов; 2b — внутренние сегменты фоторецепторов; 3 — внешняя границная мембрана; 4 — внешний ядерный слой; 5 — внешний плексиформный слой; 6 — внутренний ядерный слой; 7 — внутренний плексиформный слой; 8 — слой ганглиозных клеток; 9 — нервноволоконный слой; 10 — внутренняя границная мембрана.

Деление на слои и их именование предложено Поляк (Polyak, 1941).

К счастью, все ретинальные клетки и их аксоны высокопрозрачны, благодаря чему качество оптического изображения сцены никак не страдает.

Однако же вдобавок к прозрачным клеткам, лежащим на пути светового потока, имеется множество кровеносных сосудов, выходящих из хороида и питающих сетчатку. Поскольку сосуды достаточно темные, то можно предположить, что они должны быть видимыми (рис. 1.3.9). Однако в норме мы их не замечаем и причина в том, что мозг полностью адаптируется к их присутствию, «заливая» те фрагменты изображения на сетчатке, на которые сосуды отбрасывают тени. В некоторых ситуациях удастся разглядеть эти сосуды, например, когда офтальмолог светит в глаз под необычным углом. Сосуды собственной сетчатки можно намеренно увидеть, если зафиксировать взгляд на равномерно окрашенной поверхности (к примеру, стене или листе бумаги) и при этом подсветить фонариком, скажем, правый глаз со стороны его внешнего угла; если при этом начать аккуратно потряхивать фонарик вверх-вниз, то можно четко разглядеть сосудистый рисунок сетчатки, поскольку тени от сосудов будут падать теперь на те рецепторы, по которым мозг не адаптирован.

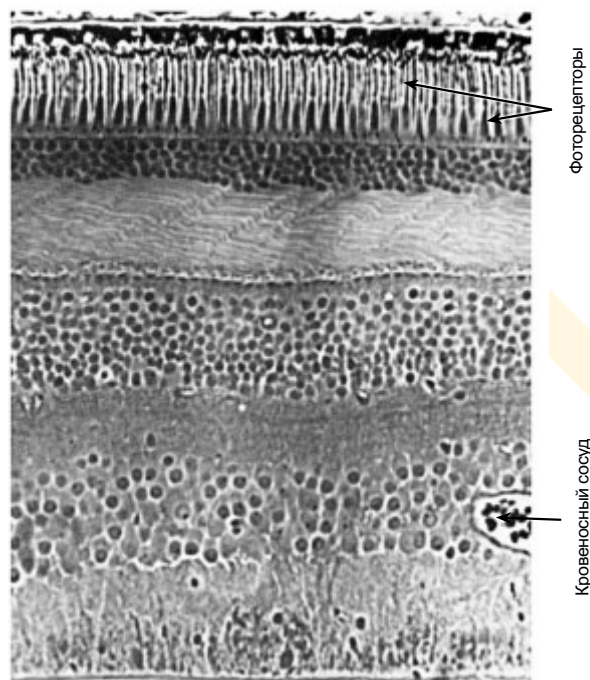


Рис. 1.3.8 Гистологический срез сетчатки человека вне центральной ямки. (Boycott & Dowling, 1969).

**Слой сетчатки.** Как уже было сказано выше, наружным слоем сетчатки является слой пигментного эпителия, не реагирующий на световое воздействие, но выполняющий три очень важных функции: *светопоглощающую*, *трофическую* (питательную), и функцию *утилизации отходов обмена* (Young, 1970, 1971).

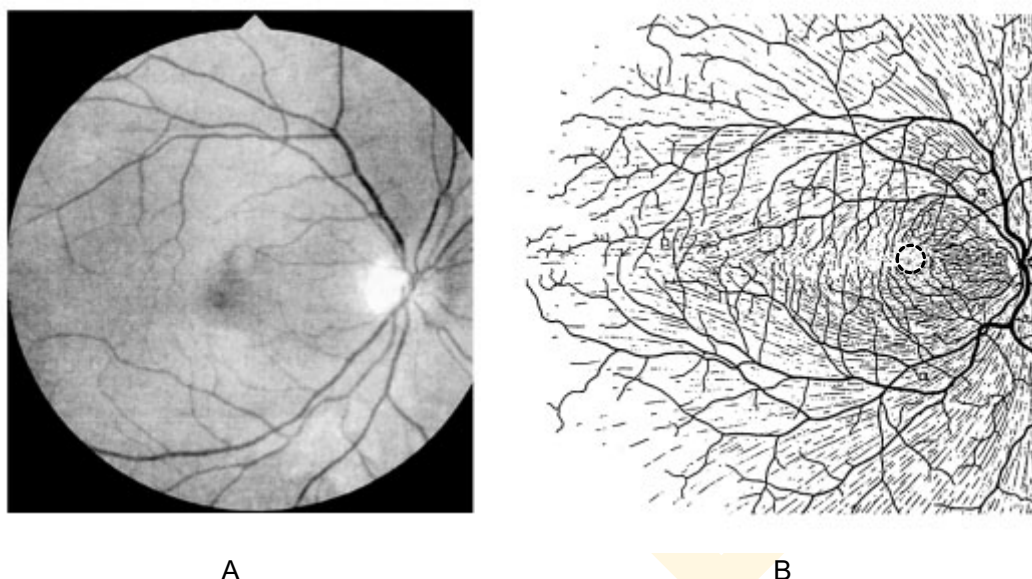
Как будет показано ниже, фоторецепторы (палочки и колбочки) состоят из двух сегментов — внутреннего и внешнего. Внешние сегменты фоторецепторов, содержащие фотопигменты, формируют отдельный слой сетчатки.

Следующий ретинальный слой образован внутренними сегментами фоторецепторов и содержит множество органелл (клеточных включений), не считая клеточных ядер.

Далее располагается т.н. *внешняя граничная мембрана*, образованная т.н. *клетками ретинальной глии* или *мюллеровскими клетками*, отделяющими внутренние сегменты фоторецепторов от их ядер.

Ядра фоторецепторов образуют отдельный сетчаточный слой, именуемый *внешним ядерным слоем*.

Первые ретинальные синапсы появляются во *внешнем плексиформном слое*, состоя-



**Рис. 1.3.9** А — фотография дна правого глаза здорового человека. Сосуды, питающие внутренние слои сетчатки, выходят из оптического диска, расположенного назально от аваскулярной центральной ямки. В — схема кровоснабжения глазного дна макаки. Аваскулярная область (центральная ямка) отмечена пунктиром. Полоски имитируют направление хода волокон в нервно-волоконном слое.

щем из дендритов биполярных клеток, дендритов горизонтальных клеток и терминалей фоторецепторов.

За внешним плексиформным слоем следует (в направлении стекловидного тела, напомним) *внутренний ядерный слой*, состоящий из тел биполярных, горизонтальных и амакриновых клеток.

Второй эшелон синапсов образуется во *внутреннем плексиформном слое*, состоящем собственно из синапсов биполярных, амакриновых и ганглиозных клеток.

Последний *клеточный* слой сетчатки именуется *слоем ганглиозных клеток* или *ганглиозным слоем*.

Здесь следует отметить, что существует два больших класса ганглиозных клеток: мелкие ганглиозные клетки, именуемые *карликовыми* или *миджет-клетками*, или *парво-клетками*. В ганглиозном слое примерно 80% клеток — парво-клетки. Оставшиеся клетки — это крупные ганглиозные клетки, именуемые *зонтичными* или *парасоль-*, или *магно-клетками*. На магно-клетки приходится 10%. Оставшиеся 10% — это т.н. *малые бираслоенные ганглиозные клетки*.

Аксоны ганглиозных клеток образуют *слой нервных волокон* или *нервно-волоконный слой* и покидают глаз, в виде плотного нервно-волоконного пучка — *оптического нерва* (вторая пара черепно-мозговых нервов), завершающегося синапсами в т.н. *среднем мозге*. Любопытно, что аксоны ганглиозных клеток внутри глаза не покрыты миелиновой оболочкой (как говорят, не миелинизированы) и поэтому практически прозрачны. Однако по выходе из глаза они покрываются миелиновой оболочкой, повышающей проводимость нервных импульсов.

Финальный слой сетчатки называется *внутренней граничной мембраной*, которая выступает в роли интерфейса между сетчаткой и стекловидным телом.

**Латеральные взаимодействия.** Ключевой особенностью гистофизиологической организации сетчатки является ее способность к т.н. *латеральным* (боковым, горизонтальным) взаимодействиям: горизонтальные и амакриновые клетки собирают информацию в горизонтальном направлении сетчатки, то есть в рамках того или иного слоя, (но не междуслойно), организуя тем самым пространственную размерность обработки оптического изображения.

Поскольку диаметр оптического нерва и, следовательно, количество аксонов ганглиозных клеток в него входящих ограничено анатомией черепа, далеко не вся оптическая информация, принятая из света сетчаткой, передается в собственно мозг: несмотря на то, что в сетчатке одного глаза более 100 млн. фоторецепторов, на них приходится всего 1 млн. ганглиозных клеток, обеспечивающих высокую степень т.н. *нейронной конвергенции* (Osterberg, 1935; Curcio & Allen, 1990).

**Слепое пятно.** Еще одной любопытной особенностью анатомии глаза является то место, где аксоны ганглиозных клеток собираются в оптический нерв и покидают глазное яблоко. Эта область именуется *оптическим диском* или *слепым пятном*. Расположено слепое пятно на горизонтальной плоскости под углом в 12-15° от центральной ямки к носу, и характеризуется тем, что в нем совсем нет фоторецепторов. Однако нельзя буквально говорить о слепоте в этом месте (за исключением особых ситуаций): во-первых, потому, что оптические диски расположены так, что рецепторы одного глаза регистрируют то, чего не видел второй глаз; во-вторых, зрительная система заливает этот участок как говорят соответствующим *сенсорным качеством*, к примеру таким же как давеча залила тени от сосудов. Мы до сих пор толком не знаем, как это происходит, но работает сие весьма эффективно.

Убедиться в наличии слепого пятна, можно если поднести нашу книгу близко к лицу, закрыть правый глаз и зафиксировать взгляд левого на правом верхнем крестике рис. 1.3.10; затем начать плавно отдалять страницу, продолжая удерживаться на крестике. Когда кружок попадет на слепое пятно, он станет невидимым. Стоит обратить внимание на то, что зрительная система заливает область слепого пятна белым фоном страницы. Сие удивительный факт, поскольку совершенно неочевидно, как зрительная система может «видеть» белое там, где его нет, не замечая при этом точку, которая есть.

Высшие мозговые процессы (идушие в зрительной коре), заполняют участок поля зрения, приходящийся на слепое пятно, соответствующей информацией, которая впоследствии осознается. Мы можем убедиться в этом, используя нижний правый кре-



**Рис. 1.3.10** Обнаружение слепого пятна. Чтобы убедиться в его наличии, нужно закрыть правый глаз, сосредоточить взгляд левого на верхнем крестике и постепенно отодвигать рисунок от себя. На расстоянии порядка 20-25 см левый кружок исчезнет. Если проделать то же с нижней частью рисунка, то можно наблюдать иллюзию «заливки» разреза нижней линии.



стик и разорванную линию слева от него: вновь закроем правый глаз и зафиксируем взгляд левого на правом нижнем крестике и вновь начнем плавно отодвигать страницу: когда разрыв «сядет» на слепое пятно, то станет невидимым — иллюзия линии заполнит зону разрыва. Итак, мы убедились в том, что принцип WYSIWYG срабатывает не всегда!

**Центральная ямка** (*fovea*) — высокоспециализированный участок сетчатки, которым, собственно, мы и читаем эти строки. Угловой размер центральной ямки по разным данным составляет от  $1.2^\circ$  до  $2^\circ$  (Palmer, 1999; Schwartz, 2004), что, к сравнению, примерно равно двойному угловому размеру полной Луны (Pirenne, 1967).

Центральная ямка — это морфологический и функциональный базис высокой *остроты зрения* (*разрешающей способности зрения*) человека. Давайте посмотрим на литеру «Е» рис. 1.3.11 и убедимся, что она идеально четкая; при этом, не сводя пристальный взгляд с нее, сосредоточимся на других буквах рисунка — они слегка размыты. Так происходит потому, что изображение литеры «Е» образуется на центральной ямке сетчатки, тогда как изображения угловых литер — на ее периферии.

У центральной ямки множество особенностей: во-первых, в ней находятся только колбочки, палочек нет; во-вторых, плотность залегания фовеальных колбочек выше, чем в каком-либо ином районе сетчатки (см. рис. 1.3.12). Все нейронные элементы сетчатки в этом месте рассредоточены, дабы свет мог непосредственно воздействовать на пигментсодержащие внешние сегменты колбочек. Следствием рассредоточения является т.н. *фовеальная лунка* или *фовеола* (рис. 1.3.13).

Отметим, что васкулярная (сосудистая) сеть покрывает всю сетчатку, за исключением именно центральной ямки, которая, как говорят, *аваскулярна*. Сие предотвращает рассеяние света сосудами, доводя до максимума разрешающую способность центральной ямки.

Трофику (питание) фовеальных и внефовеальных колбочек осуществляет хороид (сосудистая оболочка глаза).

Итак:

- отсутствие палочек;
- максимальная плотность залегания колбочек;
- рассредоточение внутренних ретинальных элементов по сторонам;
- отсутствие сосудов —

все это предпосылки великолепной разрешающей способности центральной ямки сетчатки. Итак, центральная ямка, несмотря на то, что физически мала, играет важнейшую роль в зрительном восприятии.

**Макула.** Центральную ямку окружает и покрывает образование, именуемое *желтым пятном* (*macula lutea*) или *макулой*, и содержащее желтый нечувствительный пигмент. Сей пигмент поглощает коротковолновые фотоны (пик поглощения приходится на 460 нм) и за счет этого уменьшает рассеяние «синей» составляющей оптического изображения на сетчатке, чем в определенной степени компенсирует хроматические аберрации (Wald, 1945). Поглощая жесткий ультрафиолет, желтое пятно защищает ретинальную ткань от ожога.<sup>1</sup>

<sup>1</sup> К слову сказать, сегодня полным ходом идут исследования по разработке пищевых добавок, содержащих лютеин и зеаксантин (в свое время были обнаружены в составе макулярного пигмента) и направленных на профилактику и/или лечение т.н. *возрастной макулярной дегенерации*, которая в старости может привести к полной слепоте (Gale, 2003).



F

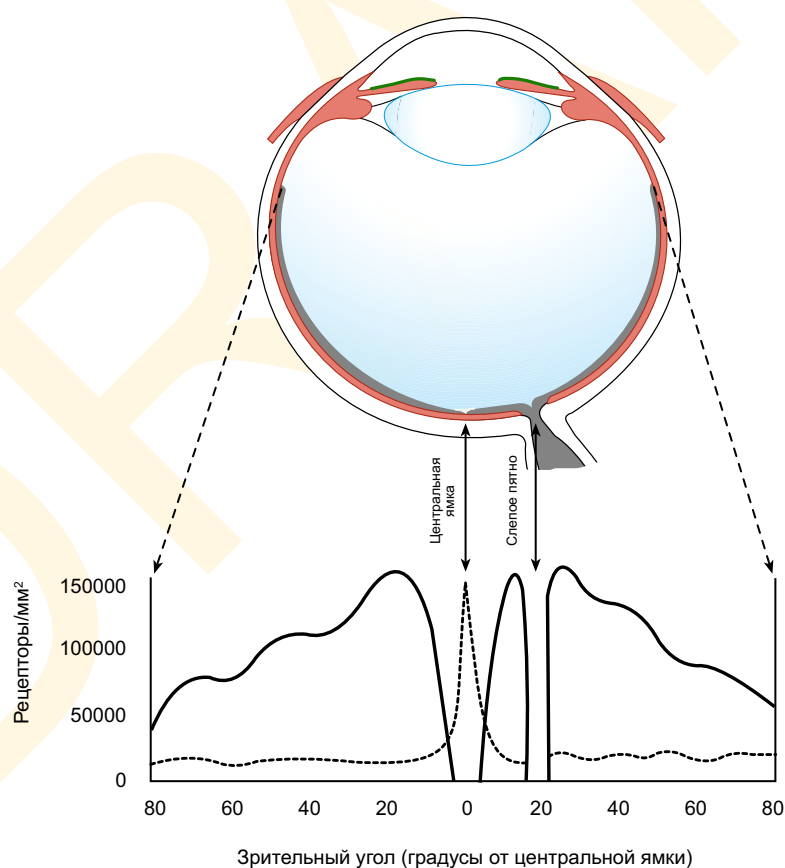
B

E

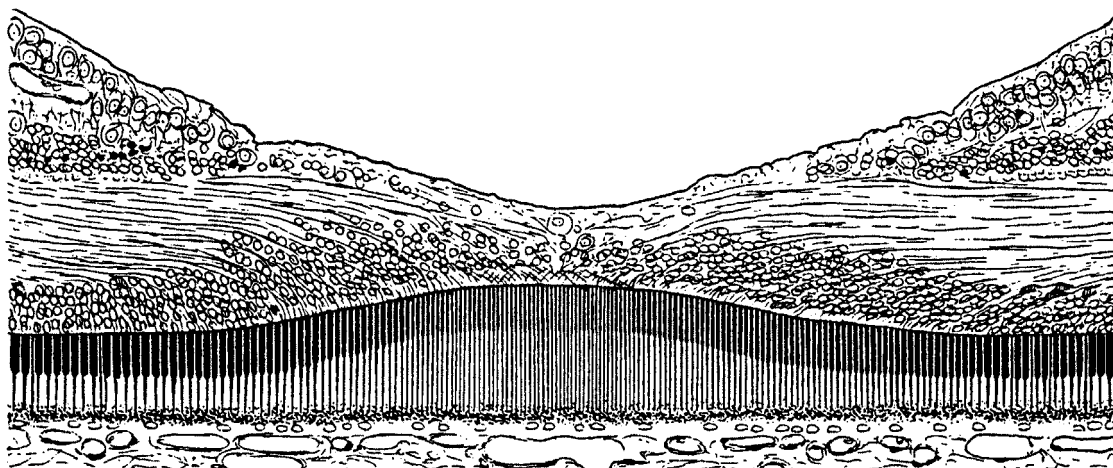
H

M

**Рис. 1.3.11** Если сосредоточить взгляд на литере «Е», но при этом прислушаться к своим ощущениям в отношении угловых литер, то легко обнаружить, что последние не столь четки: острота зрения максимальна в центральной ямке сетчатки и падает на периферии.



**Рис. 1.3.12** Распределение палочек (сплошная линия) и колбочек (пунктирная линия) в сетчатке глаза человека. Отметим, что центральная ямка заселена исключительно колбочками и что палочек на периферии сетчатки больше, чем колбочек.



**Рис. 1.3.13** Схематичный гистологический срез центральной ямки сетчатки человека. Хорошо видно, что в результате рассредоточения внутренних ретиальных элементов образуется лунка, позволяющая свету почти беспрепятственно достигать колбочек.

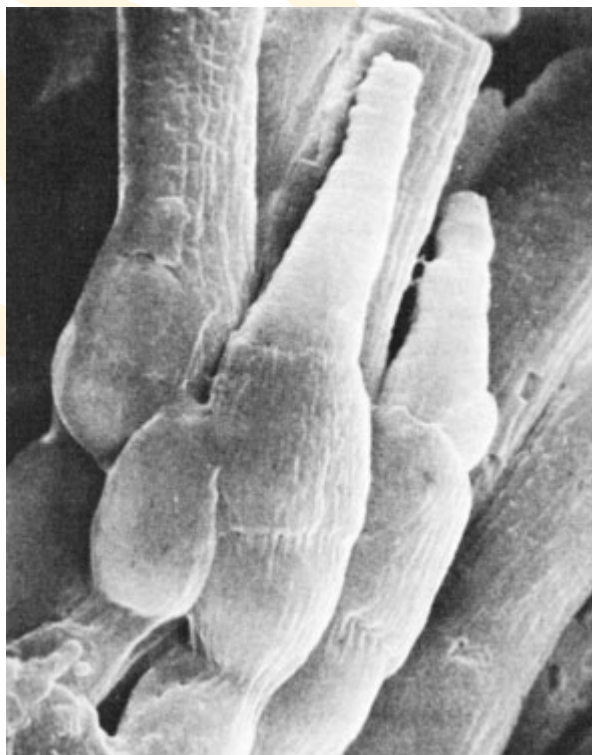
### 1.3.2.3 Фоторецепторы

Существуют два класса фоторецепторов — *палочки* и *колбочки*. Названия они получили благодаря своей форме, что хорошо видно по электронной микрофотографии сетчатки рис. 1.3.14.

**Палочки** обычно длиннее колбочек и имеют затупленные концы. Палочек много больше (120 миллионов), они предельно чувствительны к свету и заполняют собой всю сетчатку, за исключением ее центральной части (см. рис. 1.3.12). Максимальная плотность залегания палочек приходится на область примерно в  $20^\circ$  вокруг центральной ямки сетчатки и составляет порядка 150000 палочек/ $\text{мм}^2$  (см. рис. 1.3.12).

Палочки используются зрительной системой при очень низких уровнях освещенности сцен (т.н. *скотопические условия*): ночью, в сумерках, либо в слабо освещенных помещениях.

**Колбочки** короче палочек, толще и с конусообразными концами. Колбо-



**Рис. 1.3.14** Сканирующая электронная микрофотография палочек (на заднем плане) и колбочек (на переднем плане) сетчатки человека. Внешние сегменты палочек представляют собой цилиндры, тогда как внешние сегменты колбочек имеют коническую форму.

чек много меньше — «всего лишь» 6-8 миллионов; они менее чувствительны к свету и в основном расположены в центре сетчатки, хотя некоторые обнаруживаются и на ее периферии (см. рис. 1.3.12).

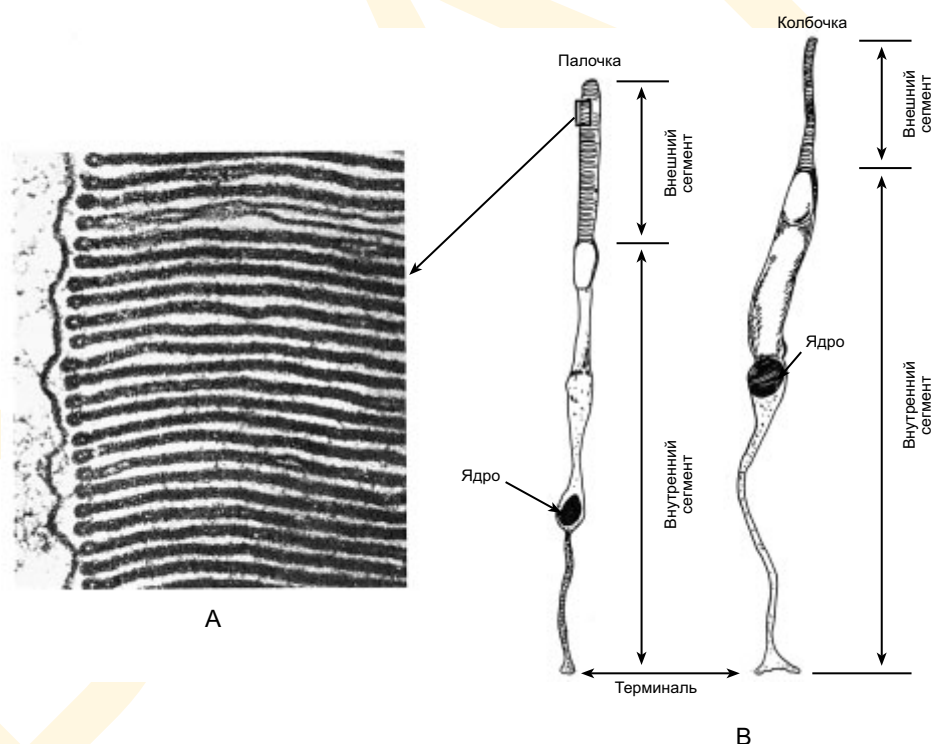
Колбочки отвечают за наши зрительные ощущения в нормальных условиях освещения (т.н. *фотопические условия*) и за все ощущения цветковые.

И палочки, и колбочки состоят из двух частей (рис. 1.3.15): *внутреннего сегмента*, содержащего *ядро* и прочие *клеточные включения (органеллы)*, и *внешнего сегмента*, содержащего миллиарды светочувствительных пигментных молекул. Молекулы пигмента упакованы в тонкие мембранные диски, расположенные в стопку и перпендикулярно длиннику внешнего сегмента.

Палочковый пигмент имеет белковое строение и называется *родопсином*; колбочковых же пигментов три вида и именуются они соответственно как *цианолэйб*, *хлоролэйб* и *эритролэйб* (т.е. сообразно своей спектральной чувствительности — см. далее).

Отметим, что важнейшим является вопрос о том, *как именно* фоторецепторы преобразуют электромагнитную энергию фотонов в нейронную активность (нервные импульсы). Сие очень сложный и совершенно уникальный процесс, который к сегодняшнему дню уже неплохо изучен (и о котором речь пойдет чуть позднее).

**Родопсин** — светочувствительный пигмент (фотопигмент), заключенный в дисках внешнего сегмента палочек; в одном диске находится примерно 10000 молекул родоп-



**Рис. 1.3.15** Морфология палочек и колбочек. Внешние сегменты фоторецепторов содержат диски с вмонтированными в них молекулами пигмента. Поглощение света этими молекулами приводит в действие каскадную биохимическую реакцию, исходом которой являются электрические изменения на мембране внутреннего сегмента. Электрические сигналы с внутреннего сегмента поступают на терминали, преобразуются и передаются другим нейронам. (Allen, 1967).

сина. Поскольку каждая палочка несет в себе порядка 1000 дисков, а всего в глазу 120 миллионов палочек, на глаз приходится примерно  $10^{15}$  молекул родопсина (Boynton, 1979), каждая из которых способна поглотить один фотон. Отметим, что огромное число молекул родопсина придает глазу человека невероятную светозахватывающую способность и обеспечивает наше ночное зрение.

На рис. 1.3.16 показана кривая спектра поглощения родопсина и способ ее получения: если, предположим, количество фотонов монохроматического света<sup>1</sup> упавших на некую емкость с родопсином известно, то отношение пропущенного света к упавшему легко вычислить; эксперимент необходимо повторить по всем длинам волн видимого спектра (например, 400 нм, 401 нм, 402 нм и т.д.), а полученный результат вычертить в виде кривой, демонстрирующей долю пропущенного света как функцию от длины волны (рис. 1.3.16B).

Те фотоны, что упали на родопсин, но не прошли сквозь него — поглотились. Следовательно, всегда есть возможность построить кривую спектра пропускания родопсина, которая будет обратна кривой его спектра поглощения (рис. 1.3.16C). Отметим, что пик поглощения родопсина приходится на 507 нм.

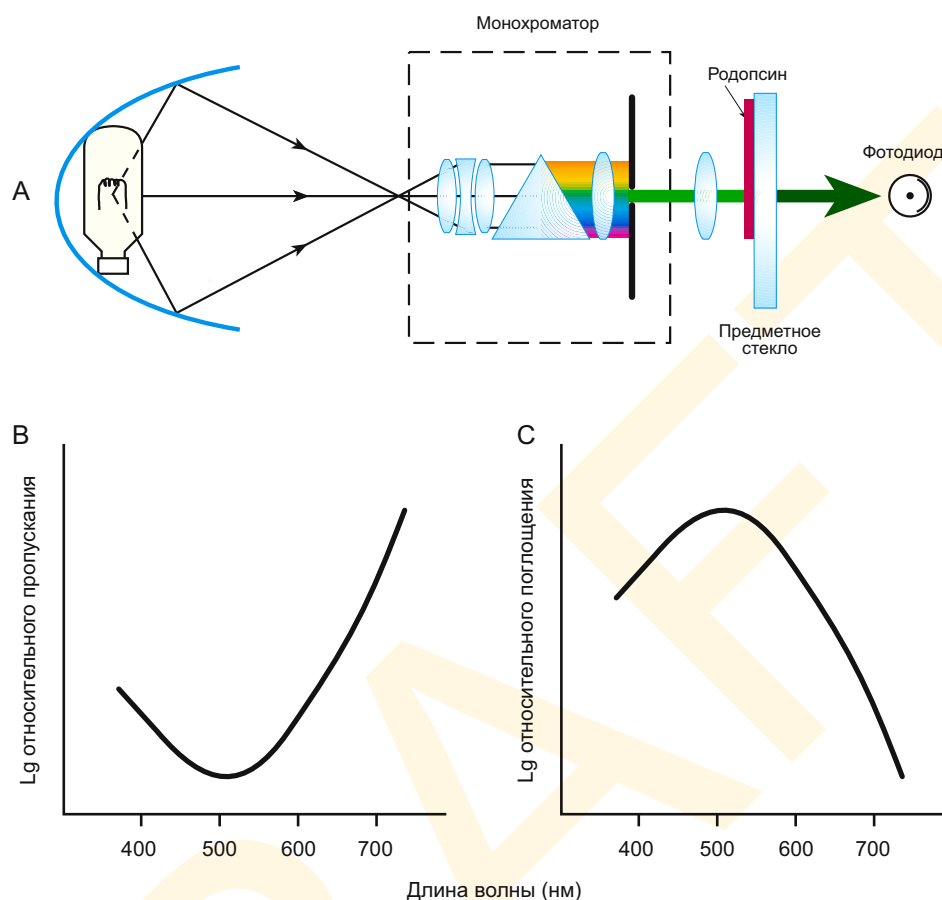
Далее. Молекула родопсина отбеливается, когда поглощает свет, т.е., как говорят, теряет оптическую активность. Для отбеливания одной молекулы родопсина требуется всего лишь один фотон (Hetch et al, 1942). Когда молекула родопсина находится в отбеленном состоянии, она не может поглотить следующий фотон и пропускает его. Однако через некоторое время отбеленная молекула спонтанно возвращается в исходное, т.е. неотбеленное состояние. Вероятность того, что молекула отбеленного родопсина вернется в исходное состояние в течение 5 мин составляет 0.50, то есть 50% (Rushton, 1965a). Следовательно, если какое-то количество родопсина было отбелено, то половина от этого количества *регенерирует* (т.е. возвращается в исходное состояние) через 5 мин. Мы говорим при этом, что время *полурегенерации* родопсина составляет 5 мин.

Теперь давайте детально рассмотрим спектр поглощения родопсина, который несет информацию о том, какова *вероятность* поглощения фотона с той или иной длиной волны (рис. 1.3.16C). Мы видим, что вероятность поглощения фотона с длиной волны в 507 нм максимальна. Сие следует из квантовой механики: молекула родопсина и фотон с длиной волны 507 нм, что называется, «хорошо подогнаны друг к другу», благодаря чему растет вероятность поглощения.

Однако сказанное вовсе не означает, что родопсин не в состоянии поглощать фотоны иных длин волн: фотоны иных частот безусловно поглощаются, но с меньшей вероятностью. Сие проиллюстрирует следующий пример: сравним воздействие от 1000 фотонов с длиной волны 507 нм и 1000 фотонов с длиной волны 580 нм, падающих на один и тот же родопсиновый диск внешнего сегмента палочки. Предположим, что кривая поглощения родопсина дает вероятность в 0.20 по фотону «507» и 0.10 по фотону «580». Простое умножение 1000 фотонов на вероятность поглощения покажет, что фотоны «507» отбелят 200 молекул родопсина, тогда как фотоны «580» — 100.

Если интенсивность потока фотонов «580» удвоить, то вероятность поглощения вырастет вдвое и сравняется с таковой у фотонов «507». Следовательно, 1000 фотонов с длиной волны 507 нм окажут на палочку точно такое же воздействие что 2000 фотонов с длиной волны в 580 нм.

<sup>1</sup> Спектральное распределение энергии монохроматического света представляет собой узкую полосу.



**Рис. 1.3.16** А — принцип определения спектральных коэффициентов пропускания и поглощения родопсина: монохроматы разных длин волн подаются на емкость с родопсином; с противоположной стороны емкости находится спектрометрический прибор, фиксирующий энергию света по длинам волн видимого спектра. В — кривая спектрального коэффициента пропускания родопсина. С — кривая спектрального коэффициента поглощения родопсина; хорошо видно, что она обратная кривой пропускания.

Отметим, что когда фотон поглощен, вся информация о длине его волны оказывается утраченной. Сие есть принцип *универсальности*: не имеет значения, фотоны какой длины волны отбелили данное количество молекул родопсина — 580 или 507 нм — результирующий эффект один и тот же.

**Скотопическая спектральная чувствительность.** Способность к детекции стимулов в скотопических условиях определяется кривой поглощения родопсина. Сие можно продемонстрировать построением кривой *скотопической спектральной чувствительности* наблюдателя (т.е. чувствительности как функции от длины волны). Т.е. нам предстоит определить минимальное количество энергии стимулов разных длин волн, при котором их начинает воспринимать испытуемый.

⇒ Минимальная энергия стимула, при которой наблюдатель начинает воспринимать этот стимул, именуется *порогом чувствительности* к данному стимулу или коротко — *порогом*.

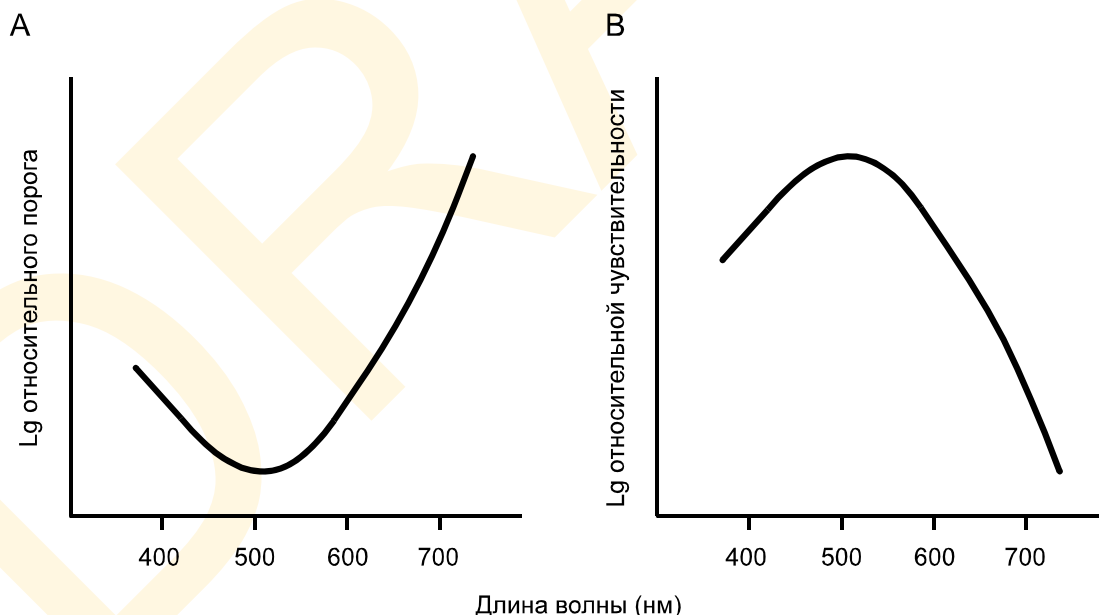


Однако вначале следует попросить испытуемого пройти темновую адаптацию в течение 45 минут (т.е. побыть это время в полностью затемненном помещении), дабы достичь максимальной регенерации родопсина.

На рис. 1.3.17 показаны кривая порога и кривая чувствительности как функции от длины волны, полученные в скотопических условиях. Мы видим, что кривая чувствительности — это попросту обратная кривая порога. Отметим, что кривая скотопической спектральной чувствительности обладает точно такой же формой, что и кривая спектра поглощения родопсина (Wald, 1945).<sup>1</sup> Сходство форм кривых указывает на то, что скотопическая спектральная чувствительность зрения человека определяется поглощательной способностью родопсина.

**Абсолютная чувствительность зрения.** Напомним, что для активации одной палочки достаточно поглощения одного фотона. Десять активированных палочек могут активировать одну ганглиозную клетку, что ведет к *детекции* (т.е. зрительному обнаружению) стимула. Однако же стимул, состоящий всего лишь из 10 фотонов невидим, поскольку часть из них поглотится тканями, окружающими фоторецептор (либо отразится от них) и не достигнет родопсина. Действительно, едва лишь 20% фотонов, упавших на сетчатку, поглощаются родопсином (Hetch et al, 1942). Таким образом чтобы 10 поглощенных фотонов привели к детекции стимула, они должны появиться в одно время в одном месте. Сие есть несовершенство механизмов пространственной и временной суммации в зрительной системе (о чем мы поговорим в следующих главах).

**Колбочковые фотопигменты.** Как уже было сказано выше, в сетчатке неаномального человеческого глаза присутствуют три колбочковых фотопигмента: *цианолэйб*,



**Рис. 1.3.17** А — скотопический порог как функция от длины волны. В — скотопическая чувствительность как функция от длины волны.

<sup>1</sup> На самом деле кривая скотопической спектральной чувствительности слегка отличается от кривой поглощения родопсина благодаря поглощению оптическими средами глаза.

хлорола́йб и эритрола́йб, пики поглощения которыми приходятся соответственно на 426, 530 и 557 нм (рис. 1.3.18А). Каждая колбочка содержит только один вид фотопигмента, и исходя из этого их делят на три функциональных класса: колбочки, содержащие цианола́йб, относят к коротковолновочувствительным колбочкам (*SWS*- или *S*-колбочки); содержащие хлорола́йб, относят к классу средневолновочувствительных колбочек (*MWS*- или *M*-колбочки) и колбочки, несущие эритрола́йб — к классу длинноволновочувствительных (*LWS*- или *L*-колбочки).<sup>1</sup> Колбочковые фотопигменты восстанавливаются быстрее, чем родопсин: период их полурегенерации составляет порядка 1.5 минут (Rushton, 1963b).

**Распределение колбочек** в сетчатке радикально отлично от палочкового: как уже было сказано, колбочки концентрируются в центральной ямке, где плотность их залегания составляет порядка 150000/мм<sup>2</sup>, что эквивалентно максимальной плотности залегания палочек вокруг центральной ямки. Отметим, что хотя плотность залегания колбочек вокруг центральной ямки прогрессивно падает по мере удаления от последней, они встречаются по всей сетчатке (рис. 1.3.19) и только небольшая доля от общего числа колбочек сосредоточена собственно в *fovea* (Boynton, 1979). Наконец отметим, что отношение числа *L*-колбочек к числу *M*-колбочек в центральной ямке сетчатки варьирует от особи к особи в диапазоне от 1:1 до 3:1 (Roorda & Williams, 1999).

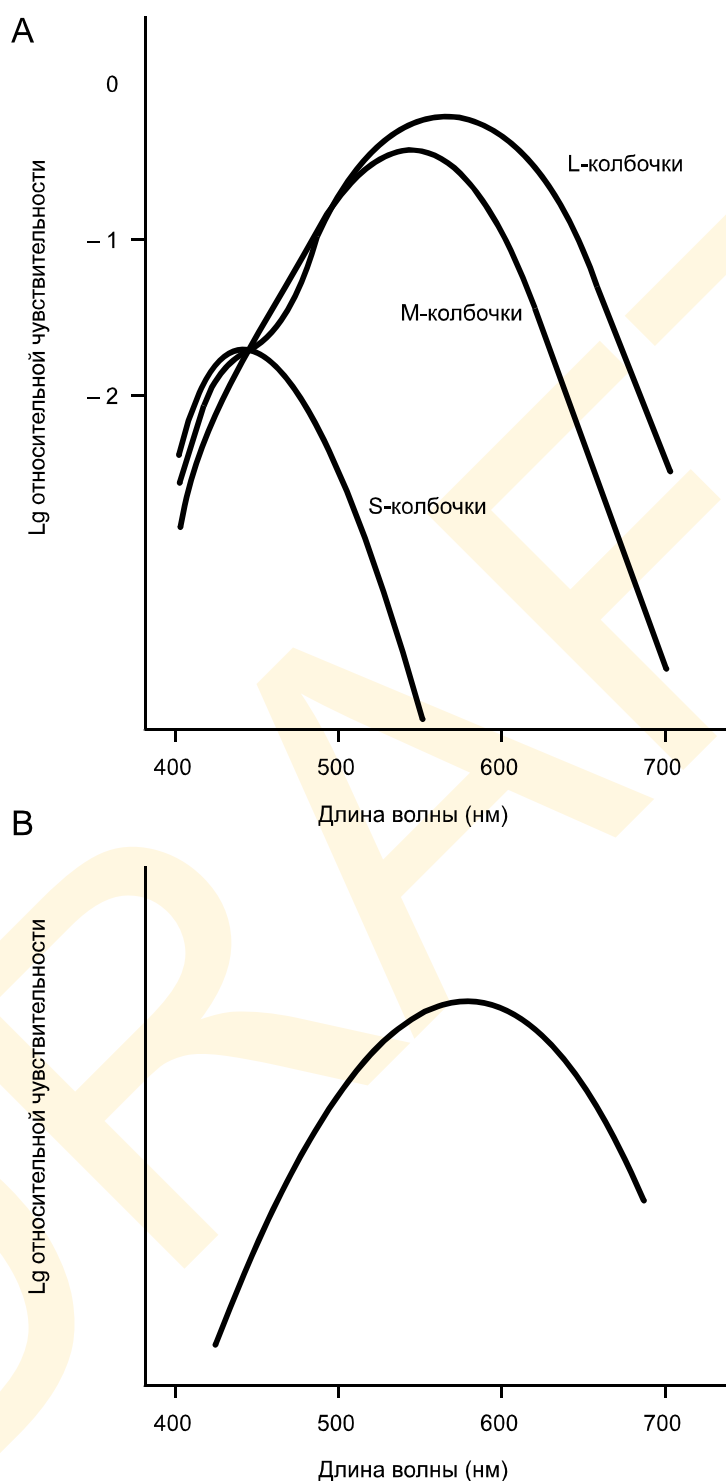
Распределение *S*-колбочек (т.е. коротковолновочувствительных) радикально отлично от распределения колбочек *L* и *M*: *S*-колбочек не только меньше (и составляет 5-10 % от общей колбочковой популяции), но их нет в центральной ямке (Curcio et al, 1991; Roorda et al, 2001) — пик плотности их пространственного залегания лежит за пределами *фовеа* и поэтому мы плохо распознаем мелкие синие объекты.

*S*-колбочки в сетчатке распределены весьма рассеянно (рис. 1.3.20) и причиной тому хроматические аберрации, вызываемые оптическими средами глаза и проявляющие себя тем, что коротковолновый («синий») компонент оптического изображения на сетчатке оказывается сильно размытым (Wandell, 1995).

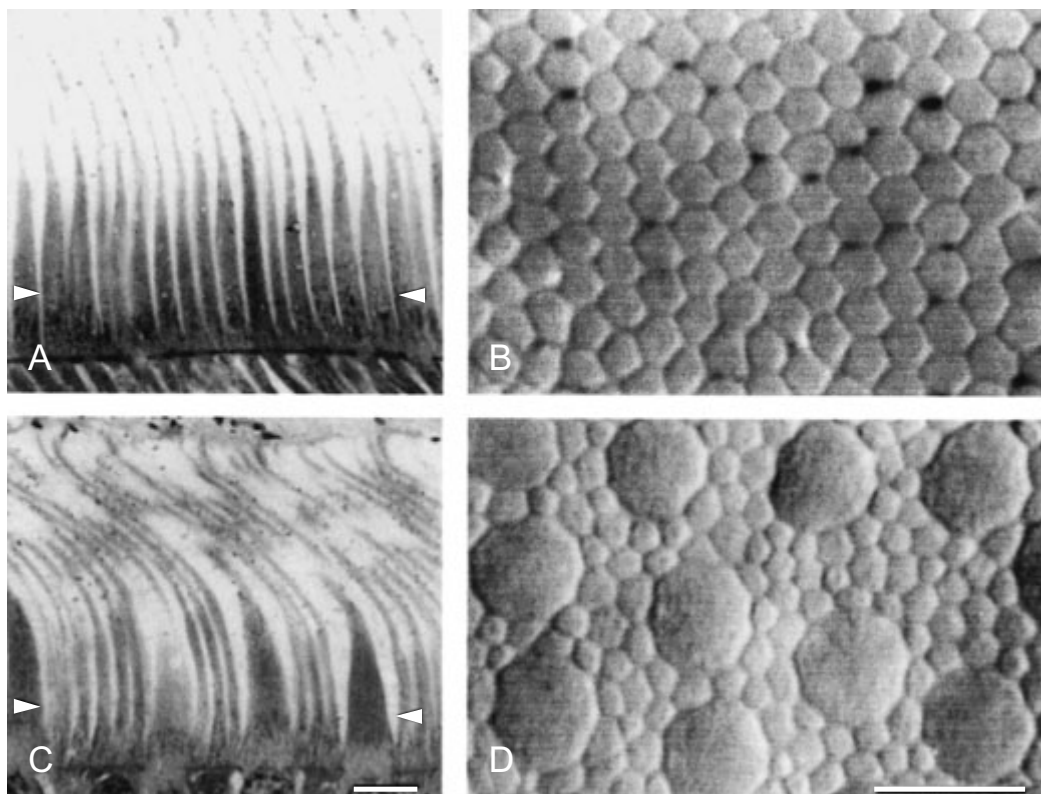
**Фотопическая спектральная чувствительность.** Кривую фотопической спектральной чувствительности получают тем же способом, что и скотопическую кривую с тем только отличием, что измерения проводятся в хорошо освещенных условиях. Фотопическая чувствительность как функция от длины волны дана на рис. 1.3.18В. Отметим, что пик чувствительности приходится на 555 нм (Wald, 1945). Очевидно также, что кривая фотопической спектральной чувствительности представляет собой сумму спектров поглощения *L*- и *M*-колбочек; вклад *S*-колбочек в эту кривую очень мал, если вообще имеет место (Cavanagh et al. 1987).

**Фотохроматический интервал.** На рис. 1.3.21 показаны кривые скотопической и фотопической спектральных чувствительностей глаза, нанесенные на один общий график. Пороги были определены после того, как наблюдатель прошел темновую адаптацию, гарантирующую то, что родопсин и колбочковые фотопигменты полностью восстановлены. Кривые представляют абсолютные чувствительности, полученные в ситуациях максимально достижимых чувствительностей; пороги представлены по каждой из длин волн: порог ахроматического скотопического зрения и порог хроматического цветового зрения.

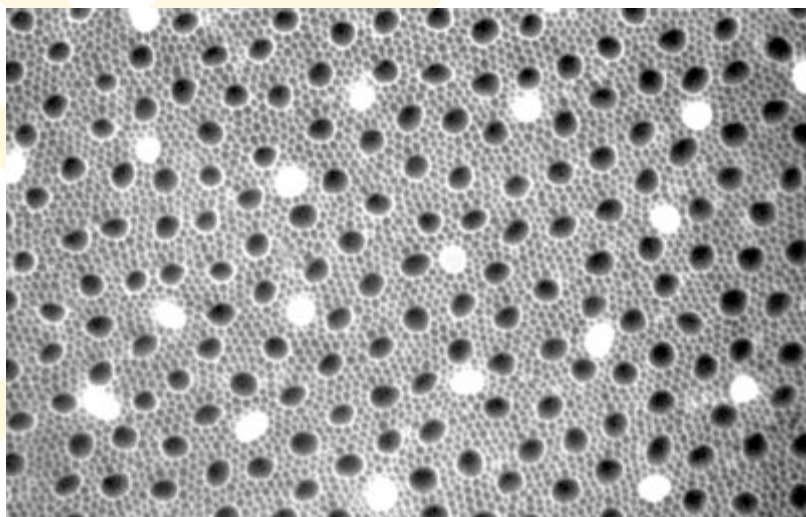
<sup>1</sup> Колбочки трех классов ошибочно (и очень часто) именуют как «синие», «зеленые» и «красные» колбочки. Поскольку колбочки не имеют таких окрасок и совершенно необязательно, что выдают таковые сигналы — данная терминология неприемлема.



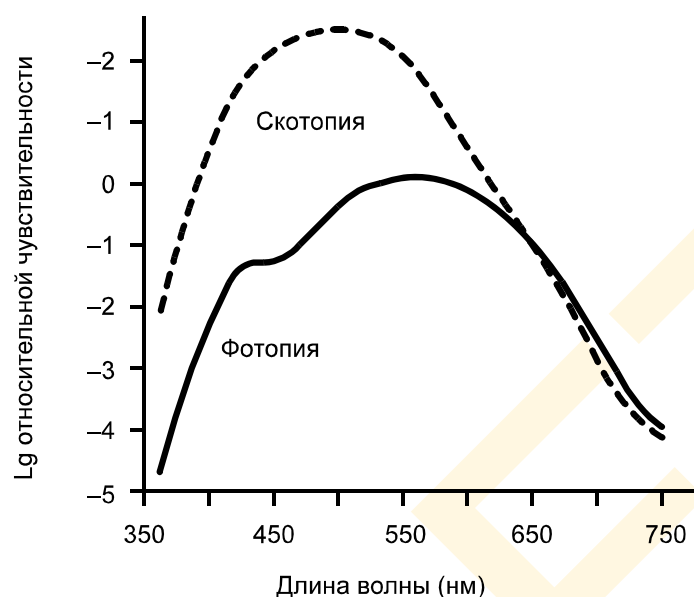
**Рис. 1.3.18** А — аппроксимированные спектры поглощения цианолэйба (S-колбочки), хлоролэйба (M-колбочки) и эритролэйба (L-колбочки). В — аппроксимированная фотопическая спектральная чувствительность зрительной системы человека как функция от длины волны. (Smith & Pokorny, 1975).



**Рис. 1.3.19** А — гистологический срез центральной ямки сетчатки глаза человека; хорошо видно, что палочки отсутствуют. В — то же, но в лицевом срезе на уровне стрелок в А; все клетки — колбочки. С — гистологический срез периферийного отдела сетчатки. D — то же, но в лицевом срезе на уровне стрелок в С; мелкие множественные клетки — палочки; крупные клетки — колбочки; отметим, что на периферии сетчатки колбочки крупнее, нежели в центральной ямке (микрофотографии д-ра Christine Curcio).



**Рис. 1.3.20** Мозаика S-колбочек в сетчатке макаки (срез на уровне внутренних колбочковых сегментов). Клетки покрашены специальным желтым красителем (светлые пятнышки). Темные пятна — внутренние сегменты прочих колбочек. Мелкие пятнышки — палочки.



**Рис. 1.3.21** Функции фотопической и скотопической спектральных чувствительностей. Данные функции представляют абсолютные чувствительности двух систем — фотопической и скотопической. Отметим, что по длинным волнам видимого спектра фотохроматический интервал практически равен нулю (Wald, 1945).

Рассмотрим монохроматический стимул с длиной волны 500 нм, интенсивность которого приподнялась над скотопическим порогом: этот стимул будет воспринят скотопической подсистемой и будет выглядеть как бесцветный (ахроматический). Если интенсивность стимула повышать дальше, то в конечном счете он начнет выглядеть окрашенным, то есть будет зафиксирован фотопическим компонентом зрения. Разница в чувствительностях скотопической и фотопической систем по данной длине волны именуется *фотохроматическим интервалом*.

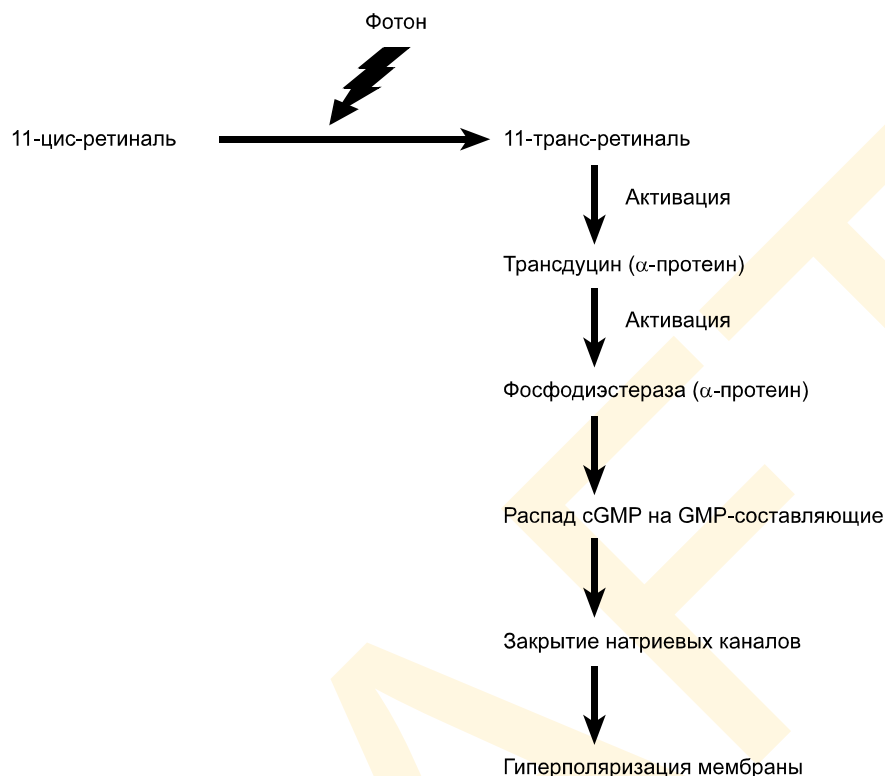
Отметим, что скотопическая подсистема много чувствительнее фотопической по всем длинам волн, за исключением длинноволнового участка спектра. В этой области фотохроматический интервал примерно равен нулю, то есть палочки и колбочки примерно равновеликочувствительны (после 650 нм колбочки даже чуть более чувствительны).

**Сдвиг Пуркине.** При переходе от скотопических условий наблюдения к фотопическим пик чувствительности зрительной системы сдвигается с 507 нм на 555 нм. Сие есть т.н. *сдвиг Пуркине* — относительный рост субъективной яркости длинноволновых стимулов при переходе от скотопических яркостей к фотопическим и наоборот: при обратном переходе растет субъективная яркость стимулов коротковолновых.

**Биохимия фоторецепторного отклика.** Теперь, на примере родопсина, рассмотрим то, как светочувствительный пигмент, конвертирует световую энергию в энергию электрохимическую (рис. 1.3.22), т.е. реализует начальный этап зрительного восприятия.<sup>1</sup>

<sup>1</sup> «Сакральная» точка взаимодействия неживой материи с живой, переход от неживого к живому. — *Прим. пер.*





**Рис. 1.3.22** Каскадная биохимическая реакция, инициируемая фотоном и приводящая к гиперполяризации мембраны фоторецептора.

И палочки, и колбочки в сравнении с обычными нейронами слегка деполяризованы: в отличие от мембранного потенциала покоя в  $-70$  мВ, потенциал фоторецепторов составляет  $-50$  мВ. После захвата фотона фоторецептор гиперполяризуется — его потенциал с  $-50$  мВ меняется на  $-70$  мВ (Tomita, 1970). Обычному физиологу сие удивительно, поскольку стимуляция нейронов обычно ведет к деполяризации их мембран, нежели к гиперполяризации.

Уровень фоторецепторной гиперполяризации зависит от интенсивности стимула: чем мощнее стимул, тем выше гиперполяризация (кстати, это одна из причин того, почему потенциал фоторецептора принято называть *градуальным*).

Схема этапов, ведущих палочку (колбочку) к гиперполяризации, показана на рис. 1.3.22 (Lamb, 1986; Pugh & Cobbs, 1986; Stryer, 1986): в темноте ионы натрия ( $\text{Na}^+$ ) непрерывно текут внутрь внешнего сегмента палочки через мельчайшие поры — т.н. *натриевые каналы*; сие движение именуется *темновым током*, создающим легкую гиперполяризацию клеточной мембраны; поглощение фотона родопсином инициирует серию биохимических реакций, приводящих, в итоге, к блокировке натриевых каналов и результирующей гиперполяризации внешнего сегмента.

Молекула родопсина состоит из двух частей: *опсина* и *хромофора*. Опсин — это инертное звено аминокислотной цепи, заключенное между мембранами дисков внешнего сегмента палочки. Опсины определяют характер поглощения фотопигментом, т.е. кривую его спектрального поглощения. Хромофор — это вещество, единое у всех фоторецепторов и собственно отвечающее на световое воздействие. Содержит *ретиаль* — альдегидное производное ретинола (витамина А).

Когда ретиналь не отбелен, он находится в состоянии 11-*цис* изомера (на 11 атоме углерода молекула ретиналя имеет перегиб).

Поглощение фотона переводит молекулу ретиналя в полный *транс*-изомер (*транс*-ретиналь).

Транс-ретиналь активирует белок *трансдуцин*.

Трансдуцин активирует фермент *протеин-фосфодиэстеразу (PDE)*.

Протеин-фосфодиэстераза разбивает сложное органическое соединение — *циклический гуанозин-монофосфат (сGMP)*, содержащееся в цитоплазме фоторецептора, на ациклические компоненты.

Снижение концентрации циклического гуанозин-монофосфата в клетке ведет к закрытию натриевых каналов во внешнем сегменте палочки, следствием чего является гиперполяризация мембраны.

Отметим, что число натриевых каналов во внешнем сегменте палочки ограничено, что сдерживает рост потенциала при гиперполяризации (Baylor et al, 1984). Когда всего лишь 10% палочкового родопсина отбелено — закрывается критичное число натриевых каналов и дальнейшее отбеливание родопсина уже не ведет к дальнейшей гиперполяризации. Сие позволяет объяснить психофизический феномен палочкового насыщения: когда 10% родопсина отбелено — палочки более уже не могут гиперполяризоваться — они *насыщены*.

Коль скоро в результате отбеливания молекула родопсина утрачивает оптическую активность (т.е. становится прозрачной), то она уже не в состоянии поглотить другой фотон. Однако со временем эта молекула возвращается в свое исходное (неотбеленное) состояние благодаря действию особых ферментов, содержащихся в т.н. *пигментном эпителии* — тонкой окрашенной выстилке глаза, расположенной позади сетчатки.

Наконец отметим, что биохимический механизм колбочкового отклика можно полагать аналогичным механизму палочкового.

Далее. Разность электрических потенциалов, возникшая под действием множества фотонов, поглощенных данным рецептором, представляет собой *отклик* этого рецептора. Разность эта может *плавно* меняться по величине (т.е. не дискретно, градуально), что отличает фоторецептор от клеток прочих отделов нервной системы.

В количественном выражении отклик фоторецептора есть *логарифмическая функция* от числа поглощенных фотонов: сие значит, что одно и то же приращение сигнала на выходе потребует очень небольшой добавки фотонов при слабом световом воздействии много больше — при сильном (см. рис. 1.3.23).

Было установлено при этом, что полная регенерация родопсина наступает через 30 мин., а колбочковых пигментов примерно через 6 мин.

Отличие во времени восстановления палочкового и колбочковых фотопигментов приводит к неравномерности хода темновой адаптации, о которой мы упоминали в разделе 1.1.3. Точную схему наступления темновой адаптации можно измерить объективно, путем определения т.н. *абсолютного порога* — минимальной интенсивности света, который еще может быть воспринят наблюдателем через разные промежутки времени (после того, как наблюдатель этот был помещен в темную комнату).

#### 1.3.2.4 Темновая адаптация

После экспозиции адаптирующим светом палочки и колбочки восстанавливают свою чувствительность по-разному. Динамика этого восстановления была изучена в экспериментах по темновой адаптации и их результаты мы обсудим ниже.

На рис. 1.3.24 показана типичная кривая темновой адаптации: чтобы получить та-

кую кривую зрительная система испытуемого была подвергнута воздействию яркого адаптирующего света, отбеливающего большую часть фотопигмента; затем адаптирующий свет выключался и выявлялся порог восприятия как функция от времени предъявления тестового стимула малой интенсивности на черном фоне. Тестовый стимул при этом был достаточно крупным, центрально фиксированным, монохроматическим с длиной волны 420 нм. Важно также то, что на протяжении эксперимента помещение оставалось полностью затемненным, то есть наблюдатель не подвергался воздействию внешнего света.

У кривой темновой адаптации есть несколько важных особенностей, и, возможно наиболее очевидна та, что в течение 35 минут порог чувствительности снижается почти на 5 логарифмических единиц, то есть, спустя 35 мин. чувствительность зрения возрастает в 100000 раз.

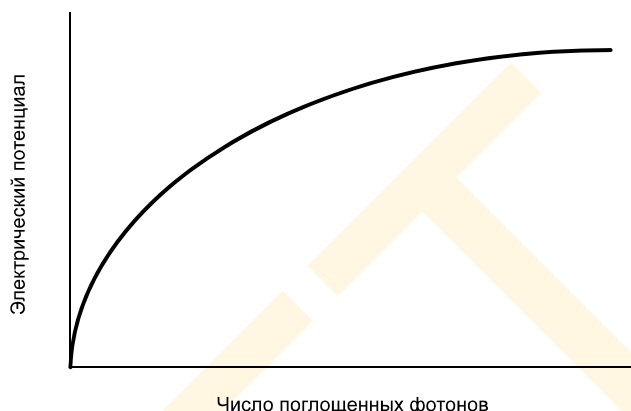
Второй и очень серьезной особенностью кривой является ее разделенность на две секции.

*Первая секция*, демонстрирует лавинообразное снижение порога в течение первых 10 мин., за которым кривая выходит на т.н. *колбочковое плато*. Физиологический смысл первой секции состоит в динамике фотопических порогов. Примерно на 12-й минуте теста происходит резкое изменение наклона кривой, именуемое *палочко-колбочковым изломом*, за которым начинается плавное снижение порога в течение примерно 35 мин с исходом в т.н. *палочковое плато* (Hetch et al, 1937, 1942).

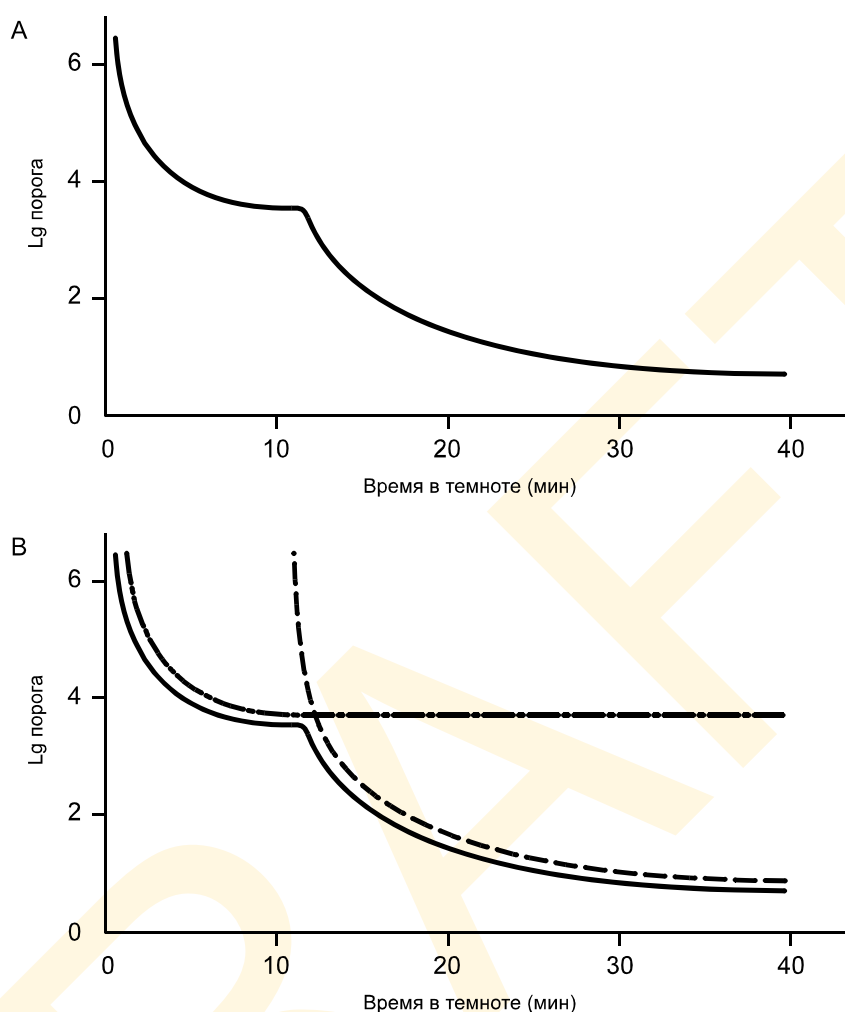
*Вторая секция* кривой, как видим, представляет скотопические пороги. Палочко-колбочковый излом возникает в тот момент, когда палочки становятся столь же чувствительны, что и колбочки. До этого момента лишь колбочки реагируют на стимул; после — только палочки.

Далее. Из кривой темновой адаптации невозможно напрямую получить фотохроматический интервал, но только опосредованно: он будет представлять собой разность между колбочковым плато (представляющим минимальный фотопический порог по данному стимулу), и палочкового плато (представляющего минимальный скотопический порог по тому же стимулу). По 10-градусному, центрально фиксированному монохроматическому стимулу с длиной волны 420 нм, использованному при построении кривой темновой адаптации рис. 1.3.24, фотохроматический интервал равен примерно 3 логарифмическим единицам.

**Влияние длины волны стимула на форму кривой темновой адаптации.** Кривая темновой адаптации рис. 1.3.24 получена с использованием монохромата с длиной волны 420 нм; однако каковой будет форма кривой при стимуле в 650 нм? На



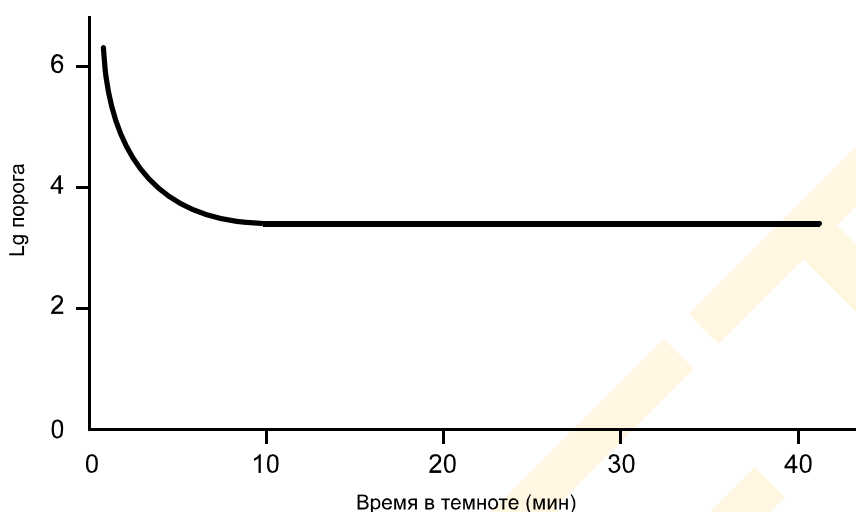
**Рис. 1.3.23** Зависимость между количеством фотонов, поглощенных фоторецептором и его выходным сигналом логарифмическая: когда поглощено небольшое количество фотонов — это приводит к выраженному изменению выходного сигнала рецептора; когда поглощено большое число фотонов — изменения выходного сигнала невелики.



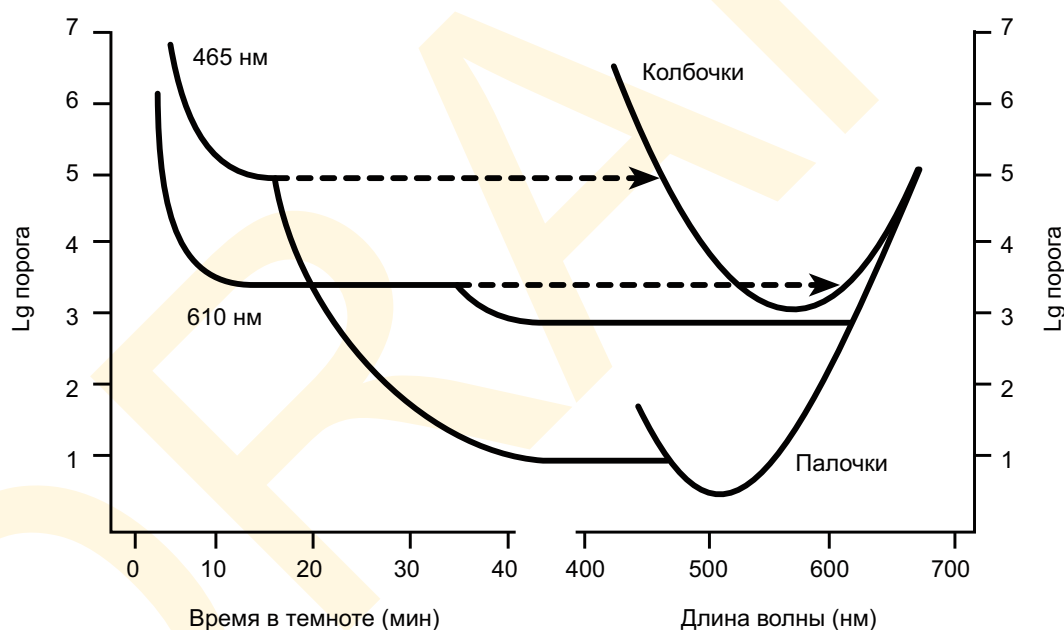
**Рис. 1.3.24** А — типичная кривая темновой адаптации. В — кривая темновой адаптации по стимулу, на который реагируют только палочки (точечный пунктир) и стимулу, на который реагируют только колбочки (точка-тире пунктир). Стимул, на который реагируют и палочки, и колбочки — сплошная линия.

рис. 1.3.25 показана такая кривая и мы видим, что она всерьез отлична от предыдущей. Если взглянуть на кривую рис. 1.3.21, то становится понятным, что результат был вполне прогнозируем, поскольку фотохроматический интервал по 650 нм равен нулю. Поскольку в этой части спектра скотопическая система не чувствительнее фотопической, палочко-колбочкового излома не возникает (Hetch et al, 1937): напротив, монохроматический стимул с длиной волны в 420 нм дает высокий фотохроматический интервал, следовательно кривая темновой адаптации, построенная с его помощью, будет иметь выраженный палочко-колбочковый излом (см. рис. 1.3.24).

Далее влияние длины волны стимула на форму кривой темновой адаптации рассмотрим по рис. 1.3.26: слева даны кривые темновой адаптации по монохроматам с длиной волны 465 нм и 610 нм; справа скотопическая и фотопическая пороговые функции, представляющие собой обратные функции соответственно скотопической и фотопической спектральных чувствительностей зрительной системы (см. рис. 1.3.21).



**Рис. 1.3.25** Кривая темновой адаптации, построенная по монохроматическому стимулу с длиной волны в 650 нм. Отметим, что палочко-колбочковый излом отсутствует.



**Рис. 1.3.26** Кривые темновой адаптации по стимулам 465 нм и 610 нм и их соотношение с функциями фотопического и скотопического порогов.

Возьмем стимул с длиной волны 465 нм: спустя 15 мин после наступления темноты появляется колбочковое плато, представляющее минимальный порог (максимальную чувствительность) фотопической системы. Пунктирная линия, продолжая это плато и пересекающая кривую функции фотопического порога по стимулу в 465 нм, демонстрирует то, что колбочковое плато соответствует некоей специфической точке фотопической пороговой функции. Палочковая порция кривой темновой адаптации по стимулу в 465 нм выходит в ноль примерно на 40 минуте и продолжается пунктиром до кривой скотопической пороговой функции.



Предлагаемый тип анализа очень удобен, поскольку иллюстрирует фотохроматический интервал в логике как кривых адаптации, так и пороговых кривых. У кривых адаптации фотохроматический интервал определяется колбочковым и палочковым плато; у пороговых кривых он представляет собой разность между скотопическим и фотопическим порогами.

Теперь озадачимся вопросом: почему палочко-колбочковый излом по монохромату с длиной волны в 465 нм более выражен и наступает раньше, чем по стимулу с длиной волны в 610 нм? Часть ответа относится к поглощательным свойствам родопсина: как показано на рис. 1.3.16С вероятность поглощения молекулой родопсина фотона с длиной волны в 465 нм относительно высока. Следовательно, на ранних этапах темновой адаптации, после того как относительно небольшое количество родопсина восстановится, чувствительность скотопической системы превысит таковую у системы фотопической. В случае стимула с длиной волны в 610 нм ситуация иная: вероятность, что молекула родопсина поглотит фотон с длиной волны в 610 нм существенно меньше, и только после того, как относительно большое количество молекул родопсина восстановится (т.е. на поздних стадиях темновой адаптации) скотопическая система обретет большую чувствительность, нежели система фотопическая.

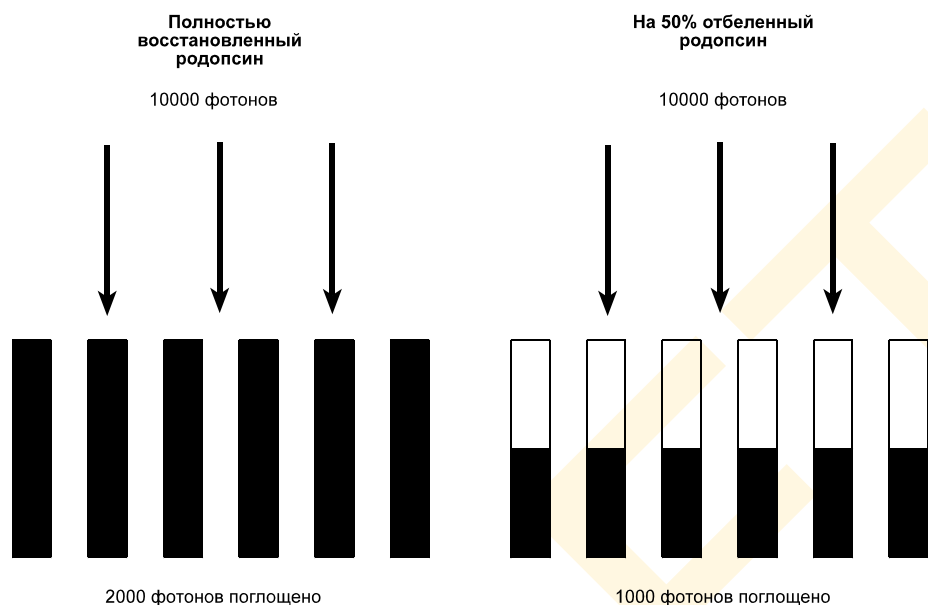
**Размеры и местоположение стимула.** Предыдущие примеры строились на том, что предъявляемый стимул центрально фиксирован, а его пространственный размер таков, что целиком кроет центральную ямку сетчатки и частично захватывает периферийное ее окружение. К примеру, стимул диаметром в  $10^\circ$  дает те результаты, что обсуждались выше.

Взяв мелкий стимул, скажем  $0.5^\circ$  и предъявив его строго центральной ямке, получим вполне предсказуемый эффект — только колбочковую функцию (рис. 1.3.25). При этом отсутствие палочковой функции лишний раз докажет нам, что в центральной ямке нет палочек.

**Физиологический базис темновой адаптации.** Чем можно объяснить медленное восстановление чувствительности после экспозиции ярким адаптирующим светом? В процессе адаптации к темноте восстановление фотопигмента повышает вероятность захвата фотонов, то есть растет чувствительность. Регенерацией фотопигмента можно целиком истолковать восстановление чувствительности зрения, что будет являться т.н. *фотомеханическим объяснением* данного феномена (Hetch, 1937). Фотомеханический принцип проиллюстрирован рис. 1.3.27, на котором схематично изображены палочки, у которых отбелена только половина родопсина. В сравнении с неотбеленной стадией имеет место лишь половинная вероятность того, что фотон будет захвачен родопсином. Следовательно, основываясь только лишь на «фотопигментных» рассуждениях 50% отбелики прогнозируют удваивание порога. Но сие рассуждение неверное — на самом деле порог растет в  $10^{10}$ !

Полые кружки рис. 1.3.28 демонстрируют динамику восстановления родопсина у т.н. *палочковых монохроматов* — больных, в сетчатке которых наличествуют одни только палочки.<sup>1</sup> Сплошная линия — это кривая темновой адаптации таких лиц;

<sup>1</sup> Изменения в состоянии пигмента могут быть оценены *in vivo* (прижизненно) путем т.н. *ретинальной денситометрии* (Rushton & Campbell, 1954): световой пучок известной интенсивности фокусируется на сетчатке; часть света поглощается рецепторами и пигментным эпителием, часть отражается назад. Интенсивность отраженного света легко измерить, и, выполнив такое измерение через разные промежутки времени после предъявления глазу очень яркого све-



**Рис. 1.3.27** Половинное отбеливание родопсина понижает вероятность поглощения фотона в два раза. Однако порог при этом возрастает на десять порядков, т.е. в  $10^{10}$  раз.

пунктирная — кривая темновой адаптации здоровых людей. Мы видим, что 50% -е отбеливание родопсина (правая ось ординат) ведет к росту порога на 10 логарифмических единиц, но никак не к удваиванию, как то следует из фотомеханического принципа.<sup>1</sup> Итак, не только и не столько регенерация фотопигмента вносит основной вклад в сложную динамику темновой адаптации, но также еще ряд факторов, причем не только рецепторных, но и *пострецепторных*, действующих в высших отделах сетчатки и постретинальных (послесетчаточных) отделах зрительной системы (Schnapf & Baylor, 1987).

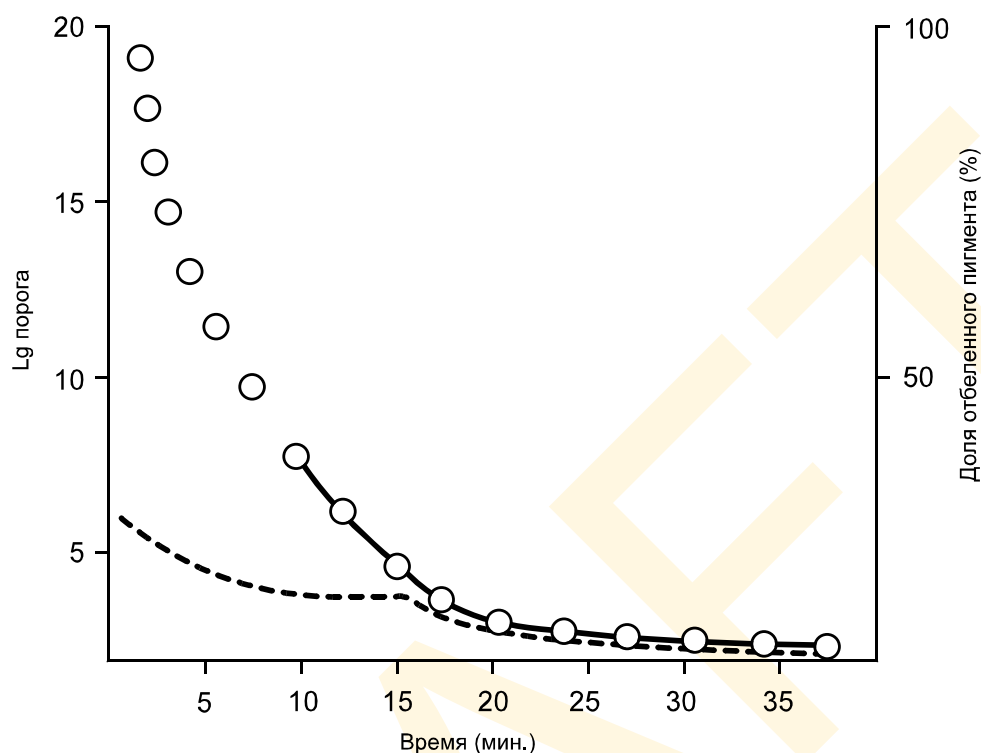
### 1.3.2.5 Световая адаптация

Когда мы выходим из дома на залитую солнцем улицу, освещенность нашей сетчатки увеличивается в несколько тысяч раз. Вопреки такому гигантскому изменению в освещенности результат восприятия объектов внешнего мира практически не меняется, к примеру: клетчатая рубашка выглядит на солнце почти так же как и дома. То есть, в течение очень короткого периода (столь короткого, что мы не успеваем осознать его) наша зрительная система адаптируется к резко возросшему уровню освещенности сцен. Сей процесс именуется *световой адаптацией*.

Световую адаптацию можно изучать с помощью методики *порогового приращения* или как ее еще называют методики определения *едва заметного отличия* (рис. 1.3.29): на некоем фоне (определенной интенсивности свечения) наблюдателю предъявляется стимул, мощность которого повышают до тех пор, пока стимул этот не станет визуаль-

тового стимула, отбеливающего большинство молекул фотопигмента, можно рассчитать скорость его регенерации.

<sup>1</sup> Отметим, что в момент палочко-колбочкового излома кривой адаптации 90% родопсина оказывается восстановленным.



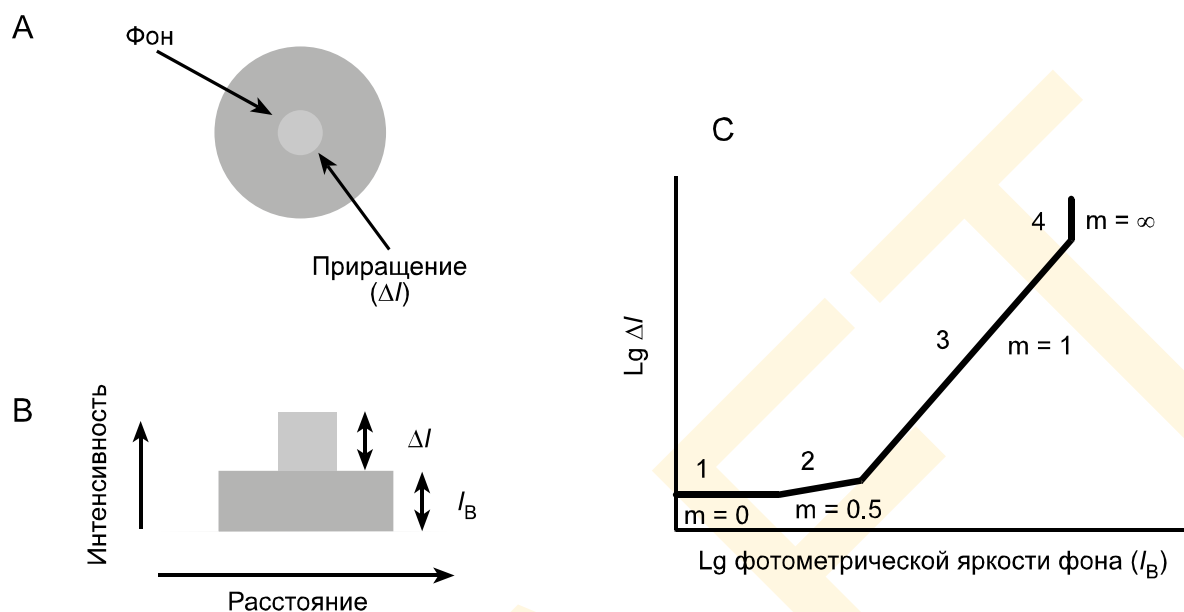
**Рис. 1.3.28** Кружками отмечены результаты ретиальной денситометрии палочковых монохроматов; сплошной линией отмечена динамика порога восприятия по мере темновой адаптации этих больных. То же отмечено пунктиром у здоровых людей (Rushton, 1965a).

но отличим от фона. После того, как порог определен, интенсивность фона увеличивают и повторяют измерение. Опыт проводят в диапазоне фотометрических яркостей фона от максимально возможного черного до экстремально яркого, что, в итоге, приводит к построению графика световой адаптации, на котором пороговые приращения ( $\Delta I$ ) даны как функция от адаптирующей интенсивности фона ( $I_B$ ).

**Скотопическая световая адаптация.** На рис. 1.3.29 показаны четыре участка функции скотопической световой адаптации. Объяснения по первым двум участкам сугубо технические и мы опишем их кратко. Участок 1, наклон которого ( $m$ ) равен нулю — это диапазон фотометрических яркостей фона, при которых детекция тестового стимула ограничена неизбежным *нейронным шумом* зрительной системы: фон совершенно черный, но нейронный шум продуцирует то, что принято называть «темновым светом»<sup>1</sup> (Fechner, 1860; Barlow, 1956).

Второй участок кривой, имеющий наклон порядка 0.5, отражает фотонные флуктуации фона (DeVries, 1943; Rose, 1948): фон достаточно тусклый и флуктуации, присутствующие источнику света играют ключевую роль в определении порога. Сей факт часто выражают в виде корневидного закона *Де-Вриза — Розе*, прогнозирующего то, что при низких яркостях фона  $\Delta I = \sqrt{I_B}$ .

<sup>1</sup> Чтобы убедиться в существовании «темнового света» достаточно закрыть глаза и приглядеться к своим зрительным ощущениям — полной темноты не возникнет, но перед «взором» будут появляться разнообразные зрительные эффекты.



**Рис. 1.3.29** Чтобы построить кривую световой адаптации стимул организуют как фон с участком приращения фотометрической яркости. А — стимул. В — профиль стимула по интенсивности. С — упрощенная кривая световой адаптации (только скотопический фрагмент).

Третий участок кривой, покрывающий диапазон в 4 логарифмических единицы, имеет наклон примерно 1, что соответствует закону Вебера (Aguilar & Stiles, 1954; Barlow, 1965; Walraven & Valetton, 1984). По мере роста фотометрической яркости фона пороговое приращение должно расти так, что отношение порогового приращения ( $\Delta I$ ) к фотометрической яркости фона ( $I_B$ ) остается постоянным. Участок кривой с наклоном в 1 принято называть *веберовской фракцией*, а величину  $\Delta I/I_B$  — *веберовской постоянной*.

Веберовская постоянная по скотопическому зрению примерно равна 0.14 (Cornsweet, 1970). Если яркость фона равна неким 100 единицам, то пороговое приращение будет равно 14 единицам ( $\Delta I = 14$ ). Если яркость фона выросла до 1000 единиц, то по Веберу приращение должно вырасти до 140 единиц ( $\Delta I = 140$ ). Хотя *относительная* чувствительность зрительной системы (0.14) действительно не меняется по мере роста яркости фона, имеет место снижение ее абсолютной чувствительности (порог вырастает с 14 до 140 единиц). Сие несоответствие между относительной и абсолютной чувствительностями именуется *эффектом Вебера*.<sup>1</sup>

Отношение приращения интенсивности к яркости фона именуется *контрастом*. Говоря, что реакция зрительной системы отвечает закону Вебера, мы говорим, что *пороговый контраст* остается постоянным при смене уровня освещения.

Заключительная, четвертая секция скотопического фрагмента кривой световой

<sup>1</sup> Физиологическая подоплека веберовского эффекта (в контексте фоторецепции) только-только начинает проясняться. Вероятно кальциевый метаболизм играет важную роль в этом явлении (Fain & Matthews, 1990).

адаптации имеет бесконечный наклон, т.е. стоит вертикально. Сие указывает на то, что палочки достигли насыщения (Aguilar & Stiles, 1954). При такой фотометрической яркости фона палочки уже не выдают сигнала на приращение энергии стимула (каким бы оно ни было), поскольку они перевозбуждены чрезмерно ярким для них фоном.

Мы можем предположить, что *палочковое насыщение* возникает тогда, когда весь родопсин отбелен. Но исследования показали, что при полном палочковом насыщении отбеливается лишь 10% родопсина (Rushton, 1965a).

Отбеливание фотопигмента приводит к тому, что натриевые каналы во внешнем сегменте палочки закрываются и что в свою очередь ограничивает ток йонов натрия внутрь клетки, приводя к гиперполяризации ее мембраны. Степень гиперполяризации в известной мере зависит от интенсивности стимула. Поскольку число натриевых каналов конечно, величина палочковой гиперполяризации тоже конечна (Baylor et al, 1984). Когда примерно 10% молекул родопсина отбелены, все натриевые каналы оказываются закрытыми и дальнейшее отбеливание родопсина не ведет к увеличению степени гиперполяризации и, следовательно, к отклику палочек.

**Фотопическая световая адаптация.** Закон Вебера продолжает работать и в фотопических условиях. Константа составляет порядка 0.015, указывая на то, что фотопическая система более чувствительна к контрасту, нежели скотопическая (скотопическая константа, напомним, равна 0.14) (Stiles, 1953). Отметим при этом, что хотя фотопическое зрение более чувствительно к контрасту, его абсолютная чувствительность ниже.

### 1.3.2.6 Кодирование оптической информации

Прочие особенности ретинальной (сетчаточной) физиологии касаются кодирования информации об интенсивности света. Поскольку задачей глаза является конверсия световых модуляций в нейронные отклики, естественно предположить, что рецепторы не будут отвечать на темноту, но будут повышать свой синаптический выход по мере роста световой интенсивности. Однако в рецепторах позвоночных все происходит диаметрально противоположным образом (Toyoda, Nosaki & Tomita, 1969) — отклик на вспышку света снижает синаптическую активность на выходе рецептора. Несмотря на то, что такое поведение рецептора более чем неожиданно, оно, тем не менее, вполне возможно, поскольку значимым является лишь то, что нейронная активность сохраняет информацию, закодированную в *разностях* световых интенсивностей по всему полю зрения, и задача эта может быть успешно решена либо положительной, либо отрицательной корреляцией между нейронной активностью и световой интенсивностью.

У беспозвоночных все наоборот (что как раз согласуется с интуитивной логикой) — их фоторецепторы усиливают выходной сигнал по мере роста световой интенсивности. По-видимому, аномальное снижение активности рецепторов позвоночных выправляется на следующем синапсе, в котором рост световой интенсивности увеличивает выход нейротрансмиттера.

### 1.3.2.7 Проводящие пути

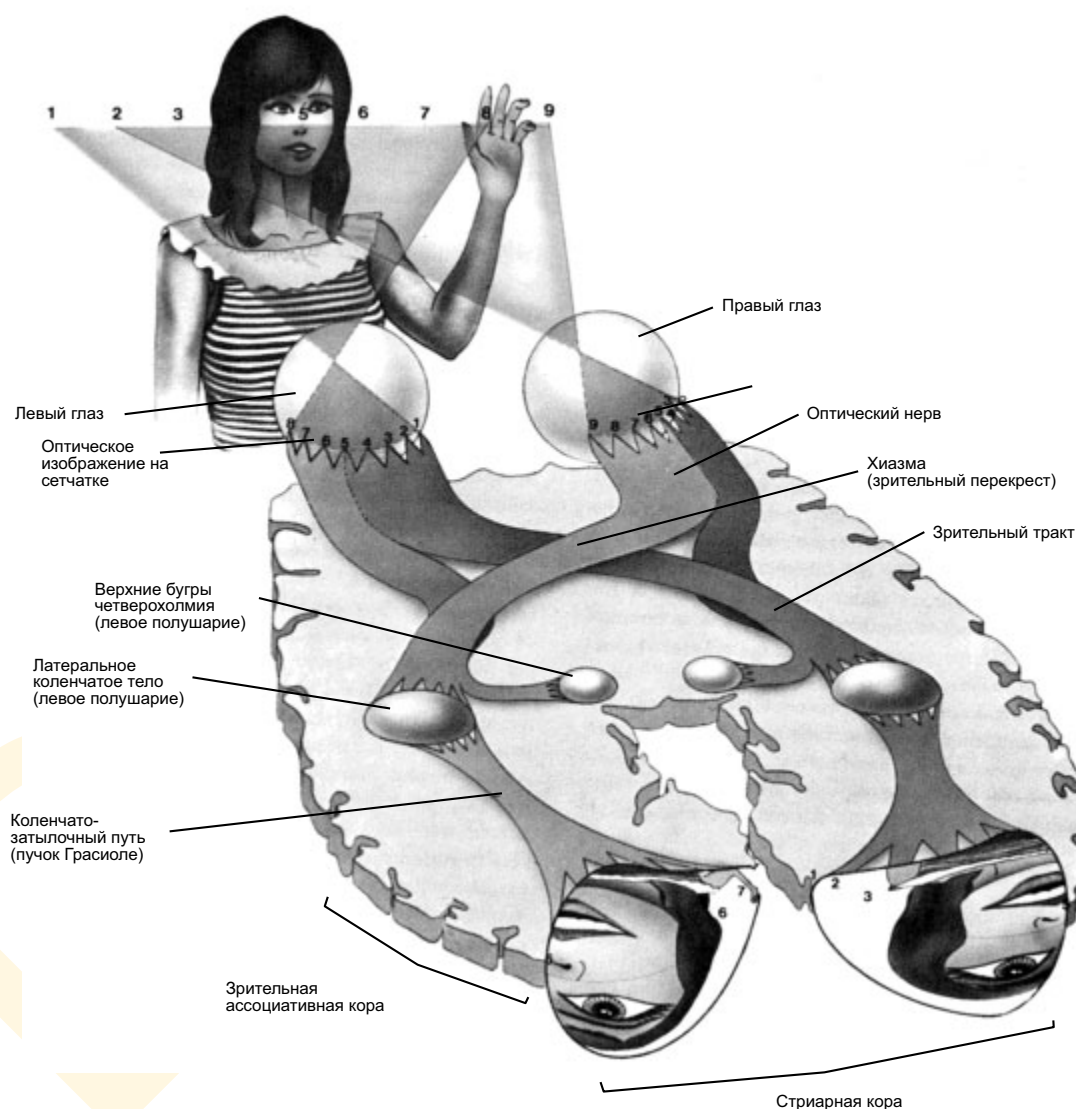
Как мы уже сказали выше, аксоны ганглиозных клеток покидают глаз в виде *оптического нерва*, образующего на своем пути *зрительный перекрест* или *хиазму*,<sup>1</sup> в кото-

<sup>1</sup> Термин «хиазма» возник из аналогии с греческой литерой «кси» — «χ».



рой волокна т.н. *назальной части* (ближней к носу) центральной ямки сетчатки каждого глаза перекидываются на противоположную сторону головного мозга, тогда как остальные продолжают ход по своей стороне (см. рис. 1.3.30). Результатом перекреста является то, что проекции внешних полей зрения в кору попадают в противоположные полушария: вся информация из левой части поля зрения идет в правое полушарие головного мозга, а вся информация из правой части поля зрения — в левое.

От зрительного перекреста в мозг отходят два отдельных проводящих пути — каждый в свое полушарие (см. ниже), и каждое дает мелкую ветвь (всего лишь несколько процентов волокон) в т.н. *верхние бугры четверохолмия* — некоего скопления ядер в



**Рис. 1.3.30** Нервные пути от глаза к зрительной коре. Информация от внутренней (назальной) порции сетчатки перекидывается на противоположную сторону мозга, в результате чего каждое полушарие получает сигналы только с противоположной стороны поля зрения. Отметим, что корковое представление центральной области поля зрения диспропорционально (Frisby, 1979).

мозговом стволе: сей зрительный центр обрабатывает первичную информацию о местоположении предметов в сцене и управляет глазными движениями.

Основные же проводящие пути вначале впадают в *латеральное коленчатое тело (ЛКТ)* т.н. *таламуса* — некоего подкоркового образования мозга, — а затем в *затылочную (первичную зрительную) кору*. В 4-й главе мы будем подробно говорить о том, как происходит расширенная обработка информации в зрительной коре, а здесь опишем лишь макроанатомическую и макрофизиологическую организации этой сложной структуры.

### 1.3.3 Зрительная кора

При внешнем осмотре коры головного мозга человека хорошо видно, что вся ее поверхность покрыта бороздами и извилинами. Борозды и извилины есть следствие того, что кора головного мозга представляет собой крупную многослойную нейронную структуру, которую необходимо уместить в маленькое и жесткое пространство черепной коробки.

Второй макроскопической особенностью коры является то, что она разделена на две половины — два более или менее симметричных полушария. В остальном кора выглядит вполне гомогенной и глядя невооруженным глазом совершенно невозможно понять, как она работает. Действительно, функция коры столь сложна, что ученые древности даже не считали мозг органом, выполняющим ментальную функцию: Аристотель, к примеру, полагал, что мыслит сердце, а мозг выступает в роли радиатора, охлаждающего кровь.

#### 1.3.3.1 Локализация функций

Как только значимость головного мозга в жизнедеятельности организма была осознана исследователями, так сразу встал вопрос о том, можно ли говорить о пространственной локализации его функций, то есть: действительно ли разные умственные способности лежат в разных анатомических отделах головного мозга, или все его функции равномерно распределены в мозговой толще?

Первое подтверждение гипотезы локализации пришло из *френологии* — паранауки изучающей связь между умственной деятельностью и строением черепа. Френологи утверждали, что размер бугров и впадин на черепе указывает на размер и развитие (или недоразвитие) подлежащих структур мозга. Задача френологов состояла в том, чтобы согласовать краниометрические данные (т.е. данные о форме и размерах отделов черепа) с такими ментальными атрибутами как «характер», «мышление» и «духовность». Итогом стали коллекции причудливых френологических карт, одна из которых показана на рис. 1.3.31. Однако же ожидавшиеся френологами корреляции не подтвердились и реальной научной поддержки концепции локализации не было до тех пор, пока в конце XIX века не началось систематическое изучение эффектов повреждения тех или иных участков головного мозга.

Научную основу идея локализации функций получила тогда, когда медики начали прибегать к посмертному анализу головного мозга больных с умственными расстройствами от разного рода травм головы. Было показано, что характер нарушений четко коррелирует с повреждением определенных участков мозга. В том числе было выявлено, что визуальные дисфункции чаще всего возникают тогда, когда повреждаются задние отделы мозга, особенно его затылочные доли (см. рис. 1.3.2). Особо ценными стали исследования молодых японских врачей, изучавших зрительные нарушения у солдат

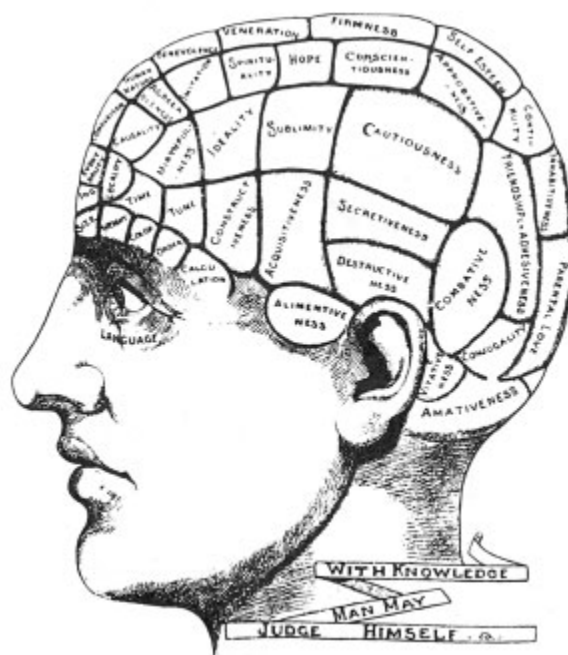
с огнестрельными ранениями в голову в период русско-японской войны 1904 — 1905 гг. (Glickstein, 1988). В результате этих (и ряда других) нейропсихологических исследований было точно установлено, что затылочные доли являются первичным корковым отделом мозга, принимающим зрительную информацию и что существуют иные корковые отделы, как говорят, *специализированные по той или иной сенсорной модальности*: слуху, вкусу, осязанию и обонянию.

Однако же имеются и данные, противоречащие концепции локализации, в частности, полученные в широко известной серии исследований американского психофизиолога Карла Лэшли и его коллег, пытавшихся локализовать память (зрительную и пр.) в экспериментах на крысах (Lashley, 1929, 1931, 1950): ученые удаляли из разных участков различные по размеру фрагменты коры головного мозга у крыс, но так и не нашли никакого специфического локуса, отвечавшего за те или иные виды памяти. Единственное, что давало корреляцию с отличиями в поведении — это общее количество удаленной коры: чем больше мозговой ткани было удалено, тем больше животное допускало поведенческих ошибок. Результат Лэшли окрестил «законом действующих масс». Открытие поставило под серьезное сомнение идею локализации памяти в каком бы то ни было регионе коры, и ученый предположил, что память более или менее равномерно распределена по всей коре головного мозга.

Итак, мы до сих пор не знаем пределов локализации той или иной функции мозга, но доказательства *усиленной поддержки* зрительной функции определенными отделами коры у нас на руках: в зрительной коре обезьян, к примеру, было выявлено более 30 зрительных полей, и каждый год обнаруживаются новые (Van Essen & DeYoe, 1995). Сегодня лучше всего изучены зрительные центры, расположенные в *окипитальных (затылочных)*, *париетальных (теменных)* и *темпоральных (височных)* долях коры (см. рис. 1.3.2).

### 1.3.3.2 Затылочная кора

Наиболее полные анатомические и физиологические данные о зрительной коре поступают к нам из исследований физиологии обезьян, в частности, макака. Анализ поведения этих животных показал, что их зрение поразительно схоже с человеческим и тем создает добротную зоомодель зрительной системы человека (De Valois & De Valois,



**Рис. 1.3.31** Френологическая карта. Френологи были уверены, что умеют определять локализации мозговых функций, путем краниометрии (измерений разных характеристик черепа) и находя корреляции ее данных с психологическими особенностями людей (Kolb & Whishaw, 1996).

1988). Фундаментальные *цитологические* (т.е. клеточные) исследования зрительной коры начались более двадцати лет назад американским нейрофизиологом Дэвидом Хьюбелом и шведским нейробиологом Торстеном Визелем (1959, 1962), благодаря чему сегодня уже довольно много известно об анатомии и физиологии зрения млекопитающих. Были исследованы структура и функция многих зрительных областей, а взаимосвязи между ними определены с помощью различных хитроумных методик. И хотя будет весьма нескромно утверждать, что мы понимаем как работает зрительная кора, у нас все же появились некоторые проблески в отношении того, какие отделы коры отвечают за зрительную функцию и каковы связи между этими отделами.

Итак, первый этап корковой обработки зрительной информации начинается в т.н. *стриарной (полосчатой) коре* — фрагменте затылочной доли, на своей стороне головного мозга, принимающей сигналы от латерального колленчатого тела. То есть, также как и у колленчатого тела, входной сигнал стриарной коры, как говорят, полностью кроссирован (перекрещен): левое поле зрения проектируется в стриарную кору правого полушария, правое — в стриарную кору левого (см. рис. 1.3.30).

Мэппинг отклика сетчатки в стриарную кору, как говорят, — *топографичен*, то есть: соседние области сетчатки проецируются на соседние области стриарной коры. Сие преобразование сохраняет *качественные* пространственные связи, но искажает *количественные*, как если бы растягивали фотографию на резине. На рис. 1.3.30 показаны мера и характер дисторсий, возникающих при топологических преобразованиях: центральная область поля зрения, которая приходится на центральную ямку или область возле нее, получает пропорционально большее представление в коре, нежели периферия. Сие именуется *кортикальным увеличением* и отражает тот факт, что пространственная информация об объектах окружающего мира наиболее детализирована в центральном регионе сетчатки, нежели на ее периферии.<sup>1</sup>

Однако же очевидно, что в восприятии объекты в центре поля зрения ничуть не крупнее периферийных. Мы можем очень легко в этом убедиться, если посмотрим на свои ладони, расположив их перед собой, но сосредоточив взгляд, скажем, на левой: если восприятие пространства искажалось бы корковым увеличением, то левая ладонь выглядела бы ощутимо крупнее правой. Но этого не происходит, и мы лишь лучше видим детали на левой руке (складки, вены, ногти и пр.).

Все сказанное выше о стриарной коре — это лишь малая толика того, что о ней сегодня известно. Но мы отложим разговор об этом отделе мозга до 4-й главы.

Стоит сказать также, что проектирование происходит и в другие отделы зрительной коры, в частности, в другие области затылочных долей, а также теменные и височные доли.

Разные участки коры отвечают за разные виды зрительной обработки: одной из важнейших особенностей, позволяющих назвать данный регион коры «зрительным» (как, скажем, стриарную кору), — это топографическая проекция ретинальных (сетчаточных) областей. Действительно, зрительная кора — это плотная мозаика мелких «карточек», кодирующих разные аспекты ретинальной стимуляции, включая субъективную яркость стимула, его хроматику, движение, глубину, текстуру и форму.

<sup>1</sup> Топологическое преобразование происходит постадийно с невысоким коэффициентом увеличения на каждой стадии: вначале некоторое увеличение центральных областей возникает на участке фоторецептор-ганглиозная клетка, затем на участке ганглиозная клетка-ЛКТ, а затем на участке ЛКТ-стриарная кора.



### 1.3.3.3 Теменная и височная кора

Один из фрагментов общей картины того, как работает зрительная кора принадлежит физиологам Мортимеру Мишкин (Mortimer Mishkin) и Лэсли Унгерлайдер (Leslie Ungerleider), которые привели убедительные доказательства того, что функции зрительных областей височной и теменной коры обезьян резко различны (рис. 1.3.32А). Т.н. *нижние височные центры* в нижних (вентральных)<sup>1</sup> отделах коры видимо выполняют функцию *идентификации объектов*, тогда как теменные центры в верхних (дорсальных)<sup>2</sup> отделах вероятно несут на себе функцию локализации. Эти проводящие пути часто называют системой «Что?» и системой «Где?» соответственно. Доказательство сего тезиса находим в экспериментах, в которых обезьянкам приходилось выполнять задания двух видов, причем после того, как одна или обе области коры были хирургически удалены (Ungerleider & Mishkin, 1982; Mishkin, Ungerleider & Macko, 1983).

Одной из задач являлся поиск предмета (см. рис. 1.3.32В): обезьяна должна была выбрать знакомый предмет из группы незнакомых, за что получала лакомство. Здоровое животное легко справляется с таким заданием и столь же легко с ним справляется обезьяна, у которой удалена часть теменной коры. Однако эта же задача предельно тяжела обезьянам, у которых отсутствует нижняя височная кора — то есть разрушена система «Что?».

Вторым заданием был т.н. выбор по метке (рис. 1.3.32С): вновь требовалось выбрать один предмет из двух, но на сей раз оба были одинаковой формы, отличаясь лишь своей пространственной приближенностью к специальной вешке. Опять же, здоровое животное легко управляется с такими заданиями, но обезьяна с удаленной теменной корой испытывает непреодолимые трудности. При этом животное с удаленной височной корой также вполне успешно справлялось с данным заданием экспериментаторов.

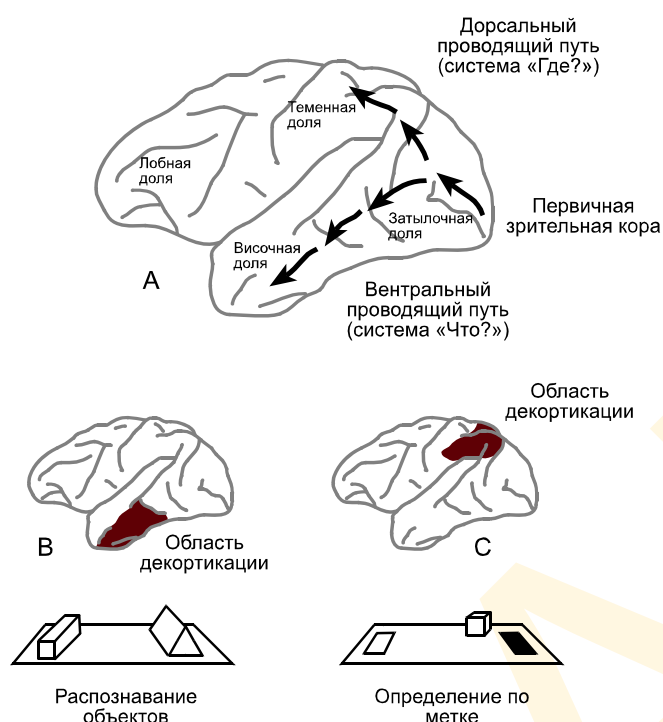
Результаты данного исследования (и ряда других) подтвердили версию Унгерлайдера и Мишкин, что в височной коре происходит обработка информации о форме объектов (что необходимо при их идентификации), и что в теменной коре обрабатывается информация о пространственной позиции предметов, необходимая при определении их местоположения. Позднее ряд исследователей показали, что функция теменной коры точнее описывается как управление пространственно-направленной моторикой (система «Как?»), то есть «дотянуться и схватить» (Goodale, 1995; Milner & Goodale, 1995). В любом случае, очевидно, что эти две совершенно разных информационных категории неизбежно должны объединяться где-то в недрах мозга, образуя связь «что-где», но по сей день неизвестно, где именно это происходит. Одним из кандидатов являются фронтальные (лобные) доли, поскольку они получают проекции как из теменной, так и из височной областей. Однако обработка информации может идти и по неким обводным путям, проходя при этом через ряд промежуточных корковых областей и упираясь в какое-то пока еще неизвестное место в коре.

Функциональное отличие между двумя описанными зрительными путями столь же значимо в физиологии зрительного восприятия людей, сколь и обезьян: у людей с поврежденными областями височной коры (например, вследствие кровоизлияния) мы

<sup>1</sup> В данном контексте «вентральный» — значит направленный в сторону ventrum — брюха. — *Прим. пер.*

<sup>2</sup> В данном контексте «дорсальный» — значит направленный в сторону dorsum — спины. — *Прим. пер.*





**Рис. 1.3.32** Два зрительных пути в коре головного мозга обезьян. Нижний (вентральный) путь начинается в затылочной коре и заканчивается в височной доле. Принято считать, что он отвечает за распознавание объектов (система «Что?»). Верхний (дорсальный) путь начинается также в затылочной коре и заканчивается в коре теменной. Принято считать, что он отвечает за местоопределение объектов (система «Где?»). Темно-красным (В и С) выделены участки коры, подвергавшиеся экспериментальному удалению.

дуют есть только с правой стороны тарелки и смогут скопировать только правую часть рисунка (если им дать такое задание) — то есть будут вести себя так, как если бы предметов в левой части поля зрения не было вообще. Несмотря на то, что многие особенности этого синдрома еще не изучены, очевидно, что он связан с нарушениями в работе «Где?»-системы.

То, что происходит при поражении определенных участков мозга помогает понять нормальную физиологию разных его отделов, но, разумеется, не дает детального представления об их работе. Углубленное изучение требует масштабных физиологических исследований с использованием разных методик, включающих анатомическую трассировку нейронных проекций из одной области в другую и запись сигналов от отдельных клеток мозга в тех или иных его отделах (что позволяет определить характер обра-

наблюдаем т.н. *визуальную агнозию* — неспособность зрительно идентифицировать предметы. Одной из форм визуальной агнозии является неспособность больных узнавать лица, причем даже лица супругов, родителей или детей (при этом больной моментально распознает своих родственников по голосу). Данная патология не ведет к недостатку зрительных ощущений — такие больные вполне способны описать лицо, которое видят во всех деталях, включая веснушки, очки и пр. — они лишь не могут ответить на вопрос «*чье оно?*». Сказанное легко истолковать в логике повреждения вентрального проводящего пути, направленного в височную кору, т.е. нарушения функции «Что?»-системы.

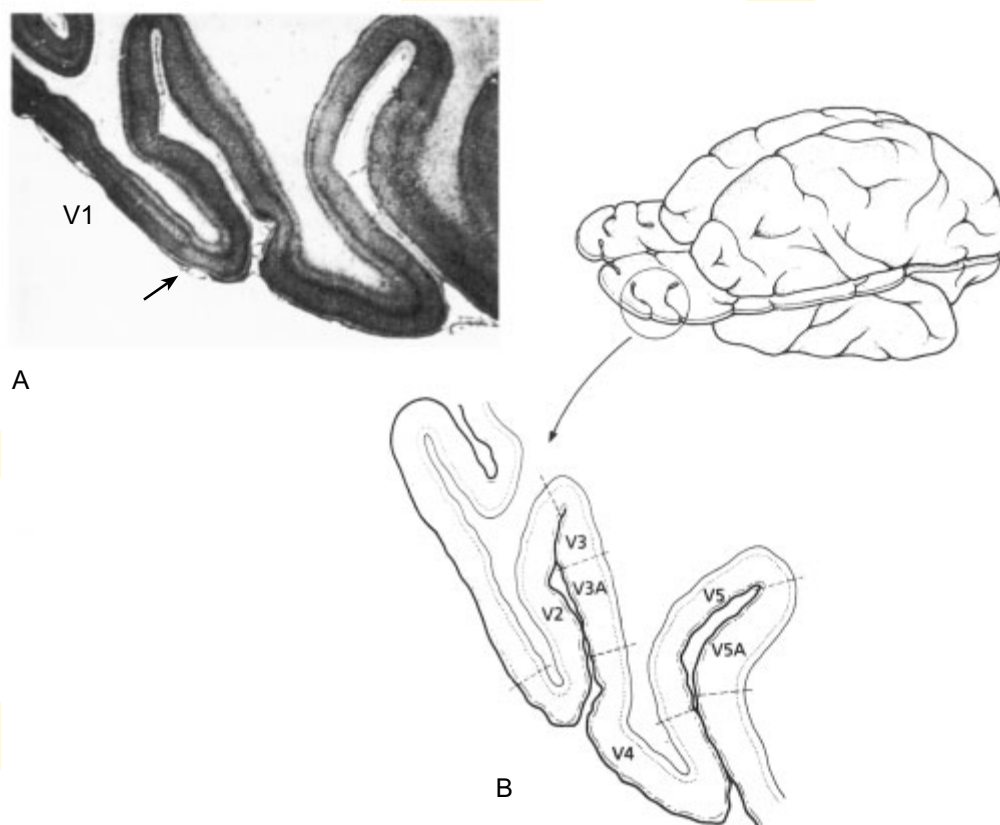
Некоторые больные, у которых повреждены височные доли страдают от т.н. *синдрома одностороннего пространственного игнорирования*. Одностороннее пространственное игнорирование — это комплекс симптомов, который мы подробно обсудим в 11-й главе, а сейчас скажем лишь о его главной особенности — неспособности больного взаимодействовать с предметами, находящимися в той части поля зрения, которая противоположна участку мозгового поражения. К примеру, больные с повреждением правой теменной доли бу-

батываемой ими информации). Мы до сих пор еще очень плохо представляем себе уровень функционирования областей мозга вне первичной зрительной коры, но коротко мы поговорим о том, что все-таки известно.

### 1.3.3.4 Картография зрительной коры

Большая часть коры головного мозга человека и близких к нему приматов скрыта в бороздах. На рис. 1.3.33 показано анатомически достоверное изображение локализации и организации некоторых областей зрительной коры в мозге макаки (поля с V1 по V5). Отметим, что V1–V5 — это часть плотно складчатого слоя корковых нейронов, большая часть из которых не видима на внешней поверхности мозга. Еще раз подчеркнем, что на рисунке показаны лишь некоторые из зрительных областей и многие находятся достаточно далеко от первичной зрительной коры.

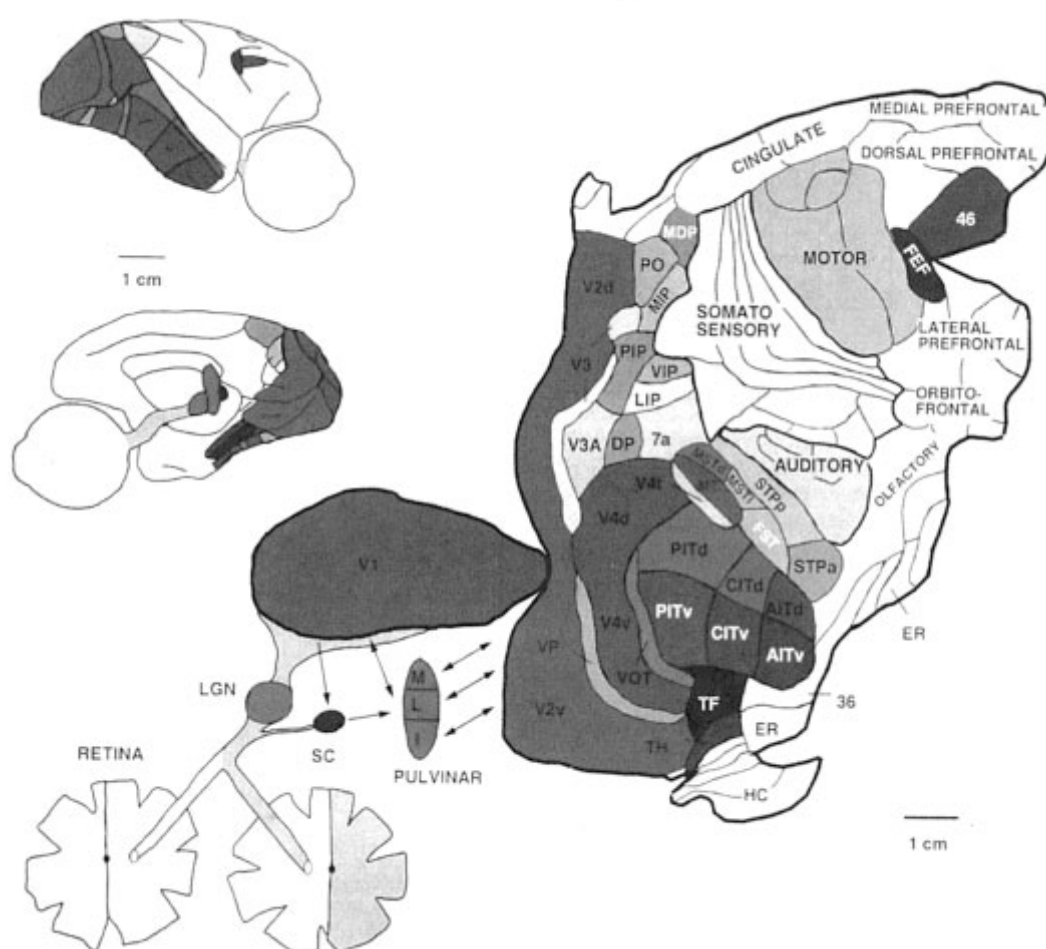
На рис. 1.3.34 показаны известные сегодня зрительные области коры макак: несколько необычное изображение мозга, поскольку складки его поверхности расправлены, в результате чего видны те участки, что в норме скрыты в бороздах; при этом пространственные соотношения между зрительными областями в целом сохранены. Чтобы иметь представление о том, как расположены зрительные области в неповрежденном мозге макаки, слева даны его виды сбоку. Мы видим, что обработка зри-



**Рис. 1.3.33** Локализация первичной зрительной коры в головном мозге макаки. Анатомические позиции стриарной коры (область V1) и некоторых престриарных областей (с V2 по V5) показаны на горизонтальном срезе мозга. Клеточное строение этих областей хорошо видно на участке фотографии (А) от V1 до стрелки, указывающей на престриарные области (Zeki, 1993).

тельной информации происходит целиком в задней половине коры. Отметим, что на уплощенной карте, граница между областями V1 и V2 удалена, благодаря чему размеры различных отделов не сильно искажены.

Итак, первая стадия обработки зрительной информации происходит в стриарной коре (первичной зрительной коре, области V1) обоих полушарий — крупнейшем отделе головного мозга, локализованном в самых задних частях затылочных долей. Стриарная кора получает бóльшую часть восходящих проекций из латерального колленчатого тела и способна выполнять первые операции зрительной обработки. Сегодня мы многое знаем об анатомии и физиологии этого отдела мозга, возможно больше, чем о других. В 4-й главе (в которой мы будем детально обсуждать обработку пространствен-



**Рис. 1.3.34** Развернутая на плоскость карта зрительных областей коры головного мозга макаки. Показаны те участки мозга, о которых известно, что они отвечают за зрительную функцию. *Слева сверху:* латеральная сторона правого полушария головного мозга макаки (вид справа). *Ниже:* медиальная (расположенная к середине) сторона правого полушария. Крупная схема: зрительные области коры макаки, развернутые на плоскость. Край между V1 и V2 удален с целью минимизации пространственных искажений. (Van Essen, Anderson, Felman, 1992).

ной информации, мы поговорим об архитектуре данной части зрительной коры; теперь же мы лишь опишем ее основные анатомические связи).

Поначалу исследователи полагали, что зрительная обработка представляет собой цепь строго последовательных этапов, т.е., что каждый предыдущий отдел, выполнив свою работу, посылает проекцию в следующий и т.д. Но гипотезы уходили в прошлое по мере того как поступали новые данные об анатомических связях между зрительными областями. Сегодня совершенно ясно, что обработка зрительной информации происходит параллельно: каждый отдел посылает волокна в несколько других, но при этом не во все. Упрощенная схема некоторых известных сегодня прямых связей между зрительными областями коры дана на рис. 1.3.35. Стоит сказать, что связи эти в основном двунаправлены, то есть, если область X проектирует данные в область Y, то Y проектирует другие данные назад в область X. Показано однако, что двунаправленные проекции несимметричны: они начинаются и заканчиваются в разных *слоях* коры.

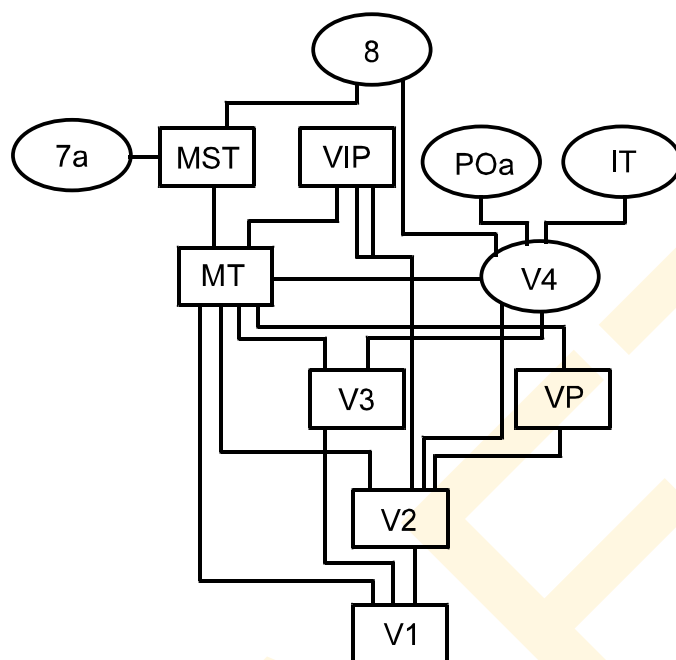
Церебральная кора (т.е. кора головного мозга) — это слоистая, или, как говорят, *ламинарная* структура. В зрительной коре выделяют шесть крупных анатомических слоев с несколькими физиологическими субслоями. При этом четвертый слой, вероятно, является входным для «прямых» или, как еще говорят, «восходящих» проекций из низших отделов нервной системы. Сие доподлинно известно в отношении V1, в которую приходят восходящие волокна из латерального колленчатого тела и образуют синапсы в четвертом слое. По прочим областям не столь очевидно, где «прямые» проекции, где «обратные» («нисходящие»). В одном направлении проекции устойчиво начинаются в поверхностных слоях коры и заканчиваются преимущественно в четвертом слое (рис. 1.3.35А) — такие проекции именуются *прямыми* (по аналогии с таковыми в ЛКТ). Проекции противоположного направления именуются *обратными*, начинаясь и заканчиваясь вне четвертого слоя (рис. 1.3.36В).

Различие между прямыми и обратными связями было использовано при построении иерархии диаграммы рис. 1.3.35: каждая область посажена на уровень, следующий за уровнем, на котором расположена область, пославшая прямую проекцию. Мы видим, что многие зрительные центры буквально опутаны сложными связями (часть из которых до сих пор не исследована). В следующих главах мы будем подробно говорить о том, что известно сегодня о физиологии некоторых наиболее изученных областей коры головного мозга.

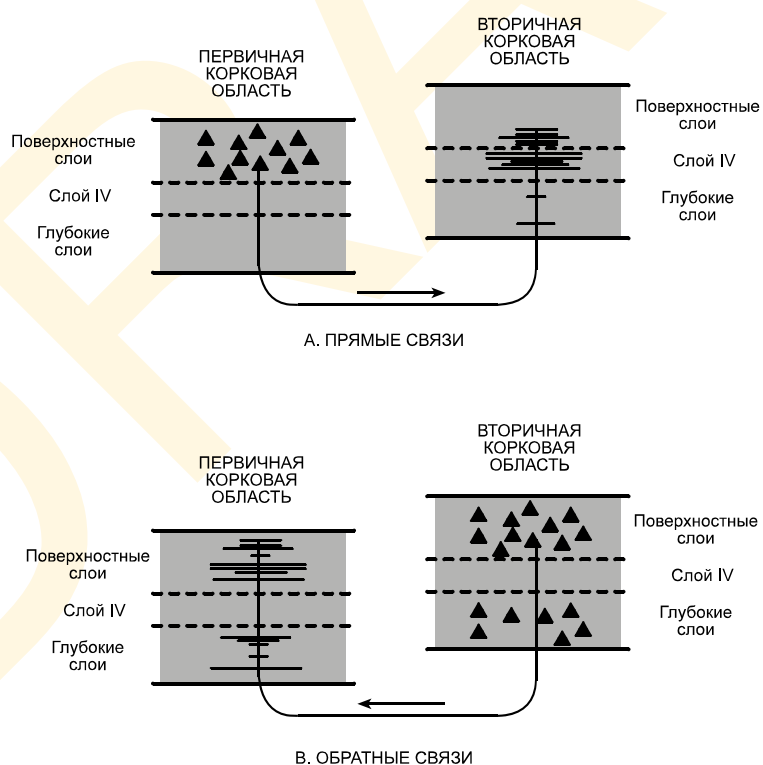
Далее: могут ли сугубо анатомические данные быть полезными в понимании нами функций коры? Известно, что существует корреляция между уровнем клетки в анатомической иерархии (рис. 1.3.20) и размером участка в сетчатке, из которого пришла информация. Клетки низших уровней корковой иерархии получают информацию из малых областей сетчатки, а клетки высших уровней — из больших ее областей. Например, нейроны области V1 могут быть стимулированы сигналом от участка сетчатки размером в 0.1-0.5° зрительного угла; область V2 обычно стимулируется сигналом от участка сетчатки размером в 0.5-1.0°; V4 — 1-4°; область IT — 25° и более (Desimone, Moran & Spitzer, 1988). Таким образом, по мере перехода от низших зрительных областей коры к высшим размер участка стимуляции сетчатки может вырасти в 100 раз и более.

### 1.3.3.5 Гипотеза о физиологических проводящих путях

Новейшие данные о возможных взаимосвязях анатомической структуры (а, вернее говорить — *гистологии*) коры с ее физиологией поступили в последние двадцать лет (или около того). Основная версия состоит в том, что существуют отдельные нейрон-



**Рис. 1.3.35** Некоторые из известных сегодня связей между зрительными областями коры макак (Van Essen & Maunsell, 1983).



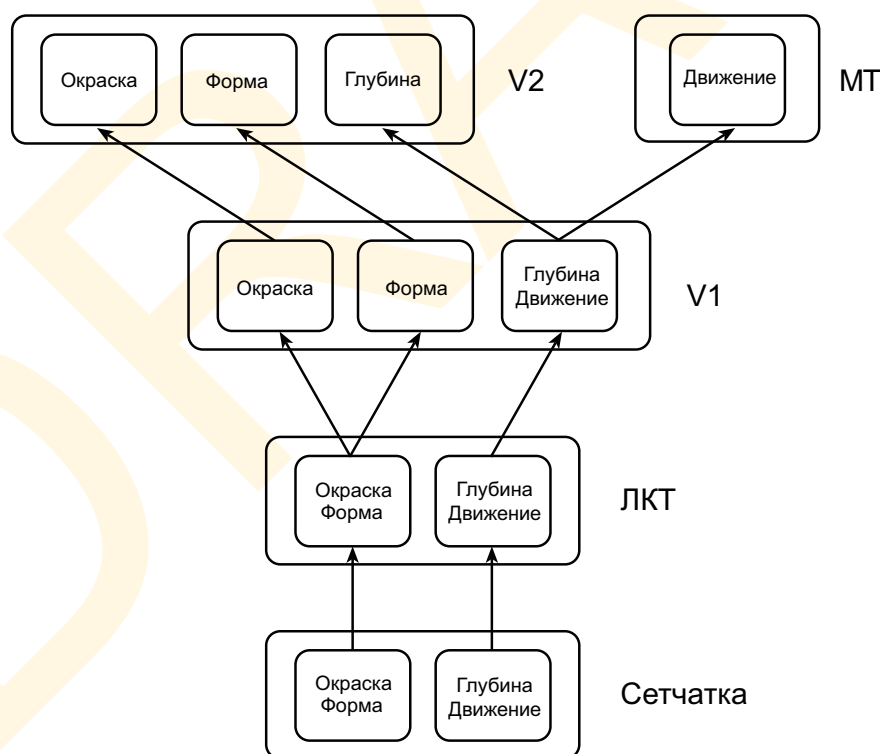
**Рис. 1.3.36** Связи между корковыми слоями. Прямые связи возникают в поверхностных слоях коры и заканчиваются в центральных, главным образом в слое IV. Обратные связи начинаются и заканчиваются вне центральных слоев (Van Essen & Maunsell, 1983).



ные пути (*проводящие пути*), предназначение которых в обработке *профильной* зрительной информации, такой как цветность стимула, его форма, глубина и движение. Идея сия — плод исследований (к примеру Zeki, 1978, 1980), результаты которых указывают на то, что разные области коры «специализируются» на обработке разных визуальных свойств объектов. Позднее стало понятно, что такая специализация очень глубоко «прошита» в зрительной системе: Ливингстон и Хьюбел (Livingstone & Hubel, 1987, 1988) суммировали многие анатомические, гистологические, физиологические и перцепционные данные о коре и предположили, что информация упомянутых четырех категорий обрабатывается в разных проводящих путях, начинающихся в сетчатке. Ученые отследили эти отличия начиная от двух классов ганглиозных клеток (один класс, отвечающий за окраску и форму, второй за удаление и движение) до ЛКТ, и от ЛКТ до разных субрегионов V1, V2 и далее.

Ливингстон и Хьюбел опубликовали доказательства того, что сигнал от единичной ганглиозной клетки, несущий информацию о цветности стимула, его форме, движении и удалении от наблюдателя, обрабатывается в разных субрегионах V1 и V2 (см. рис. 1.3.37). Эти области затем посылают различные проекции в высшие отделы коры:

- информацию о движении стимула и его удалении от наблюдателя — в область V5 (также именуемую МТ — медиальная темпоральная кора);
- информацию о цветности стимула — в область V4;
- о форме — через серию промежуточных центров (включая V4) в область IT (инферотемпоральная кора, нижняя височная кора).



**Рис. 1.3.37** Сильно упрощенная схема, иллюстрирующая гипотезу о зрительных проводящих путях. Некоторые ученые уверенно полагают, что информация об окраске, форме, движении и глубине предметов внешнего мира обрабатывается зрительной системой по отдельности и независимо.

Из перечисленных областей проводящие пути, отвечающие за обработку информации о цветности и форме стимула, могут проектироваться в вентральную «Что?»-систему (где происходит категорийная идентификация), а проводящие пути, отвечающие за обработку информации о его удалении и движении — в дорсальную «Где?»-систему (где выясняется пространственная локализация стимула). Мы вернемся к этой теории в 4-й главе, где будем детально разбирать физиологию зрительной коры.

Характер зрительной обработки в корковых областях высшего уровня изучен гораздо слабее, чем в V1. К примеру, было обнаружено, что некоторые клетки из IT активируются только когда обезьяна видит собственную руку, а другие клетки — когда мордочку (Gross, Rocha-Miranda & Bender, 1972). Характер пространственной обработки между V1 и IT по сей день остается загадкой. Анализ движения в MT-области посылает проекцию в MST — медиальную супериортемпоральную кору (т.е. срединную верхневисочную) и ряд теменных областей. Опять же, о том, как происходит обработка в этих центрах крайне мало известно.

В нашем представлении о функции коры мозга человека еще меньше ясности, чем в таковом о функции коры обезьян. Правда, исследователи нейрофизиологии человека потихоньку выявляют любопытные корреляции между локализациями травм и опухолей со зрительной недостаточностью ими вызванной, но доказательства сложны и зачастую весьма затруднительна их проверка. Применение сканирующих технологий, в частности таких как позитронная томография и функциональная магнитно-резонансная томография (см. раздел 2.2.3) только начинает поставлять данные о корковой локализации зрительных функций у человека, но перспективы, конечно, колоссальны. Мы будем ссылаться на некоторые клинически важные открытия, когда речь пойдет о данном вопросе. Некоторые из наиболее интересных нейробиологических феноменов мы обсудим в последней главе, где поговорим об очень сложной, но при этом исключительно интересной теме — теме зрительного осознания.

Столетиями исследователи тщатся понять, каким образом зрительной системе удается извлекать знание об окружающем мире из световых потоков, попадающих к нам в глаза. Вначале может показаться, что ответ прост и что нужно лишь отыскать соответствующие научные данные. Однако же наука о зрительном восприятии — это очень сложная и обширная область, которая требует не только фактов, но *теории*. Теория — это комплект представлений (*гипотез*) об основополагающих механизмах и принципах, не только организующих и объясняющих уже известные факты, но также прогнозирующих новые.

На страницах нашей книги мы обсудим большое число теорий, объединяет которые то, что все они, в конечном счете... ошибочны. Причина проста: в любой области знания может существовать лишь одна верная теория. Идея неполноты познания и несовершенства научных концепций огорчает многих студентов, которым не терпится во всем поскорей разобраться. Однако беспристрастная проверка теорий и усилия по их совершенствованию — это важнейший компонент научного исследования: даже неверные гипотезы и концепты подчас содержат элементы истины и могут давать очень серьезную пищу для размышлений.

Итак, поскольку суть научного исследования невозможно понять без изучения его теорий (равно как и фактов), настоящая глава посвящена разговору о теориях зрения, разговору, который мы организуем по принципу исторической ретроспекции.

Почему же ученые так полагаются на разные теории? Наиболее полный ответ предложил один из пионеров когнитологии Аллен Ньювелл (Newell, 1973) в своем обращении к участникам конференции по психологии зрения, напомнив при этом публике о ключевой роли научных теорий. Его речь называлась «Двадцать вопросов Природе и Создателю». «Двадцать вопросов» — это некий ритуал догадок, совершаемый особью в процессе анализа зрительного стимула, летящего от интересующего объекта сцены: «Это живое?», «Оно больше, чем буханка хлеба?», «Оно единственное в этом доме?» и т.п. Особь выйдет победителем, если догадки в отношении целевого объекта окажутся верными.

Ньювелл выстроил интересную аналогию между игрой догадок и научными методами: проводя тот или иной эксперимент мы задаем природе вопросы характера «да/нет», но получаем ответ лишь тогда, когда Природа соизволит ответить. Точка зрения Ньювелла была такова, что ответы даже на все возможные экспериментальные вопросы не дадут глубокого научного понимания предмета исследования, поскольку ответы эти никогда не выйдут за рамки обычного перечня фактов, знание которых не достаточно для научного понимания проблемы. В какой-то момент исследователь должен наконец сделать *качественный* скачок от фактов к теории, и только тогда можно выйти победителем в игре догадок, полагал ученый.

Как мы уже сказали выше, научная теория представляет собой комплект гипотез или версий и позволяет давать толкование наблюдаемым явлениям и создавать т.н. *контролепригодные прогнозы* новых. Благодаря своей двойственной роли теория, как *интегратор известных фактов* и *предсказатель новых*, экономичнее простого списка событий и явлений. Теория описывает взаимоотношения между наблюдаемыми фактами и позволяет прогнозировать их, исходя из скромного набора исходных посылов.

К примеру, знаменитый биолог Грегор Мендель, дабы истолковать закономерности появления тех или иных признаков у потомства, предположил, что существуют т.н. *гены*, несущие те или иные наследственные признаки. Как только ученый сформулировал теорию генетических комбинаций, так сразу методика прогнозирования наследственных признаков перестала нуждаться в громоздких таблицах вероятностей. Иными словами, вероятности появления тех или иных признаков стали выводиться из нескольких базовых гипотез генетической теории.

Одним важным следствием того, что теории представляют собой интегрированный набор версий или гипотез, является требование *внутренней последовательности* как самих гипотез, так и выводов из них. Теории, несущие в себе внутренние противоречия, априори некорректны и должны категорически отвергаться (кстати говоря, философы всегда озабочены поиском логических противоречий в той или иной теории). Однако логическая последовательность — это не единственный критерий верности теории: последняя должна, как говорят, удачно ложиться на большое число известных в данной области фактов, а в идеале — на все. Определяя эмпирическую пригодность той или иной научной концепции, исследователи тестируют ее прогнозы и сравнивают их со своими наблюдениями. И только ежели теория окажется одновременно и логичной, и стоящей при этом на крепкой эмпирической основе, ее начинают рассматривать всерьез.

Можно показать, что эти два критерия — логическая последовательность и эмпирическая достоверность — вместе являются достаточным основанием к тому, чтобы признать теорию «корректной» с научной точки зрения.

К сожалению, логически рассуждая, можно показать, что когда некая внутренне цельная теория адекватна некоему комплексу данных, то наверняка найдется множество аналогичных теорий. Порой две (даже более) принципиально отличных друг от друга теории могут подходить одним и тем же данным. Как при этом понять, какая из них верная? Для ответа на этот вопрос ученые используют старинный принцип «оккамовой бритвы», гласящий, что лучшая теория — это та теория, что наиболее *экономична*, то есть та, что согласуется с эмпирическими данными при наименьшем количестве версий. Следовательно, наряду с эмпирической достоверностью, экономия есть критерий оценки научных теорий.

## 2.1 КЛАССИЧЕСКИЕ ТЕОРИИ ЗРЕНИЯ

Известный гештальт-психолог Курт Коффка (Kurt Koffka) пожалуй лучше чем кто бы то ни было озвучил проблему всех теорий зрительного восприятия, задав единственный и на первый взгляд элементарный вопрос: «Почему предметы выглядят так, а не иначе?». Было предложено множество ответов, но в основе их всех лежат четыре «классических» размерности, составляющих суть всех ныне существующих физиологических теорий зрительного восприятия:

### 2.1.1 Общая схема

#### 2.1.1.1 Размерность первая: внешний мир ↔ организм

Один из возможных ответов на коффовский вопрос в рамках этой размерности звучит так: «Потому что так устроен мир». Сие защищает ту позицию, что анализ характера внешней стимуляции — это основной путь в понимании механизмов воспри-

ятия. В частности речь идет о том, что особь извлекает разного рода информацию из проксимального стимула, соответствующего стимулу дистальному.

Оппонентный ответ звучит так: «Потому что мы так устроены», «Поскольку наша зрительная система такова». Такая позиция отстаивает внутренний подход к воспринимаемой структуре, делая акцент на физиологии организма, нежели на свойствах внешнего мира. Промежуточные позиции, возможно, суть оправданный компромисс, свидетельствующий о важности как структуры стимула (внешние факторы), так и внутренней структуры организма (внутренние факторы).

### 2.1.1.2 Размерность вторая: эмпиризм ↔ нэйтивизм

Первый вариант ответа на коффовский вопрос в рамках второй размерности звучит так: «Потому что мы так научились видеть». Сие точка зрения *эмпириков*, полагающих, что мы видим так, как мы видим, потому что познание мира происходит в процессе *взаимодействия* с ним.

Оппонентный ответ в этом случае звучит как «Потому что так видеть мир — это наша врожденная способность». Сие позиция *нэйтивистов*, согласно которой у нас нет необходимости получать специфическое знание на протяжении жизни, поскольку эволюция снабдила нас всеми необходимыми нервными механизмами. Нэйтивистическая точка зрения в сущности такова, что все биологические виды в процессе эволюции приобрели должный опыт, и что каждая отдельно взятая особь не есть *tabula rasa*,<sup>1</sup> ожидающая, что опыт этот будет накапливаться за счет зрительных впечатлений.<sup>2</sup> Организмы, по мнению нэйтивистов, не нуждаются в том, чтобы обладать врожденной способностью к обучению, но что все дальнейшее развитие особи генетически запрограммировано.

### 2.1.1.3 Размерность третья: атомизм ↔ холизм

Первым вариантом ответа в рамках третьей размерности является следующий тезис: «Потому что так воспринимается каждый из фрагментов поля зрения». Согласно концепции *атомизма* результирующее восприятие цельного поля зрения прогнозируется простым сложением элементов зрительного опыта по каждому из отдельных участков сцены.

Сей подход суть антитеза следующей сентенции: «Поскольку таково цельное восприятие поля зрения». Такую позицию мы назовем *холистической*,<sup>3</sup> т.е. утверждающей то, что восприятие того или иного фрагмента поля зрения жестко зависит от влияния соседних. Отдельные области поля *не аддитивны*, то есть не могут просто складываться друг с другом, но подлежат т.н. *глобальной интеграции*. Согласно холистической точке зрения зрительная система так организует перцепты стимулов, что воспри-

<sup>1</sup> Tabula rasa (лат.) — чистый лист. — Прим. пер.

<sup>2</sup> Однако эволюционная точка зрения полагает важным умение отличать приобретенное от врожденного, поскольку механизмы обучения особи и вида принципиально отличны друг от друга. Ассоциативные механизмы обучения, которые мы обсудим в дальнейшем, относятся только к процессу обучения особи. Теория эволюции гласит, что обучение через выживание свойственно виду в целом, хотя, конечно, метафорически приложимо и к особи (Edelman, 1987).

<sup>3</sup> От англ. «whole» — целый, цельный. — Прим. пер.



ятие объектов и даже целых сцен вытесняет собой результаты восприятия отдельных участков поля зрения.

#### 2.1.1.4 Размерность четвертая: интроспекция ↔ бихевиоризм

Четвертой критично важной размерностью, задаваемой теориями зрительного восприятия является размерность *методологическая*, интересующаяся тем, как Коффа поставил вопрос, почему одну из точек зрения он предпочел другой.

Первая из этих двух концепций — следствие осознанного феноменологического анализа собственных зрительных ощущений и зрительного опыта, вторая — плод объективных измерений результатов поведения человека, его поступков. Коффа, произнеся «Почему предметы *выглядят* так, а не иначе?», фактически явился предтечей интроспективного подхода, поскольку «выглядят» — плод осознанного зрительного опыта. Бихевиористы же спросят: «Что зрение позволяет нам *делать*»? Сие безусловно законный вопрос, но принципиально отличный от коффского.

\* \* \*

Итак, памятуя об описанных четырех размерностях, рассмотрим четыре базовых теории зрительного восприятия и расскажем об ученых, отстаивавших эти теории (см. рис. 2.1.1).

Три из четырех: *структурализм*, *гештальтизм* и *экологическую оптику*, легко отличить по четырем базовым посылам (см. рис. 2.1.2). Теории эти разнятся также и по теоретическим аналогиям, подходящим к каждой из них.

Четвертая физиологическая теория восприятия — *конструктивизм* — в какой-то степени являет собой смесь предыдущих трех. Конструктивизм представляет самый современный в психологии подход к вопросам восприятия и апеллировать к нему мы будем на протяжении всей нашей книги.

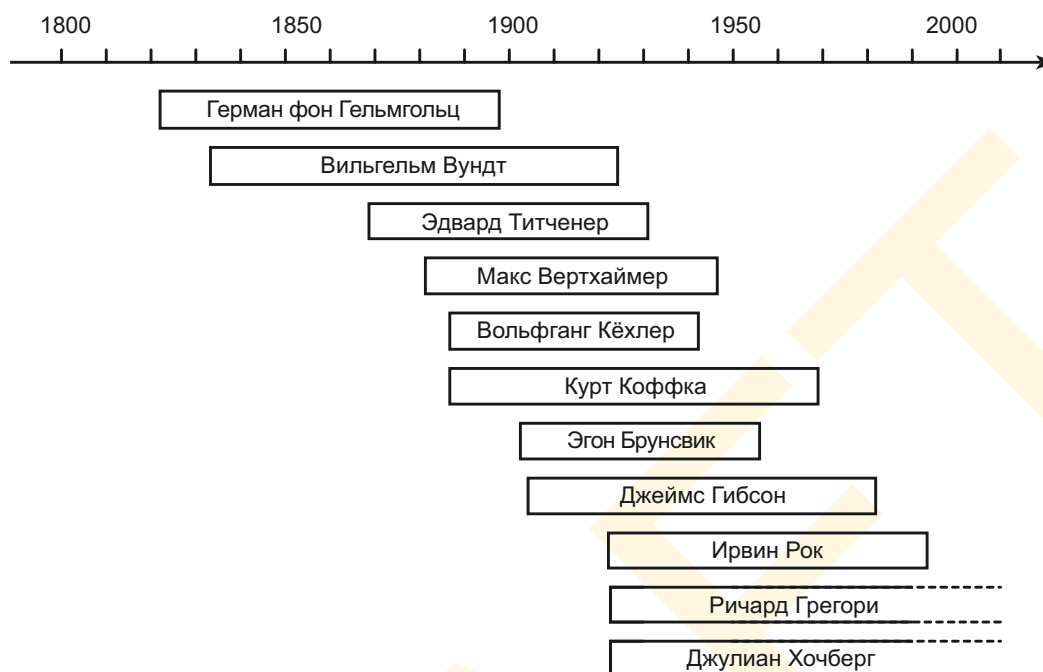
### 2.1.2 Структурализм

Первой психологической концепцией восприятия явился *структурализм*, корнями уходящий в т.н. *британский эмпиризм* — философскую школу, апологетами которой являлись в свое время Лок, Беркли и Хьюм. Отцом основателем структурализма считается немецкий психолог Вильгельм Вундт (см. рис. 2.1.3), проповедником идей которого в США выступил один из учеников — Эдвард Титченер.

#### 2.1.2.1 Атомизм и эмпиризм

Точка зрения структуралистов сводилась к тому, что восприятие начинается с некоего процесса, в котором т.н. *сенсорные атомы* — примитивные невидимые элементы опыта в данной сенсорной модальности — вызывают воспоминания о других сенсорных атомах, которые, объединившись в памяти, образуют т.н. *первичные объединенные события*.

Структуралисты полагали, что такие объединения возникают всегда, когда особь через достаточное количество стимуляций достигает полноты сенсорного опыта в данном пространстве и на данном отрезке времени. В отношении зрения структуралисты мыслили так, что сенсорные атомы — суть цветовые ощущения, возникающие по множеству сверхмелких участков поля зрения от возбуждения отдельных фоторецепторов в сетчатке. Предполагалось, что локальные ощущения объединяются в воспри-



**Рис. 2.1.1** Периоды жизни ученых, оказавших наибольшее влияние на визиологию. Вундт и Титченер — структуралисты; Вертхаймер, Кёхлер и Коффка — гештальтисты; Брунsvик и Гибсон — экооптики; Гельмгольц, Рок, Грегори и Хочберг — конструктивисты.

ятии в простую цепочку или как говорили — *конкатенацию*. Проще говоря: формирование зрительной картины, по мнению структуралистов, происходило путем простой суммации цветовых точек определенной пространственной локализации.

Полагали также, что зрительные опыты ассоциативным путем рожают воспоминания из иных сенсорных модальностей. Воспоминание о том, как выглядит собака ассоциируется с тем, как она лает, чем пахнет и т.п. И как выглядит фрагмент собаки, скажем ее голова, в сознании наблюдателя может ассоциироваться с видом иных частей ее тела: лап, туловища, хвоста и пр. Структуралисты полагали, что восприятие возникает за счет очень быстрых и неосознаваемых ассоциативных процессов, возникающих от огромного опыта общения с предметным миром. Согласно структуралистской точке зрения, чем больше наблюдатель узнает о мире через ассоциации, тем восприятие его становится богаче, точнее и организованнее.

Представление структуралистов об «атомарной» структуре сенсорного опыта очень напоминает постулаты химии, например, что соотношения между простейшими ощущениями (скажем, опыт осознания красноты данного участка поля зрения) и сложным перцептом (таким как, к примеру, перцепт яблока) таковы же, что и соотношения между атомами и сложными молекулами.<sup>1</sup> «Клеем», сцепляющим отдельные ощущение-

<sup>1</sup> Тот факт, что новейшие открытия в физике свидетельствуют о том, что и сами атомы состоят из элементарных частиц (протонов, электронов и пр.) никак не умаляет ценность атомистического подхода. Фактически это лишь укрепляет последний, поскольку новейшие теории, возникшие в результате расщепления атома свидетельствуют о существовании еще меньших атомарных компонентов. Сегодня предельной малой частицей полагают *кварк* — мельчайший элемент материи, но при этом базовые гипотезы атомарной теории стали лишь прочнее.

НАУЧНОЕ НАПРАВЛЕНИЕ	Эмпиризм ↔ Нэйтивизм	Атомизм ↔ Холизм	Внешний мир ↔ Организм	Аналогия	Метод исследования
Структурализм	Эмпиризм	Атомизм	Организм	Химия	Тренированная интроспекция
Гештальтизм	Нэйтивизм	Холизм	Организм	Физическая теория поля	Простая интроспекция
Экологическая оптика	Нэйтивизм	Холизм	Внешний мир	Механический резонанс	Анализ стимула

**Рис. 2.1.2** Три научных направления в психологии зрительного восприятия — структурализм, гештальтизм и экологическая оптика, охарактеризованные по пяти критериям:

- позиции в размерности «Эмпиризм ↔ Нэйтивизм»;
- позиции в размерности «Атомизм ↔ Холизм»;
- позиции в размерности «Внешний мир ↔ Организм»;
- подходящей аналогией из классических дисциплин;
- методом исследования.

ния в общий, результирующий перцепт, являются, по мнению структуралистов, пространственная и временная близость ассоциаций в последующих сенсорных опытах организма. Итак, мы видим, что сердцем структуралистической теории восприятия являются атомизм и эмпиризм.

### 2.1.2.2 Тренированная интроспекция

Еще одним основанием, на котором структуралисты строили свой подход к проблемам восприятия, являлся метод *тренированной интроспекции*. Структуралисты полагали, что каждый может исследовать элементарные составляющие восприятия, обратив мысленный взор внутрь себя (интроспекция) и тщательно исследовав собственные опыты, однако при этом они считали, что интроспекции следует учиться у опытного наставника. К сожалению, характер этого обучения зачастую целиком зависел личных взглядов этого наставника (Boring, 1953), к примеру: в контексте разговора о восприятии изображений, некая группа структуралистов утверждала, что ментальный компонент восприятия — суть поток сенсорных образов; в то время как оппонировавшие им структуралисты утверждали, что таких образов вообще не существует! Нестыковки и непоследовательность подорвали доверие к технике тренированной интроспекции, которую, в итоге, заменили более надежными и проверенными психофизическими методами, в частности описанными в Приложении А.



**Рис. 2.1.3** Вильгельм Вундт — основоположник научной психологии и один из главных приверженцев структурализма (Gleitman, 1981).

Итак, поскольку не имея в своем распоряжении современных научных данных, структуралисты пытались напрямую транслировать философские взгляды британских эмпиристов на активно развивающуюся психологическую дисциплину, структурализм, пожалуй, стоит рассматривать как промежуточный этап между ранним философским периодом в истории теории восприятия к более сложному психологическому периоду.

### 2.1.3 Гештальтизм

Исторически философия гештальта возникла как антитеза структурализму. Лидерами гештальтистов явились Макс Вертхаймер (рис. 2.1.4), Вольфганг Кёлер и Курт Коффка, отвергавшие всё, что имело хоть какое-то отношение к структурализму: теоретические посылы атомизма и эмпиризма, химические аналогии и, разумеется, метод тренированной интроспекции. Полемически гештальтисты оказались куда успешнее своих предшественников, но при этом гораздо менее удачливы в популяризации своих воззрений.

#### 2.1.3.1 Холизм

*Гештальт* — немецкое слово, означающее «целостная форма; образ; целое; цельное», указывающее на то, что гештальтисты гневно отметали структуралистическую идею того, что восприятие зиждется на локальных сенсорных атомах, объединяемых

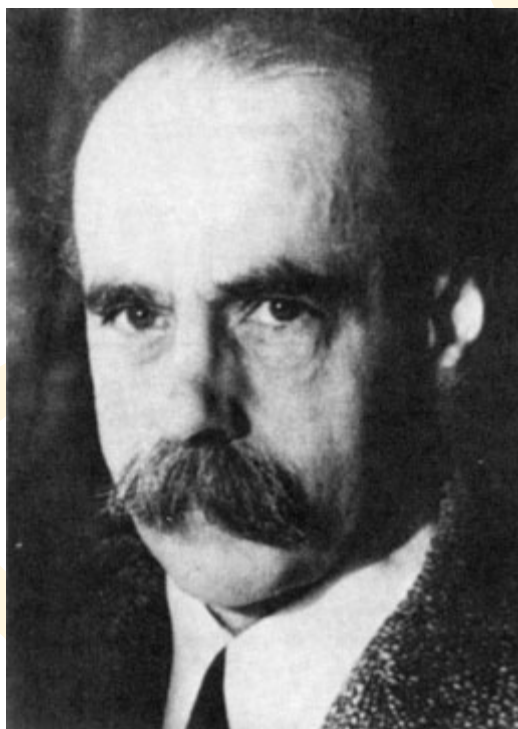
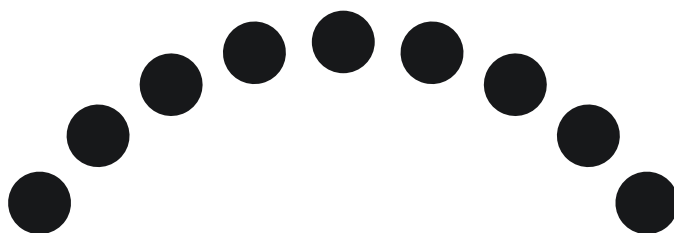


Рис. 2.1.4 Макс Вертхаймер — один из основателей гештальт-направления в психологии, внесший существенный вклад в сегодняшнее понимание механизмов восприятия (Goldstein, 1989).

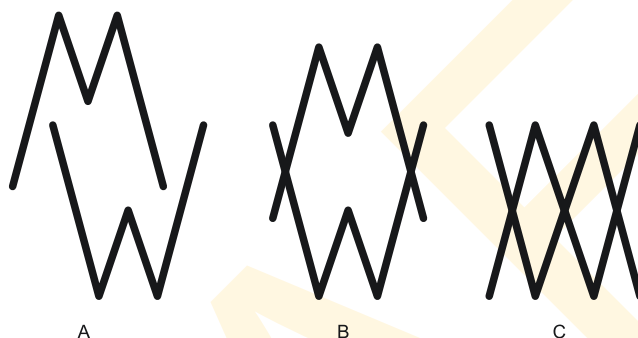
простой суммацией. Хорошо известен их слоган: «Целое не есть сумма его частей» как символ веры гештальтистов в то, что внутренняя структура восприятия — это нечто целое и неделимое на части, более того — в то, что не существует никаких взаимоотношений между условными частями этого целого.

В защиту холизма гештальтисты приводили примеры, в которых целое обладает т.н. *эмерджентными свойствами* — новыми свойствами, коих нет у частей, составляющих это целое. Проиллюстрировать это очень просто с помощью линии, построенной из отдельных точек (рис. 2.1.5): сами по себе точки обладают лишь тремя перцепционными (чувственными) параметрами: цветом, размером и позицией. Но когда множество точек объединяется в линию, новое образование приобретает дополнительные качества, такие как длина, пространственная ориентация и кривизна. Эти свойства появляются только у целого, то есть, только тогда, когда точки уложены в линию.

Примеры, аналогичные приведенному, легли в основу отрицания гештальти-



**Рис. 2.1.5** Эмерджентные свойства композиции: выстраивание отдельных кружков в линию обуславливает появление у нее эмерджентных качеств, таких как длина, направление и кривизна — свойств, которыми не обладают сами кружки.



**Рис. 2.1.6** Эффекты перцепционной организации: буквы «М» и «W» легко читаются в секции А; несколько хуже в секции В и практически не читаются в секции С, что суть результат различий в пространственных отношениях между частями. Психологам — гештальтистам известно множество аналогичных феноменов.

стами структуралистических посылов, поскольку на самом деле простая сумма частей крайне редко приводит к восприятию целого. Следовательно, гештальтисты в первую очередь изучали и описывали те аспекты восприятия, что определялись совокупностью качеств целого предмета — т.е. *гештальтом*. В частности, их интересовал сам путь, по которому целое организуется из субчастей (подробнее см. гл. 6). К примеру, почему мы легко и быстро распознаем буквы «М» и «W» на рис. 2.1.6А; чуть с большим трудом на рис. 2.1.6В и с очень большим усилием на рис. 2.1.6С? Во всех трех случаях буквы физически одни и те же, но на рис. 2.1.6В они расположены так, что накладываются друг на друга, а на рис. 2.1.6С так, что образуют некое новое целое, в котором фрагменты «М» и «W» являются продолжением друг друга.

Итак, гештальтисты первыми осознали исключительную значимость структуры восприятия и первыми начали анализировать ее свойства.

Неудивительно, что гештальт-психологи наряду с атомистической отвергли и классическую химическую аналогию структурализма. Они предпочли рассуждать о ментальных процессах, проводя аналогию с физическими силовыми полями, таким как магнитное или электрическое. Критично важным аспектом физических полей, к которому апеллировали гештальтисты, являлся их холистический характер. К примеру, единичная положительно заряженная частица образует электрическое поле, распространяющееся в пространстве как показано на рис. 2.1.7А. Но если мы добавим еще одну заряженную частицу, то структура результирующего поля изменится, и характер этого изменения окажется зависим от характера отношений между частицами (рис. 2.1.7В и 2.1.7С). Приведенный пример был коньком гештальтистов, поскольку,



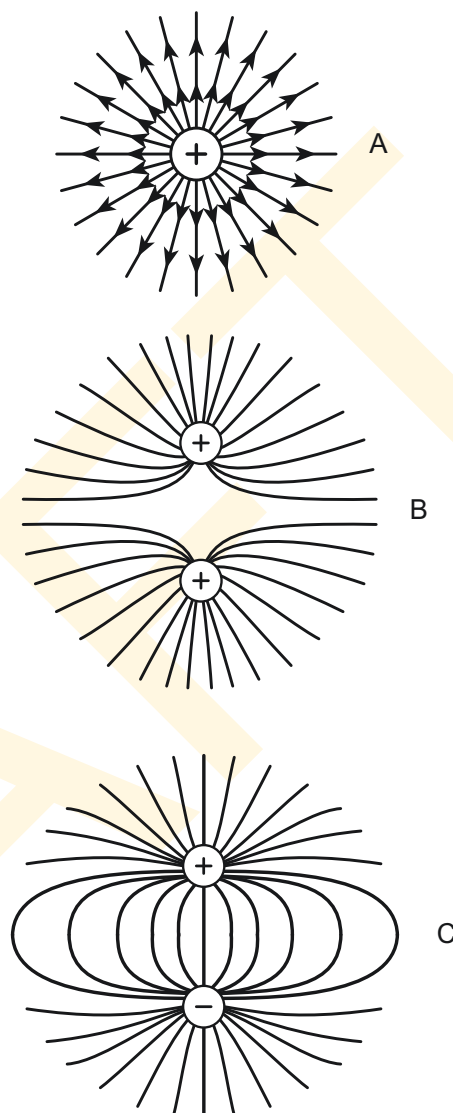
напомним, их пафос строился на отрицании структурализма и обретении целым качественно новых свойств, не присущих слагаемым этого целого.

Со временем гештальт-психологи пошли дальше и отвергли эмпиризм как базис восприятия. К примеру, они были уверены, что механизмы перцепционной организации (в частности, те, что реагируют на непрерывность линий на рис. 2.1.6С) не требуют обучения опытом, но возникают вследствие взаимодействия структур мозга со структурой стимула. Однако нэйтивистическая позиция гештальт-психологов не занимала центрального места в их философии — опорой им по-прежнему оставался холизм. Возможно, наиболее точной и полной формулировкой позиции гештальтистов-психологов является следующая: цель нэйтивизма-эмпиризма — в отрицании главной структуралистической идеи о ведущей роли опыта в восприятии; а также в утверждении того, что базисом восприятия являются процессы, не требующие обучения.

### 2.1.3.2 Психофизиологический изоморфизм

Несмотря на то, что в контексте зрительного восприятия главным предметом исследования гештальт-психологов являлась структура изображения, стимулирующего орган зрения, они апеллировали в основном к мозговым механизмам, т.е. механизмам *внутри организма*. Гештальтисты строили свои концепции на взаимоотношениях между психикой и мозгом, что сформулировано в доктрине т.н. *психофизиологического изоморфизма*. Суть последнего в том (Wertheimer, 1912; Kohler, 1947), что опыты восприятия (психология) структурно схожи с (изоморфны) лежащим в их основе мозговым процессам (физиология).<sup>1</sup>

Типичной иллюстрацией идеи психофизиологического изоморфизма является теория *оппонентной обработки* при цветовом восприятии стимулов, предложенная



**Рис. 2.1.7** Холизм силовых полей. Самостоятельная положительно заряженная частица обладает определенным силовым полем (А); после добавки еще одной заряженной частицы (В и С) результирующее поле меняется в зависимости от заряда этой дополнительной частицы.

<sup>1</sup> Предшествующие гипотезы в этом направлении предложили Мач (1865-1965) и Геринг (1878-1964), а также Мюллер (1896). Однако идеи ученых были не столь глубоки как рассуждения гештальт-психологов.



**Рис. 2.1.8** Постобразы. Примерно на одну минуту зафиксируйте взгляд на белой точке в центре левого изображения, а затем переведите его на черную точку в центре правого: возникнет иллюзия, в которой мы узнаем традиционный бело-сине-красный флаг Соединенных Штатов.

немецким исследователем Германом Герингом (Hering, 1878/1964). Суть геринговской теории в том, что существует шесть физиологически первичных и однозначных цветовых ощущений, объединенных в три пары: красное ↔ зеленое, синее ↔ желтое и черное ↔ белое. Для иллюстрации сказанного как минимум на 30 сек. сосредоточим взгляд на черной точке в центре правой половины рис. 2.1.8, а затем быстро расслабимся и отведем взгляд в сторону. Мы увидим т.н. *постобраз* американского флага в его привычном красно-бело-синем виде, а не зелено-черно-желтом, как на левом изображении. Вместе с прочими феноменами, приводящими к тем же выводам (см. раздел 3.2.1), хроматическая оппозиционность постобразов оригинальному восприятию (красный ↔ зеленый, синий ↔ желтый, черный ↔ белый) говорит в пользу геринговской оппозиционной теории.

Приняв во внимание оппозиционную структуру опыта цветового восприятия, гештальт-доктрина психофизиологического изоморфизма пришла к тому, что должна существовать некая соответствующая оппозиционная структура нейронных процессов, лежащая в основе цветового восприятия световых потоков. Сегодня на руках у нас имеется прочное доказательство верности такого предположения: как будет показано в разделе 3.2.3, действительно в зрительной системе человека существует три типа нейронов, кодирующих информацию о фотометрической яркости и цветности стимула по трем парам оппозиционных сигналов: красно-зеленому, сине-желтому и черно-белому, то есть так, как и предполагал в свое время Геринг (DeValois & Jacob, 1968). Таким образом, согласованность опыта переживания цветовых ощущений с оппозиционными нейронными процессами подтверждает гештальт-теорию психофизиологического изоморфизма. Однако сам по себе психофизиологический изоморфизм не в силах положить основу нейрологической теории функций мозга.

Немецкий психолог Вольфганг Кёлер (1920, 1950) глубже исследовал взаимосвязь между теоретическими положениями гештальт-психологии и лежащими в их основе мозговыми механизмами. Ученый предложил рассматривать мозг как пример того, что мы называем *физическим гештальтом*, т.е. как динамическую физическую систему, стремящуюся к равновесию с минимумом энергетических затрат. Для иллюстрации кёлеровской идеи гештальтисты часто прибегают к аналогии с мыльными пузырями, любопытнейшее свойство которых в том, что вне зависимости от исходной формы, со временем они неизбежно превращаются в правильные сферы: местные напряжения и растяжения на поверхности пузыря, в конечном итоге, равномерно распро-

страняются по этой поверхности, поэтому рано или поздно пузырь переходит в состояние полной стабильности, при котором распределение сил на его поверхности требует минимума энергии. Множество прочих физических гештальтов определено через данное динамическое свойство, и гештальтисты были уверены, что мозг один из них.

Вторая специфическая идея Кёлера состояла в том, что причинные мозговые механизмы восприятия — это электромагнитные поля, возникающие в результате электрических процессов в миллионах нейронов. Такие электрические мозговые поля являлись очень важным свойством физического гештальта и во времена Кёлера полагались физиологически достоверными. Эксперименты по проверке этой концепции, выполненные позднее, показали, что в противоположность кёлеровским прогнозам, разрозненные электрические поля головного мозга не оказывают серьезного влияния на способность к восприятию: ни покрытие коры высокопроводящей золотой фольгой (Lashley, Chow, & Semmes, 1951), ни внедрение электрически изолированных слюдяных электродов в кору (Sperry & Miner, 1955) не дали эффектов, прогнозируемых теорией электрических мозговых полей. Кёлер попытался оспорить результаты экспериментов в серии статей, но в целом было уже поздно и, как это принято, экспериментальные данные обрели решающее значение для научного сообщества. Итак, в основном потому, что гештальт-теория не получила физиологического подтверждения, интерес к ней утратили довольно быстро. Однако гештальт-идеи сегодня возвращаются к нам под видом т.н. *теории динамических связей*, что мы покажем в дальнейшем.

#### 2.1.4 Экологическая оптика

Еще одной широко известной теорией зрительного восприятия является концепция, предложенная преподавателем корнельского университета Джеймсом Джеромом Гибсоном (см. рис. 2.1.9). Подобно гештальтистам Гибсон решительно противостоял структурализму и был весьма подвержен идеям гештальта, в частности его холистической направленности. Гибсоновское отрицание структурализма в чем-то идет дальше гештальтного: ученый отвергал даже ту идею, что в основе восприятия лежит



**Рис. 2.1.9** Джеймс Джером Гибсон — основоположник экологической теории зрительного восприятия. Его работы по извлечению информации из структуры ретинальных изображений предзнаменовали современные разработки в сфере компьютерного зрения (Gleitman, 1981).

соматика (т.е. органное устройство). Более того, он продвигал мысль, что суть перцепта можно полнее истолковать, если проанализировать структуру пространства, окружающего особь, т.е., как говорят — *экологию* этой особи. В одной из статей о гибсоновском подходе к изучению восприятия было даже сказано: «Восприятие — это не вопрос о том, что творится в нашей голове, но о том, что происходит вокруг нее» (Масе, 1977). В сущности, гибсоновская теория — это теория *информационного базиса* восприятия, нежели его *базиса физиологического*. Отметим, что сей подход, радикально отличный от всех предыдущих, положен в основу разработок в сфере компьютерного зрения.

### 2.1.4.1 Анализ структуры стимула

Цель гибсоновской экологической теории восприятия состояла в том, чтобы описать, как объекты реального мира структурируют отражающийся от них свет в т.н. *внешний оптический массив* — *ВОМ* (см. раздел 1.2.1), воздействующий на зрительную систему человека. Гибсон показал, что характеристики *проксимального* стимула несут информацию о стимуле *дистальном*. Ученый настаивал на том, что цельный паттерн проксимальной стимуляции несет в себе гораздо больше информации о дистальном стимуле, чем ранее предполагалось, к примеру: Гибсон отмечал, что когда поверхность с равномерной структурой (дистальный стимул) расположена под углом к наблюдателю, ее изображение на сетчатке (проксимальный стимул) образует т.н. *текстурный градиент*, в котором (по мере удаления от наблюдателя текстурных составляющих этой плоскости) плавно уменьшаются пространственные размеры соответствующих текстурных элементов, растет их пространственная плотность и меняется форма (см. рис. 2.1.10).

Многие объекты с однотипной фактурой поверхности (например газоны, каменные стены, фактурные потолки, кафельные полы т.д.) создают на сетчатке наблюдателя равномерные текстурные градиенты. Гибсон полагал эти градиенты источниками пространственной информации (т.е. информации об удалении объектов от наблюдателя, их уклоне и размере), т.е. старался показать, что человек воспринимает эти свойства предметов, когда на сетчатке образуются соответствующие градиенты (Gibson, 1950).

Теорию экологической оптики важно разобрать еще и потому, что она толкует восприятие как процесс активного исследования особью окружающей обстановки, и, как следствие, обретения ею осведомленности об этой обстановке. Когда наблюдатель движется в пространстве, то паттерн стимуляции на сетчатке все время меняется: Гибсон придавал особе значение плотности стимулирующего оптического потока и даже внедрил концепцию т.н. *динамического ВОМ* (см. раздел 1.2.1). Здесь стоит сказать, что



**Рис. 2.1.10** Текстурные градиенты. Когда поверхность состоит из множества однотипных объектов, уходящих вдаль, она создает на сетчатке текстурное изображение, элементы которого плавно уменьшаются в размере, уплотняются и укорачиваются.



большинство предшествующих работ по зрительному восприятию базировались на исследованиях, в которых голова наблюдателя была иммобилизована (жестко фиксирована в пространстве), а стимул предъявлялся в сугубо искусственных и весьма ограниченных условиях просмотра. Но Гибсон и его последователи убедительно показали, что системы восприятия развивались у организмов благодаря *движению* во время поиска еды, воды, пары и убежища. Следовательно, теория экологической оптики описывает то, по какой схеме световые потоки, несущие информацию об окружающем мире, воздействуют на глаз *движущегося* наблюдателя.

Одним из наиболее важных и самобытных гибсоновских открытий, сделанных им в процессе изучения динамических внешних оптических массивов, является то, что та же самая информация, что описывает характер окружающей особь обстановки, описывает и траекторию движения наблюдателя в этой обстановке. К примеру, последовательность видов на рис. 2.1.11 указывает не только на то, что дверной проем ведет в другую комнату, но также свидетельствует о приближении к нему самого наблюдателя. То есть, и восприятие окружающей обстановки, и характер движения точки наблюдения определяются характером оптической информации, поступающей в глаз.

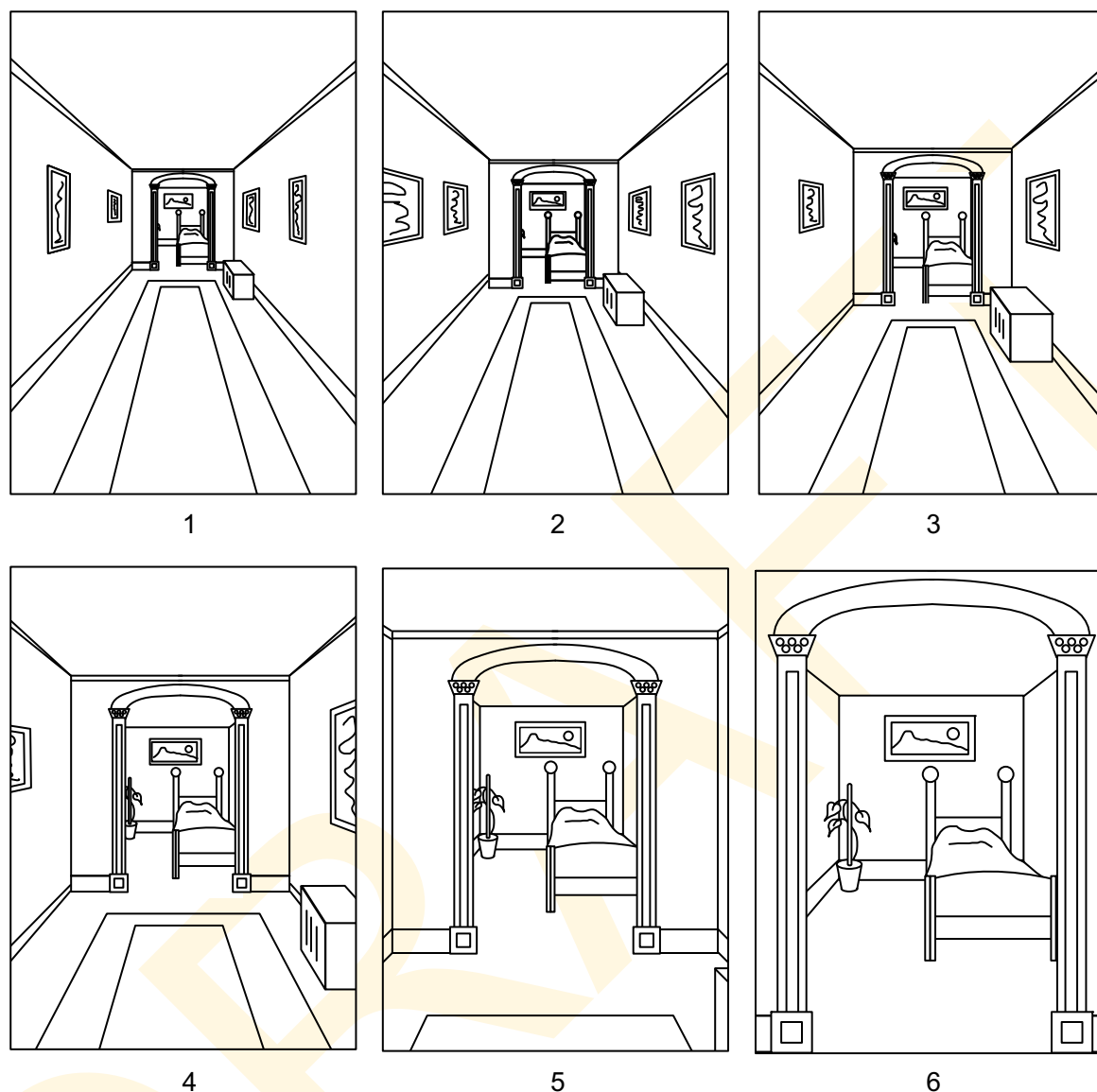
Гибсон был уверен, что информации, извлекаемой активно действующим организмом из проксимального стимула, вполне достаточно для однозначного зрительного восприятия окружающего мира (процесс извлечения данных из ретинальной стимуляции ученый назвал *информационным захватом*). Несмотря на то, что Гибсон не столь глубоко интересовался биологическими механизмами восприятия, но более информативностью стимулов, он предложил т.н. *резонансную модель*, толкующую процесс этого захвата: Гибсон уподоблял мозг камертону и говорил о том, что все перцептивные функции центральной нервной системы сравнимы с механическим резонансом (Gibson, 1966, 1979). Информация, заключенная в стимуле, по мнению ученого, возбуждает соответствующие нейронные структуры мозга почти так же как механическое колебание воздушной среды определенной частоты заставляет камертон колебаться с той же частотой. К сожалению, Гибсон не стал прорабатывать аналогию дальше. Однако идея гибсоновского резонанса — это идея, не противоречащая келеровской концепции физического гештальта, но требующая при этом тщательной проработки.

#### 2.1.4.2 Прямое восприятие

Пожалуй, наиболее противоречивым аспектом гибсоновской теории является его гипотеза прямого восприятия, суть которой сводится к тому, что зрительное восприятие окружающего мира подвижным, активно ищущим организмом полностью определяется оптической информацией на сетчатке и что никаких процессов-посредников или внутренних отображений не требуется.<sup>1</sup> В изучении гибсоновской концепции прямого восприятия нужно учитывать то, что идея эта была во многом реакцией на первые робкие предположения о том, что восприятие возникает только на базе т.н. *бессознательных инференций* (умозаключений), радикально превышающих информацию, заложенную в стимул (см. раздел 2.1.4). Ученый категорически отрицал идею инференций, предлагая взамен тезис о том, что никаких умозаключений не требуется, если особь принимает во внимание множество источников оптической информации, которые не рассматриваются при традиционном анализе статических ретинальных изо-

<sup>1</sup> Степень сегодняшнего неприятия этой идеи легко оценить по работе Ульмана «Против доктрины прямого восприятия» (Ullman, 1980) и 27 комментариям к ней.





**Рис. 2.1.11** Слияние обстановки и точек наблюдения в динамический оптический массив. Оптический поток, сформированный этими изображениями задан как взаиморасположением поверхностей в сцене, так и траекторией движения наблюдателя в ней.

бражений, в частности тех, что идут из оптического потока, когда наблюдатель активно исследует окружающую обстановку.

К сожалению, великолепный гибсоновский анализ информации, извлекаемой зрением из динамического оптического массива, с трудом согласуется с фундаментальной неопределенностью восстановления картины реального мира из плоских ретинальных изображений — т.е. с *задачей инверсии*, которую мы обсуждали в разделе 1.2.3. Как видим, гибсоновский упор на особь исследующую — это попытка решить проблему инверсии добавкой к проксимальному стимулу временной размерности. Сие безусловно многое добавляет к исходной оптической информации, но при этом не является тем рычагом, что позволил бы оправдать идею прямого восприятия в контексте однозначного распознавания объектов внешнего мира из сугубо оптической информации.

Проблема состоит в том, что время — это тоже размерность внешнего мира. Следовательно, информация, дополнительная к динамическому ВОМ, по-прежнему недостаточна для однозначного восстановления событий обстановки. Последнее фактически четырехмерно (три пространственных размерности и одна временная), поскольку информация, зашита в динамический ВОМ лишь трехразмерна (две пространственных размерности и одна временная), причем даже при нормальном бинокулярном зрении. Дополнительная информация, обретаемая благодаря временной размерности (на которую мало кто из исследователей обращал внимание), безусловно помогает в решении задачи инверсии, но при этом не дает окончательного и однозначного решения (как полагал Гибсон).

### 2.1.5 Конструктивизм

Несомненно, что одним из важнейших фактов в физиологии является то, что несмотря на отрицание тезиса о самодостаточности оптической информации в решении задачи инверсии, наука настаивает на том, что нашей зрительной системе удастся подбирать верные решения с очевидным постоянством. За счет чего? Большинство исследователей сходятся в мысли, что помимо ретинальных изображений должен существовать некий дополнительный источник информации, активно используемый в зрительном акте: зрительная система должна чем-то дополнять данные, несомые проксимальными стимулами (даже динамическими), дабы выбрать единственное наиболее верное решение из бесконечного числа возможных. Сие позиция *конструктивистов*.

#### 2.1.5.1 Методологический бихевиоризм

Конструктивизм сегодня — это доминирующий, классический подход к теории зрения, объединяющий в себе все лучшее из трех ранее нами описанных концепций. Поскольку конструктивизм — это, если угодно, эклектическое направление, то его нельзя охарактеризовать четырьмя строгими положениями, которыми мы пользовались в разговоре о прочих психологических теориях. Современный конструктивизм нацелен в первую очередь на то, чтобы описать внутренние механизмы восприятия, нежели на то, что несет в себе внешний мир. Тем не менее, внутренние механизмы теорий конструктивизма часто основаны на извлечении информации о внешнем мире из паттернов ретинальной стимуляции, которые Гибсон описал в своей теории экологической оптики.

Конструктивизм проповедует ту идею, что глобальные, общие перцепты *строятся* (*конструируются*) из локальной информации. Вместе с тем признается и важность эмерджентных свойств линий, краев, углов и даже целых фигур, на чем настаивали гештальт-психологи.

В отношении спора нэйтивистов с эмпиристами современный конструктивизм держит нейтралитет: некоторые аспекты перцепционной обработки безусловно являются врожденными, однако прочие, вне всякого сомнения приобретаются через взаимодействие с окружающим миром. Каковы последние? — эмпирический вопрос, ответ на который дают наблюдения за развитием младенцев.

Наконец, самое современное направление конструктивизма — это т.н. *методологический бихевиоризм*, строящий выводы о перцепционных процессах на основе *количественных измерений* поведения человека и/или животных, но не на интроспекции осознанных перцептов (последнее может служить дополнительным источником научных данных). В этом конструктивизм резко контрастирует как со структурализмом,



**Рис. 2.1.12** Герман фон Гельмгольц — физик, математик, физиолог, чьи предположения о функции зрения легли в основу конструктивизма и современных вычислительных теорий (Golstein, 1989).

так и с гештальтизмом, склонными к экстенсивному применению феноменологических наблюдений. Однако можно не сомневаться, что интроспективный анализ — это первый шаг большинства конструктивистских теоретических построений. Отличие же состоит в том, что конструктивисты, дабы объективно поверять свои гипотезы, накапливают *количественные* поведенческие показатели.

### 2.1.5.2 Бессознательные инференции

Отцом теории конструктивизма считается выдающийся немецкий исследователь Герман фон Гельмгольц (см. рис. 2.1.12) — физик, математик и физиолог, современник Вундта и структуралистов. Очень важно отметить, что идеи ученого, опубликованные еще в 1867 г. в «Трактате по физиологической оптике», остаются на удивление современными.

Вклад Гельмгольца в понимание нами механизмов зрительного восприятия огромен, и одной из наиболее стойких и базисных концепций современной теории конструктивизма является тезис о том, что восприятие зависит от

процесса, именуемого *бессознательной инференцией*: в отличие от Гибсона Гельмгольц признавал существование некоей логической брешки между оптической информацией, поступающей от ретиальной стимуляции, и перцепционным знанием, полученным в результате анализа этой информации (логическая брешь — это как раз обсуждавшийся нами в разделе 1.2.3 процесс зрительной инверсии и его *недоопределенность*). Гельмгольц отстаивал ту идею, что брешь эту в процессе интерпретации ретиальных изображений в перцепционные «умозаклучения» — инференции, можно переступить путем неких скрытых «допущений», версий. То есть, ученый утверждал, что в зрительном акте действует процесс некоего, условно говоря, «умозаклучения» или что-то очень похожее, преобразующее неполную оптическую 2D-информацию проксимального стимула, в перцепционную 3D-интерпретацию внешнего мира. Сей процесс бессознателен, поскольку в отличие от обычных логических процессов, имеющих место в мышлении и при решении разного рода биологических задач, особь не имеет представления о том, как, когда и где возникнут зрительные инференции и возникнут ли они вообще.

Спустя век такие известные конструктивисты как Ричард Грегори (1970), Джулиан Хочберг (1964) и Ирвин Рок (1983) дополнили и развили идеи Гельмгольца.

### 2.1.5.3 Принцип правдоподобия

Коль скоро зрительная система делает выводы о характере предметов внешней обстановки из оптической информации, извлеченной из ретиальной стимуляции, то важно понимать то, как именно делаются эти выводы. Гельмгольц озвучил позицию, что зрение интерпретирует ретиальную стимуляцию по принципу наибольшей *вероятности* событий, вызвавших эту стимуляцию: такой подход обычно именуют «принципом правдоподобия». Сие суть вероятностная точка зрения, согласно которой зрительная система рассчитывает интерпретацию по принципу наивысшей вероятности данной ретиальной стимуляции.<sup>1</sup>

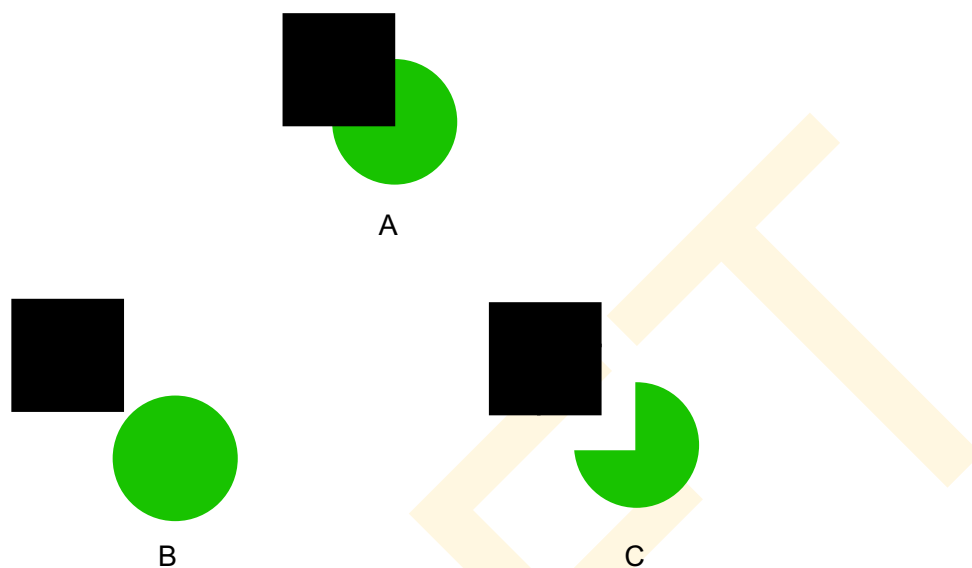
### 2.1.5.4 Принцип Прягнанца

С принципом правдоподобия контрастирует *гештальт-принцип минимализма* (*принцип Прягнанца*), утверждающий, что базисом выбора среди возможных интерпретаций является «добротность» или «простота» (Pomerantz & Kubovy, 1986). К примеру, тот факт, что мы воспринимаем рис. 2.1.13А как круг позади квадрата (интерпретация В), нежели квадрат, идеально подогнанный к фигуре в три четверти круга (интерпретация С), с позиции конструктивистской правдоподобности можно истолковать как то, что вероятность частичного прикрытия круга квадратом намного выше, нежели вероятность идеального совмещения трех четвертей круга со сторонами квадрата. Однако то же можно истолковать и с позиций прягнанцева принципа гештальт-психологии: полноценный круг в плане восприятия всегда проще, экономичнее, упорядоченней и симметричнее (см. раздел 8.3). На первый взгляд принцип правдоподобия и прягнанцев принцип принципиально отличны друг от друга, однако существует очень высокая корреляция между тем, что правдоподобно и тем, что просто. Характер объектов и событий, которые вероятнее всего создают паттерны стимуляции, также имеют тенденцию быть проще объектов и событий, свет от которых может произвести точно такое же оптическое изображение на сетчатке.

<sup>1</sup> Принцип правдоподобия, таким образом, указывает на т.н. *байесовский подход* в физиологии: в свете байесовской концепции зрительная система добивается максимально правдоподобной интерпретации, рассчитывая условную вероятность каждой из множества вероятных 3D-сцен ( $S_i$ ), давших те или иные сенсорные данные ( $E$ ) из других вероятностей, которые обычно проще получить. Теория байесовской вероятности получена из т.н. *теоремы Байеса*, суть которой можно выразить уравнением:

$$P(S_i|E) = \frac{P(E|S_i)P(S_i)}{P(E)}$$

где  $P(S_i|E)$  — это вероятность того, что данная сцена —  $S_i$  (одна из  $N$ -возможных сцен) является подлинной;  $P(E|S_i)$  — вероятность подлинности данной сцены;  $P(S_i)$  — т.н. *априорная вероятность* сцены;  $P(E)$  — априорная вероятность подлинности данной сцены. Если величины правой части уравнения известны или их можно вычислить из априорного опыта зрительного восприятия окружающего мира, то ключевая вероятность в левой части уравнения (в частности, 3D-ощущение, порожденное текущими сенсорными данными) также вычисляема. Основная трудность данного подхода в том, что он предполагает конечность возможного набора 3D-сцен и набора исходных сенсорных данных. Однако в случае наблюдателя исследующего ни первое, ни второе допущения не подтверждаются на практике.



**Рис. 2.1.13** Два варианта интерпретации одного и того же изображения: изображение «А» можно толковать как круг позади квадрата (В-интерпретация), либо как три четверти круга в стык с квадратом (С-интерпретация). То, что зрительная система человека отдает предпочтение В-варианту объяснимо как гельмгольцевым принципом правдоподобия, так и прыгнанцевым минимализмом.

#### 2.1.5.5 Эвристическая интерпретация

Общая теоретическая концепция, представленная в нашей книге хорошо согласуется с гельмгольцевским представлением о бессознательных инференциях. Мы углубим идею ученого и покажем, что путем высокоправдоподобных версий о характере предметов внешней обстановки и условиях их просмотра зрительная система выходит за рамки оптической информации. Когда эти версии соединяются с сенсорными данными, извлеченными из изображения, экспонирующего сетчатку, происходит процесс т.н. *эвристической интерпретации*, в котором зрительная система строит серию *версий* того, каковы наиболее вероятные условия внешней обстановки, свет от которых произвел данное изображение. Мы говорим, что процесс является эвристическим, поскольку задействует вероятностные эмпирические правила обычно (но все же не всегда) истинные. Если исходные версии ложны, они ведут к ложным инференциям, проявляющим себя в форме зрительных иллюзий (см. раздел 2.3.5). Однако в повседневной обстановке эти версии истинны и, соответственно, восприятие — веридикально.

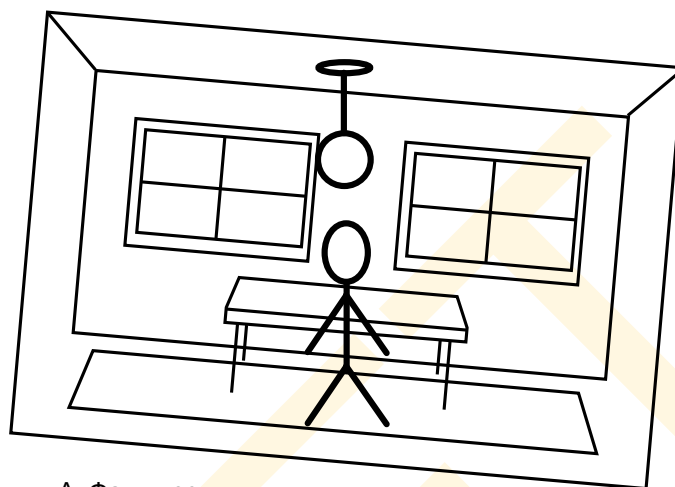
Принцип правдоподобия тесно соприкасается с конструктивистским тезисом об эвристичности восприятия: как только что было показано, оно окажется веридикальным тогда, когда версии, на которых это восприятие основано, истинны и иллюзорным — когда ложны. Иными словами, чем ближе версия к истине, тем вернее восприятие. Таким образом, эволюционная ценность зрения достигает максимума за счет наиболее вероятных версий интерпретации оптической информации. Поэтому в контексте разговора о зрительном восприятии мы будем говорить об эвристических версиях на протяжении всей книги. Эвристический подход позволит нам понять, как именно зрительная система выходит за рамки информации, данной для решения, казалось



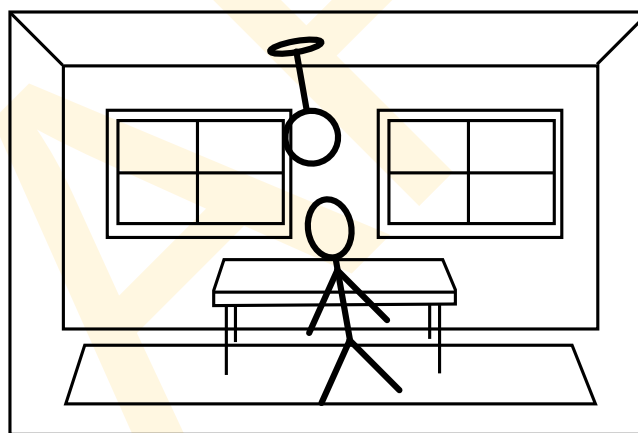
бы, неразрешимой задачи инверсии (и выполняя сие с высочайшей достоверностью и эффективностью).

Скрытых версий, строящихся зрительной системой, очень много и все они разные. Одним из примеров служит версия о том, что выраженные контуры комнатной обстановки (края потолка, дверей, стен и т.п.) расположены либо по строгой вертикали, либо перпендикулярно ей. Поскольку сия версия в большинстве случаев истинна, мы хорошо ориентируемся среди окружающих нас предметов. Но если она вдруг окажется ложной, как скажем в косой комнате луна-парка, наше зрительное восприятие станет поразительно беспомощным: в такой комнате нам не только покажется, что люстра висит неровно, но что и наше собственное тело наклонено в том же направлении (см. рис. 2.1.14). Столь иллюзорное восприятие обычно возникает в сугубо искусственных условиях и именно по той причине, что наши зрительные версии истинны в условиях естественных. Когда же эти версии оказываются ложными, как скажем в той самой косой комнате, иллюзорное восприятие наглядно демонстрирует нам факт существования обсуждаемых версий и свидетельствует о тесной взаимосвязи между реальностью, иллюзией и скрытыми допущениями (что будет обсуждаться детально в гл. 7).

Буквальная (примитивная) интерпретация бессознательных инференций может звучать так, что восприятие реализуется путем последовательного исполнения правил символической логики или путем решения математических уравнений. Действительно — это тот класс процессов, которые Гельмгольц имел в виду, когда говорил о бессознательных инференциях. Это те самые процессы, которые решительно отвергал Гибсон в своей доктрине прямого восприятия. К счастью, сегодня в рамках вычислитель-



А. Фактически



В. Результат восприятия

**Рис. 2.1.14** Иллюзия, возникшая в результате ошибочной скрытой версии: косая комната лунапарка (А) выглядит ровной, потому что зрительная система априори предполагает, что пол и потолок строго горизонтальны, а стены — вертикальны. В результате наблюдатель ошибочно воспринимает себя и люстру наклоненными (В).

ной концепции существуют достаточно веские доказательства верности общей схемы конструктивизма, чем мы и воспользуемся. Мы увидим, что, к примеру, в т.н. «нейронных» сетях результирующие перцепты частично основаны на данных сенсорной стимуляции, частично на дополнительных версиях, включенных в паттерн взаимосвязей между нейроноподобными элементами (см. разделы 5.3.3 и 6.4). Такие сети способны «делать выводы» на основе эвристических допущений (или, как их еще называют коннекционисты — «гибких ограничений») и без использования символической логики или математических уравнений.

В нашей книге мы будем толковать концепцию «бессознательных инференций» достаточно широко, в том числе и в аспекте вычислительного формализма. Однако к чести гештальтистов и Гибсона отметим, что поведение таких сетей хорошо согласуется с динамическим поведением кёлеровских физических гештальтов и гибсоновской метафорой механического резонанса при информационном захвате.

## 2.2 КРАТКАЯ ИСТОРИЯ ИНФОРМАТИКИ В КОНТЕКСТЕ НАУКИ О ЗРЕНИИ

С появлением в 50-60-х годах XX века информатики и кибернетики, в физиологии наступила новая эра: стали применяться компьютерные имитации, моделирующие когнитивные процессы разного толка; физиологи приняли логику информационного процессинга, т.е. стали рассматривать мозг как биологический процессор информации. Все это оказало огромное влияние на эволюцию науки о зрении как на междисциплинарную область знания.

### 2.2.1 Компьютерное зрение

Колоссальным прорывом в изучении зрения явилась возможность компьютерной имитации сложных процессов зрительной обработки. Зрение и прочие виды восприятия и мышления рассматривались ранее исключительно в контексте живого: исследователи строили гипотезы и тестировали их на живых существах, что было дорого, сложно и далеко не всегда гуманно. Однако появление технологий компьютерной имитации радикально изменило ситуацию: вычислительные машины позволили исследователям создавать синтетические системы, которые вначале программируются человеком, а затем их поведение сравнивается с таковым у зрячих организмов. Расстроим краткую историю вопроса.

#### 2.2.1.1 Появление компьютеров

Компьютерная эра началась в 1930 г., когда выдающийся английский математик Алан Тьюринг (см. рис. 2.2.1) определил класс гипотетических ЭВМ, получивших название «универсальных тьюринговых машин» (Turing, 1937). Машины эти программируются на автоматическую обработку информации теоретически бесконечным числом способов. Отметим, что тьюринговы компьютеры — суть математическая абстракция, однако же это не помешало инженерам быстро приступить к физическому воплощению идеи ученого. В 1946 г. математик Джон фон Нейман собрал команду, которая на базе Пенсильванского университета изготовила первый цифровой компьютер общего назначения, явивший собой монструозную вакуумную трубку огромных размеров. Аппарат окрестили «ENIAC» — Electronic Numerical Integrator and Calculator — элек-

тронный цифровой интегратор и калькулятор. Несмотря на то, что современные компьютеры на многие порядки превосходят по мощности и сложности организации первые вычислительные машины, все они в конечном счете суть образцы универсальных тьюринговых машин, выполняющих определенное число операций, заложенных в них программистами.

В начале 40-х годов XX века Тьюринг сам осознал грандиозные перспективы своего изобретения, в частности, в плане имитации мыслительных процессов (Turing, 1950). Идея эта дала толчок появлению такого направления как *искусственный интеллект (ИИ)* — ветвь компьютерной науки, в которой вычислительная машина программируется на имитацию разумного поведения. Вначале исследователи сосредоточились на попытках имитации сложных интеллектуальных задач, таких как игра в шахматы и доказательство математических теорем (к примеру Newell, Shaw & Simon, 1958; Newell & Simon, 1963), и лишь потом пришло понимание того, что программирование компьютеров на визуальное восприятие окружающего мира — это куда более обещающее и перспективное направление исследования. Сие дало толчок развитию научного направления, которое сегодня именуют *компьютерным зрением*, т.е. изучению того, как следует программировать ЭВМ на извлечение полезной информации об окружающем мире из оптических изображений.

Компьютерное зрение инициировало два фундаментально важных шага в нашей науке, навсегда и радикально изменивших ее теоретическую составляющую: началась работа с реальными изображениями и появились т.н. *эксплицитные теории*.

### 2.2.1.2 Начало работы с реальными изображениями

Теории зрения, воплощенные в первых компьютерных программах, позволили применять эти программы к черно-белым изображениям реальных сцен, полученным с помощью видеокамер. Заметим, что классические теории зрительного восприятия в целом строились на учете параметров стимула, никогда не встречающегося в реальной обстановке — такой стимул предполагался идеальным, лишенным шумов и создавался светом, отраженным от идеальных объектов. Компьютерное зрение позволило опробовать имевшиеся теории на *реальных изображениях реальных сцен*.



Рис. 2.2.1 Алан Тьюринг — выдающийся британский математик, которому принадлежит изобретение компьютера, а также авторство идей искусственного интеллекта и тестирования методом «черного ящика» (т.н. функциональный тест), именуемый ныне «тьюринговым тестом» (Turing, 1959).

### 2.2.1.3 Появление эксплицитных теорий

До наступления эры компьютерных имитаций теории зрительного восприятия были туманны, произвольны и неполны, перегружены лишними концепциями и деталями. Сие в аккурат случай четырех классических теорий зрения, которые мы обсуждали в разделе 2.1. Электронные имитации радикально изменили ситуацию, поскольку одним из характерных признаков компьютерных программ является то, что они *явны* или как еще говорят — *эксплицитны*.

Первым открытием, сделанным на основе описанных шагов, явилось понимание того, что *зрение запредельно сложно*. Оказалось также, что научить компьютер «видеть» даже простейшие вещи — невероятно трудно: процессы, которые до того считались физиологами сами собой разумеющимися (к примеру, детекция краев и определение того, к какой части трехмерного целого относится данный участок и т.п.), требовали колоссальных вычислительных ресурсов. Оказалось, что даже самая современная компьютерная программа, исполняемая самым мощным компьютером, не в состоянии достичь скорости, точности и гибкости человеческих органов чувств и решать простейшие зрительные задачи. В то же время эти же задачи зрительная система отрабатывает с огромной скоростью, точностью и без видимых усилий.

### 2.2.1.4 Мир из кубиков

В первых теориях компьютерного зрения сцена толковалась как фрагмент примитивного мира, в котором все объекты просты, равномерно окрашены, геометрически правильны и лежат на общей плоской поверхности. С целью упрощения задачи число и виды объектов сократили до минимума, что позволило компьютеру хоть как-то работать. Но если в сцену попадала чашка кофе или даже кусочек бумаги, компьютерная программа не справлялась и «висла».

Одной из первых сколько-нибудь значимых блочных зрительных имитаций явилась программа Робертса (Roberts, 1965), которая, с одной стороны, явилась отправной точкой исследований в сфере компьютерного зрения, а с другой — наглядно продемонстрировала то, что первые программы такого плана могли работать лишь в сугубо искусственных, упрощенных условиях. Компьютер, оснащенный такой программой мог распознавать черно-белые изображения, но только при условии, что изображения эти состоят из ненакладывающихся многогранников и что набор этих многогранников более чем скуден (см. рис. 2.2.2А).

Функционировала робертсовская программа двухэтапно:

- на первом этапе по видеоизображению путем поиска т.н. *локальных яркостных краев* — изменений в количестве света, отраженного от соседних элементов изображения вычерчивались линейчатые контуры (см. рис. 2.2.2Е). Края регистрировались только в четырех пространственных ориентациях с шагом в 45° (см. рис. 2.2.2С), а затем соединялись вместе, образуя плавные линии контуров объекта (рис. 2.2.2D);

- на втором этапе программа подгоняла к имеющемуся чертежу определенный набор объемных базовых элементов (см. рис. 2.2.2F).

### 2.2.1.5 Вычислительные подходы к экологической оптике

Недавние успехи в сфере компьютерного зрения позволили раздвинуть рамки условий, при которых становится возможным формальный анализ информации, поступающей из оптических изображений сцен на поверхности светочувствительных сенсоров. Напомним, что сей подход в свое время горячо отстаивал Гибсон в своих первых

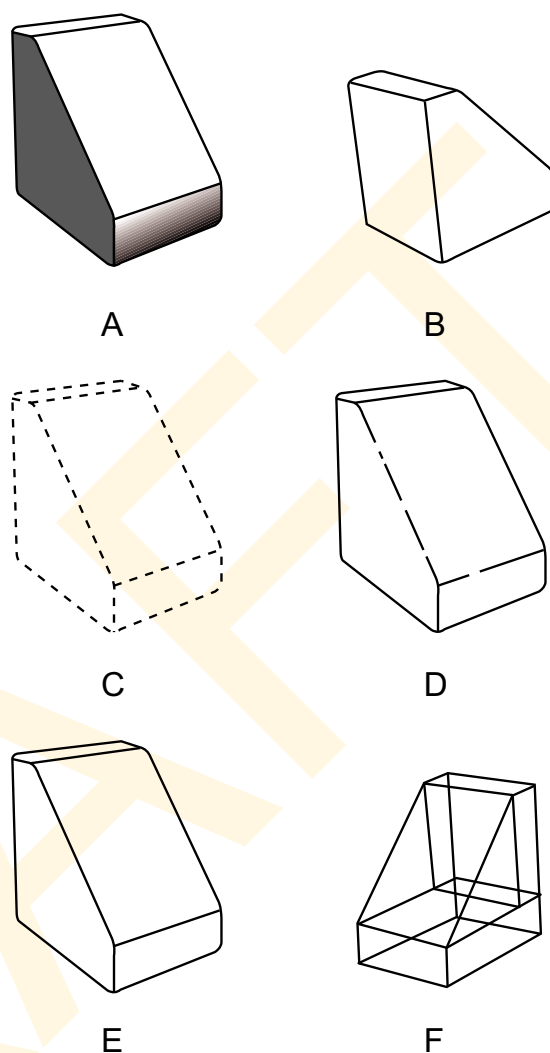
трудах по экологической оптике (Gibson, 1950, 1966, 1979). Итогом явилось то, что исследователи стали тратить уйму времени, пытаясь угадать пути восстановления более полной информации о сцене непосредственно из ее первичного двумерного изображения на сенсоре, в частности информации об удалении и отклонении поверхностей от точки наблюдения.

По многим аспектам исследования пионерами в данной области стали голландские психофизиологи Жан Кёндеринк и Андреа Ван Дорн, применившие сложные методы дифференциальной геометрии к решению таких вопросов как:

— восприятие движения из оптического потока (Koenderink & Van Doorn, 1976a), восприятие расстояния до объекта из стереоскопических данных (Koenderink & Van Doorn, 1976b)

— 3D-ориентации поверхностей из информации о затенениях (Koenderink, Van Doorn, & Kappers, 1992) и т.п. Отметим, что сами ученые не писали компьютерных программ, но их исследования положены в основу работы программистов.

Математический подход к созданию программ компьютерного зрения яснее и эффективнее всего сформулирован в стенах Массачусетского технологического института (МТИ)<sup>1</sup>, в частности Дэвидом Марром (см. рис. 2.2.3) и его талантливыми коллегами. Исследование Марра давало математический анализ пути, по которому информация о структуре поверхностей и объектов трехмерного пространства извлекается из структуры яркостных модуляций двумерной проекции на сетчатку (Marr, 1982), что, кстати говоря, весьма созвучно с гибсоновским экологическим принципом. Успех исследователей был огромен и их



**Рис. 2.2.2** Роберстовская программа «блочного мира». Чтобы получить изображение А, вначале выполнялась обработка локальных краев (С), которые затем объединялись в замкнутые контуры (D и E). Затем к полученной прориси подбирался подходящий ей элемент (примитив), дабы получить правильную форму предмета (F) и предсказать его возможный вид с другой точки наблюдения (B) (Roberts, 1965).

<sup>1</sup> МТИ (M.T.I.) — Массачусетский технологический институт. Ведущий частный научно-исследовательский и учебный институт США в г. Кембридже — пригороде Бостона, шт. Массачусетс. Основан в 1861 в г. Бостоне, переведен в Кембридж в 1916. Известен исследованиями в области физики, химии, биологии, технических наук.





**Рис. 2.2.3** Дэвид Марр: выдающийся физиолог, чьи вычислительные теории зрения и размышления об их взаимосвязи с физиологией зрительного восприятия помогли утвердить междисциплинарный подход в физиологии.

подход к проблеме стал ведущим в сфере изучения компьютерного зрения в последние два десятилетия. На страницах нашей книги мы детально разберем результаты исследований, проведенных в МТИ.

### 2.2.1.6 Коннекционизм и нейронные сети

Сравнительно новой разработкой в сфере компьютерной имитации зрения стало появление моделей т.н. *нейронных сетей* (Feldman, 1981; Feldman & Ballard, 1982; Grossberg, 1982; Hinton & Anderson, 1981).

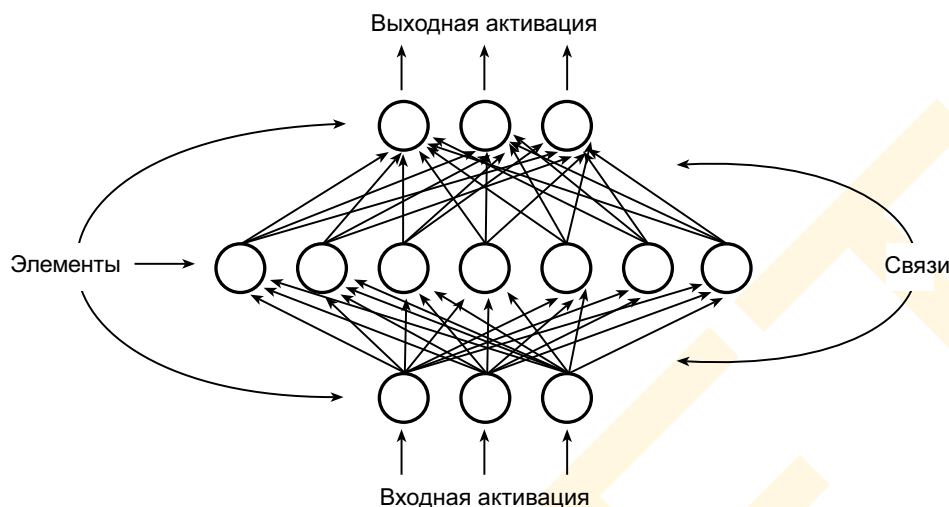
Модели эти строятся на том предположении, что человеческое зрение сильно зависит от деятельности т.н. *нейронных цепей с массовым параллелизмом*. Из самого названия следует, что такие модели представляют собой сложные сети, состоящие из множества вычислительных элементов (*единиц*) с богатыми взаимосвязями. Каждый элемент работает как отдельный «нейрон». Текущее состояние элемента характеризуется *уровнем* его *активности*, который примерно соответствует частоте пульсации нейронов. Сигнал распространяется по всей сети за счет связей, которые (являясь либо возбужда-

ющими, либо подавляющими) сильно напоминают т.н. *синапсы* — межклеточные контакты, которыми связаны между собой нейроны (см. рис. 2.2.4).

Несмотря на то, что модели нейронных сетей можно описать математически, такое описание строится в основном на нелинейных уравнениях, трудно разрешимых аналитически (см. Приложение В). То есть, чтобы определить, как поведет себя данная сеть, ее работу следует проимитировать средствами ЭВМ.<sup>1</sup>

**Перцептроны.** Историческими предшественниками современной теории коннекционизма явились т.н. *перцептроны*. Перцептроны — это некий специфический класс моделей нейроноподобных сетей, активно разрабатывавшихся Фрэнком Розенблаттом (1962) и его коллегами. В структуре перцептронов особенно интересным и важным являлось то, что они могли научиться идентификации объектов новых категорий путем управления весами их связей согласно неким простым правилам. Розенблатт до-

<sup>1</sup> Несмотря на то, что работа большинства современных компьютеров последовательна, а модели нейронных сетей обладают высочайшим параллелизмом, первое никак не влияет на результат имитации, но лишь снижает ее скорость. Параллельные процессы также можно проимитировать последовательными вычислениями путем т.н. *гирляндования* или как еще говорят — *стрингинга* тех процессов, которые проходят одновременно.



**Рис. 2.2.4** Трехслойная прямопоточная нейронная сеть с обучением по алгоритму обратного распространения (ошибок). Первый слой элементов представляет вход, последний — выход, а промежуточный «скрытый» слой выполняет преобразование входных данных в выходные. Обучение реализуется за счет управления весами связей, направленного на минимизацию ошибок.

казал, что этих правил вполне достаточно для того, чтобы научить сеть любому категориальному различению (если таковое в принципе выполнимо).

Примерно в то же время психолог Дональд О. Хебб, опираясь на идеологию нейронных сетей, в рамках которых ассоциации образовывались внутри и между сложными «клеточными ассамблеями», работающими одновременно (Hebb, 1964), провел огромное этологическое<sup>1</sup> и психологическое исследование обучаемости человека и животных.

Однако, несмотря на очевидный прогресс, исследования в сфере нейронных сетей подверглись тяжелому удару, когда математики Марвин Мински и Сеймур Паперт (1969) продемонстрировали ряд серьезных недочетов в розенблаттовских перцептронах, что остановило исследования в данной области более чем на десятилетие.

Тем не менее отдельные ученые-энтузиасты, раскиданные по всему миру, сохранили верность идее моделирования нейронных сетей и в начале 80-х годов XX века группа когнитологов, трудившихся на кафедре психологии Калифорнийского университета в Сан-Диего, получила серию важнейших данных, реанимировавших это направление исследования. Назвав свой подход *моделированием параллельной обработки*, Дэвид Румельхарт, Джеймс Мак-Клелланд, Джеффри Хинтон, Павел Смоленский и Майкл Джордан продемонстрировали то, что модели нейронных сетей обладают рядом серьезных преимуществ перед традиционными компьютерными программами (McClelland & Rumelhart, 1986, 1988; Rumelhart & McClelland, 1986). Позднее мы обсудим некоторые примеры таких моделей.

<sup>1</sup> Этология — наука о поведении. — Прим. пер.

### 2.2.2 Информационный подход в психологии

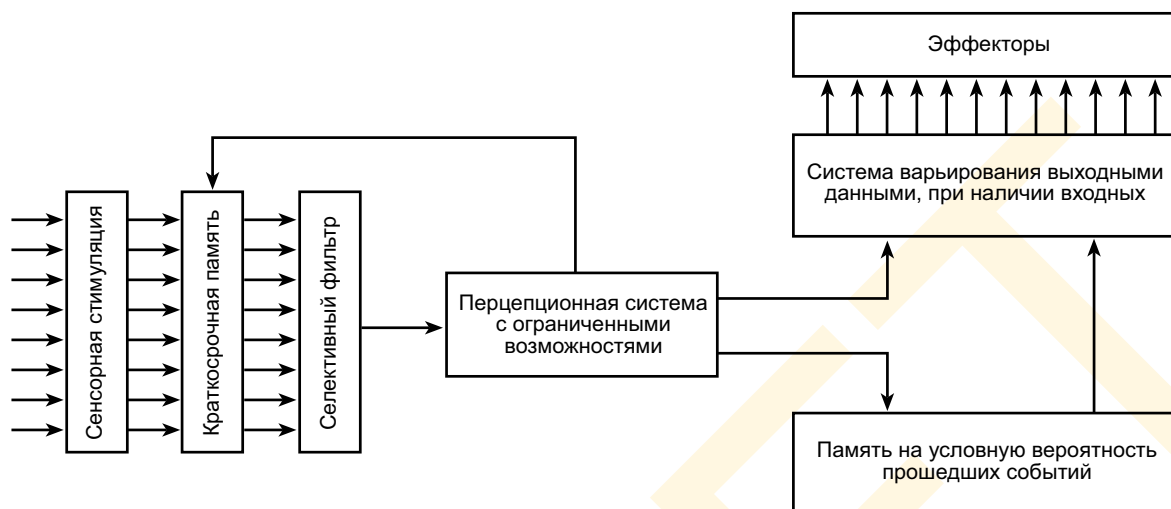
Еще одной вехой в эволюции современной физиологии явилось появление информационного подхода в психологии. Начиная с 20-х годов XX-века умы психологов захватил *бихевиоризм*.<sup>1</sup> Бихевиористы полагали, что путеводным принципом в психологии является непосредственное наблюдение за поведением особи. Методологически сие есть диаметрально противоположность интроспекции (аналогично тому как гештальтизм противоположность структурализму). Теоретически это означало полное отрицание такого психологического направления как *ментализм*. В своем крайнем выражении бихевиоризм требовал полного устранения из психологических концепций любых ссылок на восприятие, поскольку восприятие в основе своей — плод внутреннего познания внешнего мира. Возможно именно поэтому теоретический бихевиоризм не нашел поддержки в физиологическом сообществе.

Расцвет бихевиоризма пришелся на 40-е и 50-е годы XX века, особенно в контексте вопросов обучения. Однако в конце 50-х появилась концепция, поставившая под сомнение догмы бихевиористов и утверждавшая ту позицию, что ментальные процессы можно толковать как процессы информационные. Толчком к такому пониманию послужили успехи в электронике, вычислительной технике и теории информации. Сие позволило психологам строить теории зрения в рамках нового, более точного понятийного языка, который теснее соприкасался с языком компьютерного программирования. Для проверки информационной теории были разработаны новые экспериментальные методы, в частности использующие показатель времени отклика как объективный критерий добротности этой теории (Sternberg, 1966, 1969). За десять лет информационный подход крепко утвердился в когнитивной психологии и вскоре стал доминирующей парадигмой в понимании механизмов зрительного восприятия и прочих проявлений ментальной активности.

Одной из эпохальных публикаций, ознаменовавшей приход новой эры в когнитологии, явилась монография Дональда Бродбента «Восприятие и коммуникация», вышедшая в 1958 г. В этой книге автор анализирует процесс отбора и обработки информации, заключенной в звуковых стимулах (т.е. стимулах, воздействующих на наш орган слуха) от информации из прочих стимулов. Бродбент показал, что внимание работает как некий переключатель, свитчер, выбирающий из нескольких информационных каналов нужный, причем *после того* как выполнен анализ общей сенсорной стимуляции (см. рис. 2.2.5). Несмотря на то, что отдельные концепции Бродбента впоследствии заменили более сложными информационными теориями (в контексте восприятия и внимания), ученый был одним из первых, кто предложил психологическую концепцию в форме функциональной схемы, описывающей *темпоральную* (временную) структуру событий информационной обработки (см. раздел 2.3).

Знаменательной вехой в истории информационной теории восприятия явилось исследование краткосрочной зрительной памяти, выполненное психологом Джорджем Сперлингом, трудившимся под эгидой Bell Telephone Laboratories. Разработав хитрую методику, позволявшую не перегружать зрительную систему наблюдателя в ее реакции на сверхкороткие зрительные стимулы (меньшие 0.1 сек), Сперлинг изучил ту форму зрительной памяти, которую ранее именовали *иконической памятью*, действующей не более полусекунды. Весьма элегантные эксперименты ученого выявили множество удивительных свойств данного вида памяти и породили сотни дочерних ис-

<sup>1</sup> От англ. behavior — поведение. — Прим. пер.



**Рис. 2.2.5** Фильтровая бродбендовская теория слухового восприятия. Предложенная в 1958 г., явилась одной из первых теорий восприятия, построенных в информационной логике и логике блок-схем. Селективный фильтр работает как свитч, пускающий информацию по одному из сенсорных каналов на вышестоящий уровень т.н. *перцепционной системы с ограниченными возможностями* (Broadbent, 1958).

следований ее механизмов (см. раздел 12.1). Однако главное то, что исследования Сперлинга побудили многих психологов рассматривать вопросы зрительного восприятия с позиций информационной обработки.

В итоге, исследования Сперлинга и последующие концептуальные разработки в психологии привели к отказу от бихевиоризма как от доминирующего подхода к пониманию механизмов познания. Водоразделом явился в 1967 г. выход в свет книги Ульриха Найссера «Когнитивная психология». Бихевиористы отвергли найссеровский информационный подход сразу и по всем направлениям: от общих вопросов познания, до самых специфических его областей, в частности восприятия и внимания. Но как стало ясно в дальнейшем, у найссеровской концепции множество плюсов в плане понимания нами механизмов познания, включая зрение (что будет показано в следующих главах).

### 2.2.3 Информационный подход в биологии

Третьим научным событием, утвердившим информационную парадигму, явилось изобретение физиологических методов изучения нейронной активности в зрительной системе. Эти методы позволили ответить на вопрос о том, *как* происходит обработка зрительной информации в сетчатке и зрительных центрах головного мозга. При этом нейрон принят за единицу анализа в зрительной системе, то есть зрение полагают объяснимым в логике некоего *пускового паттерна*, появляющегося в результате взаимодействия между множеством отдельных нейронов (Barlow, 1972). Благодаря тому, что существуют методы изучения активности отдельных нейронов, исследователи могут клетка за клеткой строить функциональные схемы связей в зрительной системе, описывая функцию каждого нейрона в отдельности.

Появлением методов исследования нейронной активности отмечено начало инфор-

мационного подхода в биологии — области, в которой зрение играет ведущую роль. Последние разработки в технологии визуализации структуры и функции живого мозга радикально расширили возможности исследования физиологии зрительного восприятия, в частности у здоровых наблюдателей.

### 2.2.3.1 Первые опыты

До начала 50-х годов XX-века мозг не рассматривался как система обработки информации, но просто как орган, чья физиология непонятна. Действительно, были потрачены годы упорных исследований и отшумели сотни жарких научных дебатов, прежде чем биологи поняли, что нейроны не соединены непосредственно друг с другом, но посредством отдельных структур — синапсов, реализующих передачу сигналов за счет химических агентов в тонких межклеточных щелях. Но даже тогда идея мозга как процессора информации не была принята научной общественностью и не принималась до тех пор, пока не провели аналогию между мозгом и компьютером. Первым к такой аналогии пришел математик Джон фон Нейман (1951), и, в частности к тому, что нервные импульсы, идущие по аксонам — это в сущности то же самое, что и двоичный цифровой код вычислительных машин (1 и 0). Сегодня однако некоторые нейрофизиологи уже не рассматривают всерьез эту аналогию, поскольку налицо явные отличия между мозгом и компьютером. Но при этом самая идея, что **мозг — это биологическая система обработки информации** полностью принята, и результаты большинства современных физиологических исследований функции мозга интерпретируются только в информационном ключе.

**Деструктивные эксперименты.** Первые исследования функции центральной нервной системы строились на базе т.н. *деструктивных экспериментов* — экспериментов, в которых та или иная область головного мозга животных удалялась хирургическим путем или каким-либо иным способом. Плодом этих экспериментов стал ряд удивительных открытий касаясь локализации церебральных функций (см. раздел 1.3).

**Мозговая электростимуляция.** То же можно сказать и о методах т.н. *мозговой электростимуляции*, суть которых в том, что в мозг вводился специальный микроэлектрод, слабый ток от которого воздействовал на окружающие нейроны; экспериментатор при этом следил за изменением в поведении животного или, в случаях нейрохирургического лечения людей, записывал словесное описание ощущений больного. Оба варианта экспериментов оказались исключительно важными в оценке масштабности структуры зрительных центров головного мозга, но при этом не дали никаких полезных сведений о характере информационной обработки в нормально функционирующем мозге, поскольку таковая не измерима электрохимическим путем.

### 2.2.3.2 Экстраклеточная регистрация

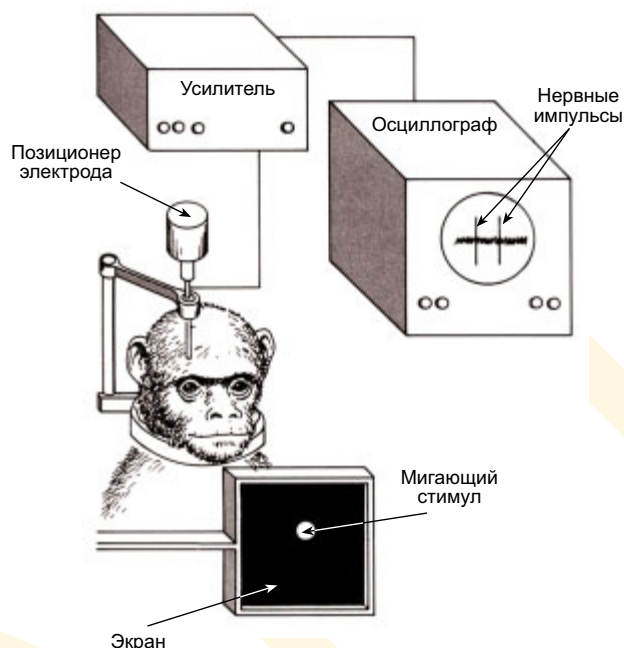
В 50-х годах XX века с целью изучения характера информационной обработки в одном отдельно взятом нейроне, и с целью выявления зависимости между подачей стимула и усилением/ослаблением нейронной активности, была разработана физиологическая методика, получившая название *экстраклеточной регистрации активности единичного нейрона*: сверхтонкий электрод подводили предельно близко к аксону, благодаря чему удавалось зарегистрировать ничтожные изменения электрического потенциала, возникавшие в момент всплеска активности этого нейрона (см. рис. 2.2.6). Выходной сигнал электрода записывался и соотносился с характером стимула, акти-



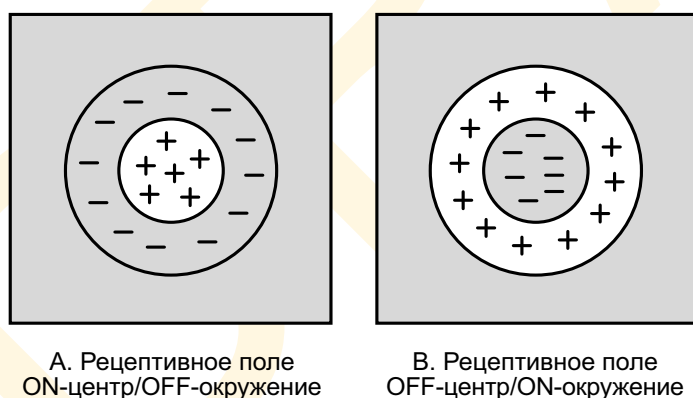
вировавшего нейрон. В контексте зрения в роли стимула обычно выступал световой паттерн, проектируемый на сетчатку подопытного животного.

В 50-60-е годы XX века были проведены беспрецедентные по значимости исследования, вызвавшие всплеск интереса к биологической обработке информации в зрительной системе млекопитающих, в частности котов и обезьян: наиболее значимыми явились работы Стивена Кюфлера, Дэвида Хьюбела и Торстена Визеля, проведенные в Гарвардском университете. Кюфлер (1953) применил метод экстраклеточной регистрации активности единичного нейрона для определения размеров и формы т.н. *рецептивных полей* ганглиозных клеток сетчатки.

⇒ **Рецептивное поле ганглиозной клетки сетчатки** — это некая пространственная область сетчатки, световое воздействие на которую либо ак-



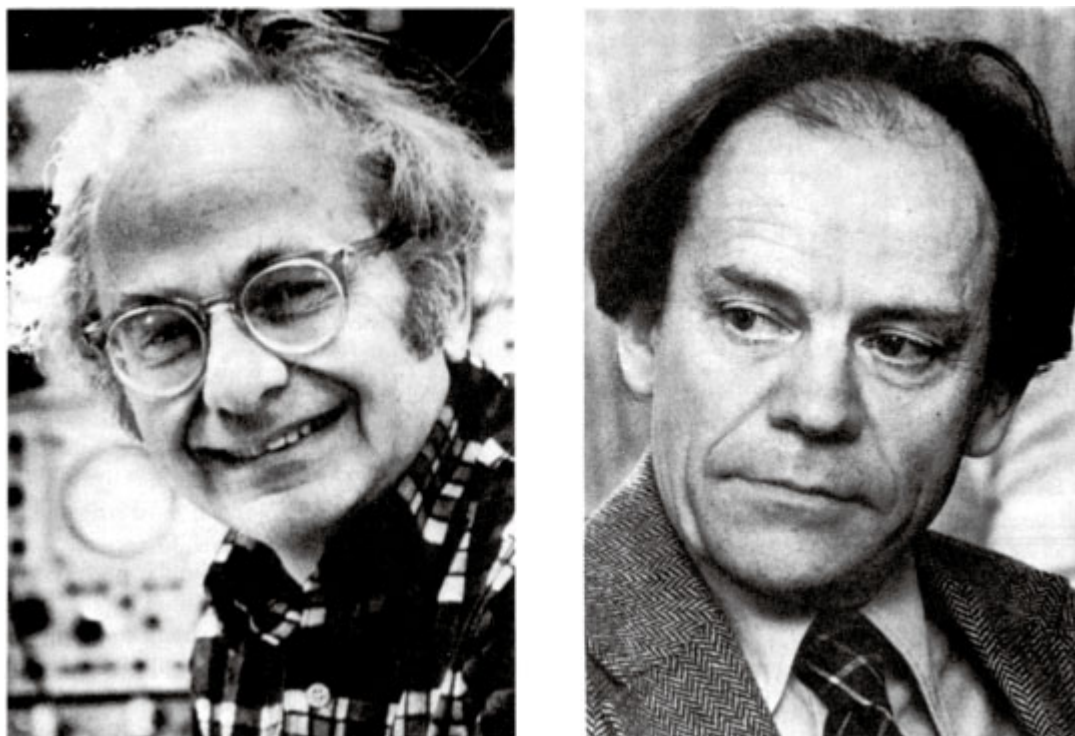
**Рис. 2.2.6** Экстраклеточная регистрация активности единичного нейрона: микроэлектрод внедрен в глаз животного, а на сетчатку воздействует стимулирующий световой паттерн; электрод регистрирует электрическую активность клетки, результат выводится на экран осциллографа (Sekuler & Blake, 1985).



**Рис. 2.2.7** Рецептивные поля ганглиозных клеток сетчатки. Частота пульсации клеток класса ON-центр/OFF-окружение (слева) растет, когда свет стимулирует центральную область рецептивного поля; падает, когда свет стимулирует его периферию; не меняется, когда свет оказывает воздействие за пределы рецептивного поля. Частота пульсации ганглиозных клеток класса OFF-центр/ON-окружение (справа) меняется противоположным образом.

тивирует, либо подавляет одну ганглиозную клетку.

Следует отметить, что аналогичные исследования (и в тот же период) проводились неким Барлоу (1953) на сетчатке лягушек. Исследования показали, что оптимальный стимулирующий паттерн, вызывающий высокую активность ганглиозных клеток — это либо яркая точка с темным круглым окружением, либо наоборот — темная точка с ярким окружением (рис. 2.2.7). Рассуждения строились в логике нейронных межсоединений, согласуемых с наблюдаемыми результатами, и описы-



**Рис. 2.2.8** Дэвид Хьюбел и Торстен Визель — нобелевские лауреаты — пионеры в исследовании коркового отдела зрительной системы и выявлении множества свойств рецептивных полей клеток первичной зрительной коры. Труд ученых дал толчок множеству исследований физиологии зрительной коры с применением техники экстраклеточной регистрации (Golstein, 1989).

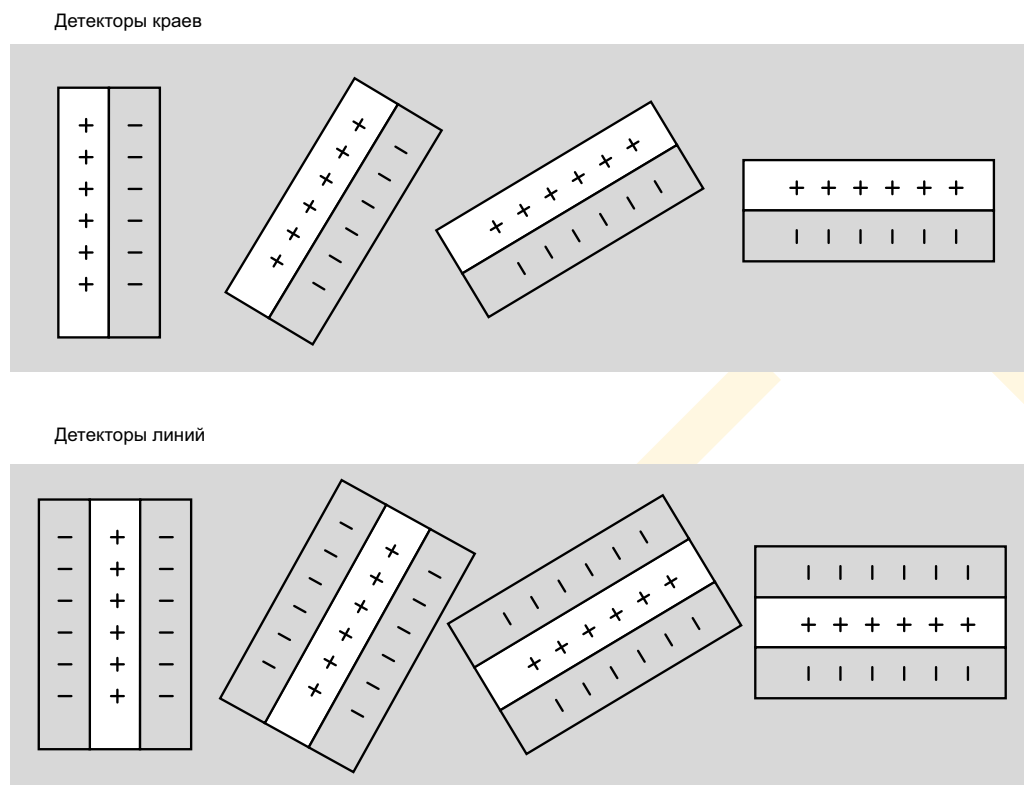
вали то, как информация передается возбужденными ганглиозными клетками и каким образом она обрабатывается.

Нобелевские лауреаты Дэвид Хьюбел и Торстен Визель (рис. 2.2.8) применили сходные методы при составлении карт более сложных рецептивных полей клеток зрительной коры (Hubel & Wiesel, 1959, 1962). Ученые обнаружили вытянутые рецептивные поля, несущие информацию о пространственной ориентации стимулов (см. рис. 2.2.9). Исследования Хьюбела и Визеля дали научное объяснение механизмов обработки зрительной информации и стоит отметить, что благодаря этому объяснению именно зрительная кора — это наиболее изученный участок коры головного мозга человека. В дальнейшем мы будем подробно говорить о свойствах коры, в частности когда речь пойдет о нейронной обработке пространственной информации (гл. 4).

Стоит отметить однако, что у метода экстраклеточной записи сигналов единичных нейронов имеются существенные недостатки: чтобы понять какова одновременная работа серии клеток требуется множество электродов, что чисто технически затруднительно. Проще говоря, в деле изучения общей архитектуры коры метод экстраклеточной регистрации неуклюж и непрактичен.

### *2.2.3.3 Авторадиография*

К счастью, удалось разработать серию новых и гораздо более эффективных, чем предыдущий, физиологических методов исследования, позволяющих изучать функ-



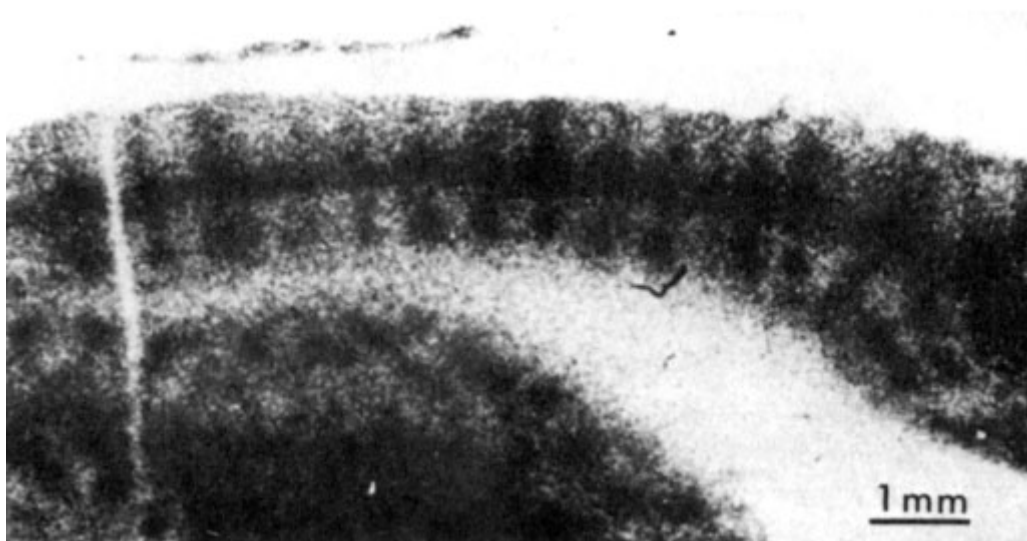
**Рис. 2.2.9** Рецептивные поля кортикальных клеток. У клеток области V1 зрительной коры вытянутые рецептивные поля, дающие максимум отклика, когда стимулируются краями или линиями определенной ориентации и позиции.

циональную архитектуру коры головного мозга. Один из них — метод *авторадиографии*.

Чтобы построить авторадиограмму, животному вводится определенное количество радиомаркера — радиоактивно меченного вещества, обычно из класса сахаров (например, 2-дезоксиглюкоза). Такое вещество захватывается возбужденными нейронами и накапливается в них, поскольку не участвует в обмене веществ. Спустя некоторое время после инъекции животному предъявляют стимул, одно из свойств которого представляет интерес, например, набор линий определенного направления. Радиомаркер тем быстрее сорбируется клетками, чем выше их активность в ответ на стимул с данными свойствами. Животное затем, увы, умертвляется, а с его мозга выполняются тонкие срезы, накладываемые на радиочувствительную «фото»-бумагу. Результирующее изображение дает картину пространственного распределения тысяч клеток, активированных стимулом (рис. 2.2.10). Стоит отметить, что в связке с методом экстраклеточной регистрации авторадиографический метод колоссально углубил наше понимание функциональной архитектуры зрительной коры.

### 2.2.3.4 Визуализация мозговых структур

Наиболее эффективными и гуманными инструментами изучения нейронных механизмов зрения являются т.н. *неинвазивные* методы исследования, предполагающие получение функциональных изображений головного мозга человека: *компьютерная*



**Рис. 2.2.10** Авторадиограмма положения нейронных колонок в зрительной коре. Кортикальные клетки прокрашиваются сильнее, если оказываются высокоактивными. Темные вертикальные полосы указывают на ориентацию колонок в верхних и нижних слоях коры. Темная горизонтальная полоса — это слой 4, получающий входной сигнал от латерального колленчатого тела и откликающийся на стимулы любых ориентаций (Hubel, Wiesel, & Stryker, 1977).

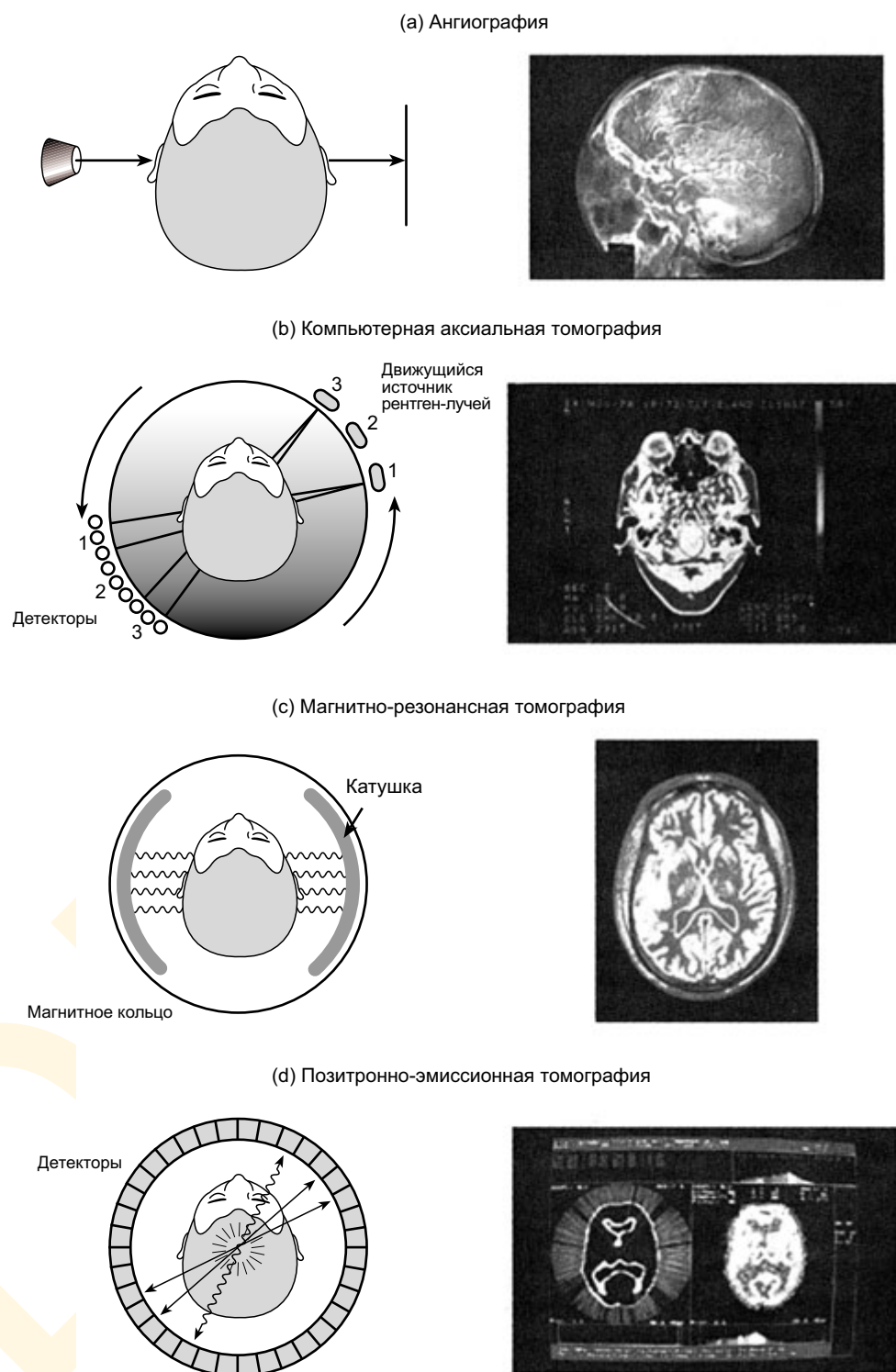
томография (КТ), позитронная эмиссионная томография (ПЭТ), магнитно-резонансная томография (МРТ). Эти великолепные технологии, разработанные в 70-е годы XX-века, явились буквально революционными в исследовании биологических тканей, в частности головного мозга, поскольку позволили получать информацию о состоянии органов без нарушения целостности кожных покровов.

**Рентген-компьютерная томография.** Первой появилась техника *рентген-компьютерной томографии*: чтобы получить томограмму головного мозга пучок рентгеновских лучей фокусировался на интересующем слое мозговой ткани (очень тонком), а прошедшие сквозь нее лучи регистрировались на противоположной стороне (см. рис. 2.2.11В). Поскольку ткани мозга разной плотности по-разному влияют на ход рентгеновских лучей, мощность излучения, проникшего сквозь ту ли иную ткань, несет информацию о ее физической плотности. Следовательно, замер интенсивности пучка, прошедшего в данном конкретном направлении, является достаточным для определения плотности в данной локале мозга.

Показано, что измерение интенсивности рентгеновских лучей по разным направлениям, позволяет составить полную карту плотности тканей на интересующем участке мозга: описанная процедура повторяется под множеством разных углов, а массив полученных данных обрабатывается мощным компьютером с использованием сложных алгоритмов. Из множества срезов, полученных на разных уровнях, можно составить более или менее полную картину анатомической структуры головного мозга (см. рис. 2.2.12).

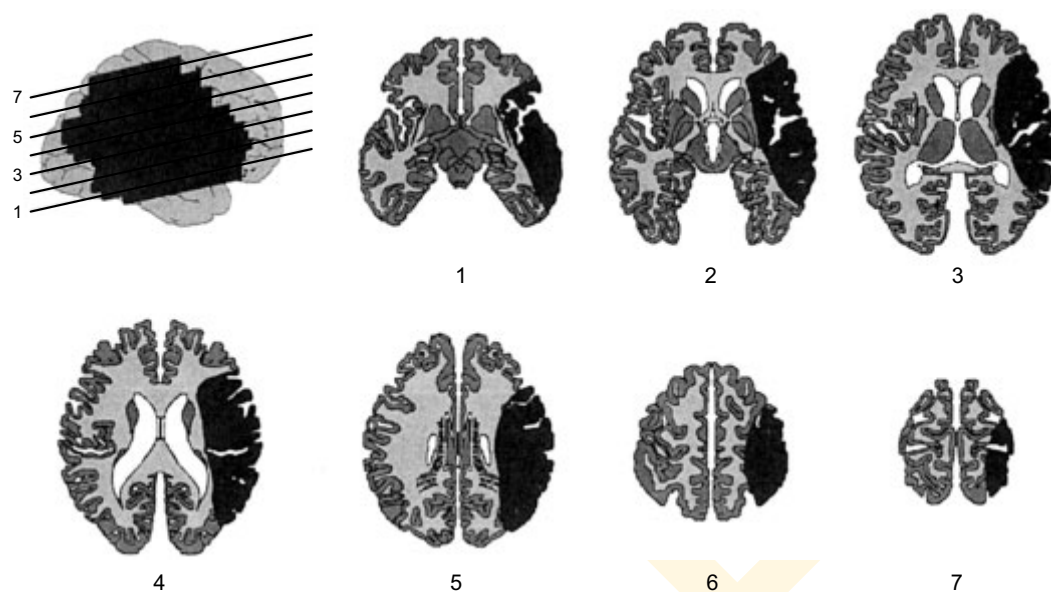
**Ядерный магнитный резонанс.** После того как в применении компьютерной томографии накопился определенный опыт, биофизикам стало ясно, что послойную картину мозга можно восстановить не только с помощью рентгеновских лучей, но путем регистрации иных физических факторов. В итоге предложили технологию, основанную





**Рис. 2.2.11** Методы визуализации мозговых структур: человеческий мозг можно исследовать неинвазивными способами, в частности, обычной рентгенографией (a), компьютерной томографией (b), магнитно-резонансной томографией (c), а также позитронно-эмиссионной томографией (d).





**Рис. 2.2.12** Множественные «срезы» головного мозга человека: продемонстрированы компьютерные изображения мозга с обширным поражением правой теменной доли. Область поражения отмечена черным (Ro & Rafal, 1996).

на принципе *ядерного магнитного резонанса*: голова пациента помещается в магнитное поле высокой напряженности (см. рис 2.2.11С), поляризуя молекулы мозгового вещества подобно стрелкам компасов. Благодаря поляризации магнитные поля молекул выстраиваются в том же направлении, что и направление магнитного поля. Когда мозговую ткань, находящуюся в таком состоянии, подвергают воздействию радиоволн, молекулы начинают испускать регистрируемые радиосигналы, несущие информацию об их химических свойствах. Эта информация восстанавливается благодаря детекторам, расположенным вокруг головы пациента. Результаты измерений обрабатываются по алгоритмам компьютерной томографии, восстанавливающим карту мозговой ткани. Отметим, что метод магнитного резонанса, не подвергая мозг воздействию вредоносных рентгеновских лучей, дает гораздо более детализированную его картину, нежели компьютерная томография.

**Позитронно-эмиссионная томография.** Сканирование мозга по методу *позитронно-эмиссионной томографии (ПЭТ)* принципиально отлично от описанных выше двух технологий: в кровоток испытуемого вводится небольшая доза радиоактивного вещества, которое в тканях мозга ведет себя также как и при автордиографии — изотопы по мере распада спонтанно испускают позитроны, что регистрируется датчиками, расположенными вокруг головы пациента (2.2.11D).

Измерив по разным направлениям уровень позитронной эмиссии и обработав полученную информацию по алгоритмам компьютерной томографии, можно восстановить картину распределения радиоактивности в мозге. Поскольку уровень радиации в каждой точке четко соотносится с величиной нейронной активности, выстраивается т.н. *радиоактивный трэйс* этой активности. Мы говорим при этом, что ПЭТ-метод отражает то, насколько активны нейроны всех отделов мозга при воздействии стимула. Отметим, что позитронно-эмиссионная томография выполняет измерение не напрямую, но

косвенно — через кровоток. Когда та или иная область мозга активна, ее кровоснабжение усилено и это усиление регистрируется через усиление эмиссии позитронов.

Даже несмотря на то, что ПЭТ-изображения менее детализированы, чем полученные с помощью ЯМР-технологии, серьезным преимуществом первых является не только структурная, но также функциональная регистрация ситуации. Сие позволяет исследователям вдобавок к анатомической картине мозга получать картину его текущей функциональной активности. К примеру, можно получить ПЭТ-изображения в момент предъявления наблюдателю сложного зрительного стимула, когда клеточная активность в затылочных долях головного мозга выше, чем в остальных областях коры. Когда испытуемому предъявлен стимул звуковой, то активность зрительных областей мозга оказывается низкой, но при этом клетки височных долей мозга активными. Таким образом, метод позитронно-эмиссионной томографии можно применять с целью выявления активности тех или иных областей мозга при выполнении сложных зрительных задач, в частности таких как чтение, зрительное воображение, напряжение внимания и память.

Слабым местом ПЭТ является то, что на сбор данных уходит не менее 40 сек. и, следовательно, невозможно отследить скоротекущие процессы, то есть те, что длятся секунды и доли секунды. Еще один недостаток метода — низкое пространственное разрешение (ощутимо меньшее, чем у ЯМР).

**Функциональный ядерно-магнитный резонанс.** Одной из весьма многообещающих разработок в технологии визуализации мозговых структур явился метод *функционального* ядерно-магнитного резонанса, позволяющий выполнять оценку как анатомической структуры областей мозга, так и их функциональной активности. Такой вариант ядерно-магнитного исследования обычно именуют «ФЯМР». Подобно ПЭТ функциональный ЯМР базируется на том, что кровообращение в высокоактивных участках мозга резко повышено, хотя некоторые разновидности метода ориентируются на концентрацию кислорода в крови, питающей различные отделы мозга. Отметим также, что ФЯМР превосходит ПЭТ-технику как по пространственному разрешению, так и по скорости. Более того, метод не требует инъекции радиоактивных веществ и не дает вредных побочных эффектов. Однако технология ФЯМР еще достаточно молода и пройдет не один год пока нейробиологи по достоинству оценят ее возможности. Наконец отметим, что метод функционального ЯМР может в итоге выступить в роли нейробиологического «микроскопа»: передовой технологии, которая навсегда изменит тактику функционального исследования мозга.

## 2.3 ТЕОРИЯ ИНФОРМАЦИОННОЙ ОБРАБОТКИ

В оставшейся части текущей главы мы дадим теоретический фундамент нашей книги: во-первых, обсудим теоретическую подоплеку информационной парадигмы, на которой строятся современные теории зрения; во-вторых, поговорим о базовых концепциях информационного подхода — концепции *представления* и концепции *обработки (процессинга)*, а также обсудим ряд важных сопутствующих вопросов. В конце главы мы сформулируем некий общетеоретический концепт, структурирующий наш разговор о зрительном восприятии и состоящий из четырех этапов информационной обработки: от ретинального изображения до распознавания объектов. В дальнейших главах книги мы углубим и детализируем знание, полученное в текущем разделе.

Следует заранее предупредить читателя о том, что материал данного раздела абстрактен и теоретичен: сие, вероятно, создает особые сложности в его восприятии, по-

сколько он расположен в самом начале книги, то есть до разбора специфических примеров. И все-таки очень важно как можно раньше дать читателю основы информационного подхода, поскольку их понимание потребуется при дальнейшем объяснении материала.

Информационная парадигма — это базис рассуждений о психической деятельности человека как о вычислительном процессе — рассуждений, успешно применимых не только к зрительному восприятию, но к широкому диапазону когнитивных феноменов в восприятии слуховом, памяти, речи, суждениях, мышлении и решении разного рода задач. Фактически информационный подход, нами обсуждаемый, многие исследователи относят к т.н. *куниановой парадигме* познания (Lachman, Lachman & Bufferfield, 1979; Palmer & Kimchi, 1986).

Вышеупомянутый идеолог науки Томас Кун (1962) определил понятие *научной парадигмы* следующим образом:

⇒ **Научная парадигма** — это комплект рабочих гипотез, на которых аналитик строит исследование по той или иной проблеме.

Гипотезы научной парадигмы обычно задействуют претеоретические и метатеоретические пути концептуализации основных проблем и оптимальных путей их решения. Кун описывает ньютоновский взгляд на физику как парадигму, которая оставалась неизменной с XVII века по начало века XIX-го. Несмотря на то, что с ньютоновых времен теоретическая физика ушла далеко вперед (к примеру, на свет появились максвелловские уравнения магнитных полей), ни одно из новых открытий не потребовало пересмотра фундаментальных ньютоновых посылов о природе физического мира. К примеру, Ньютон и его последователи, явно или неявно, но предположили, что имеет место количественное отличие между массой и энергией, что время абсолютно и что случайностей не бывает. Квантовая механика и эйнштейнова теория относительности закрыли ньютонову парадигму и открыли новую, построенную на новых гипотезах, в частности на том, что масса и энергия эквивалентны, что время относительно и что случайности носят вероятностный характер.

Тот тезис, что информационный подход — суть парадигма когнитологии, в частности, ее подраздела — визиологии, основан на широко распространенной точке зрения, что характер ментальных процессов можно истолковать в информационной логике. И хотя по сей день остаются исследователи, отрицающие информационный подход (например, большинство известных сторонников гибсоновской теории экологической оптики) — сегодня он является отправной точкой всех известных концепций зрительного восприятия.

### 2.3.1 Вычислительная метафора

Развитие теорий зрения жестко определялось методами исследования, доступными на том или ином историческом этапе. В визиологии наиболее значимой вехой безусловно стало появление компьютеров. Внедрение компьютерных технологий в нашу науку шло по двум основным направлениям:

— первое — это обкатка новых теорий зрительных процессов с целью проверки того, как эти теории работают в реальных приложениях, в частности в такой новоявленной сфере как компьютерное зрение (см. раздел 2.2), в которой перед ЭВМ ставится задача распознавания окружающего мира по возможности тем же путем, каким идет зрение человека;

— второе представление оказалось еще серьезнее и легло в основу ведущей теоретической аналогии ментальных процессов в рамках информационной парадигмы. В общих чертах эта аналогия состоит в том, что ментальные процессы (в частности, зрительное восприятие) так же относятся к мозгу, как программы к ЭВМ. Иными словами, психическая деятельность уподоблена программе, исполняемой компьютером под названием «мозг»; психика — это «софт» биологических вычислений, а мозг — «хард».

За последние 30-40 лет построили множество теорий восприятия, воплощенных в программном виде, однако еще больше были сформулированы в рамках информационной концепции, но без какого-либо фактического программного выражения. В текущей главе мы дадим краткий обзор этих теорий и поймем, отчего они оказались столь популярны.

Вычислительная аналогия по большей части заменила теоретические аналогии, описанные нами в разделе 2.1, в частности: химическую аналогию, лежащую в основе структурализма; полевую аналогию, на которой строится гештальт-теория; аналогию резонанса, положенную в основу гибсоновской теории информационного захвата. Отметим, что компьютерная аналогия легко совместима с дедуктивной аналогией конструктивизма, поскольку, фактически, строится на тех же логических ходах, что выполняются компьютером. Многие когнитологи убеждены в том, что аналогия между функцией мозга и работой компьютерных программ много глубже, чем просто аналогия и что мозг — это в буквальном смысле биологический компьютер. Ежели сие верно, то это значит, что ЭВМ может воспринимать окружающий мир тем же путем, что и человек.

Некоторые исследователи полагают, что аналогия «психика/мозг = программа/компьютер» выходит за рамки программного компьютерного «зрения», являясь фактически осознанным зрительным опытом. Такую позицию занимают приверженцы идеи т.н. *жесткого искусственного интеллекта* (Searle, 1980), полагающие, что должным образом запрограммированная машина фактически реализует мыслительный процесс, включая осознанный опыт. Сия позиция контрастирует с точкой зрения, именуемой «мягким искусственным интеллектом», приверженцы которой считают, что машина лишь имитирует ментальные процессы, течение сознания и пр. Верна ли точка зрения «жестких интеллектуалов» или нет — не столь важно, но важно то, что она дала толчок бурной дискуссии с весьма вескими аргументами с обеих сторон (см. Searle, 1980 и комментарии к этой работе).

Не принимая во внимание крайние позиции, тем не менее отметим, что ежели мыслительные процессы, действительно аналогичны работе компьютерных программ, сие подталкивает нас к изучению того, как происходит обработка информации в компьютере, поскольку последнее — плацдарм потенциального прорыва в нашем понимании работы мозга и того, как эффективнее всего изучать эту работу.

### 2.3.2 Три уровня информационной обработки

Дэвид Марр в своей выдающейся монографии «Зрение» (D. Marr, 1982) выделяет три уровня нашего толкования процессов информационной обработки: *вычислительный*, *алгоритмический* и *имплементарный*. Автор дает т.н. *метатеоретический анализ* информационной парадигмы<sup>1</sup> и доказывает, что существуют серьезные концепту-

<sup>1</sup> Метатеория — это теория о теориях, концепция, не анализирующая зрение как таковое, но суть теорий зрения.

альные отличия между предложенными уровнями и каждый из них исключительно важен в понимании механизмов зрения — если угодно — как информационных механизмов.

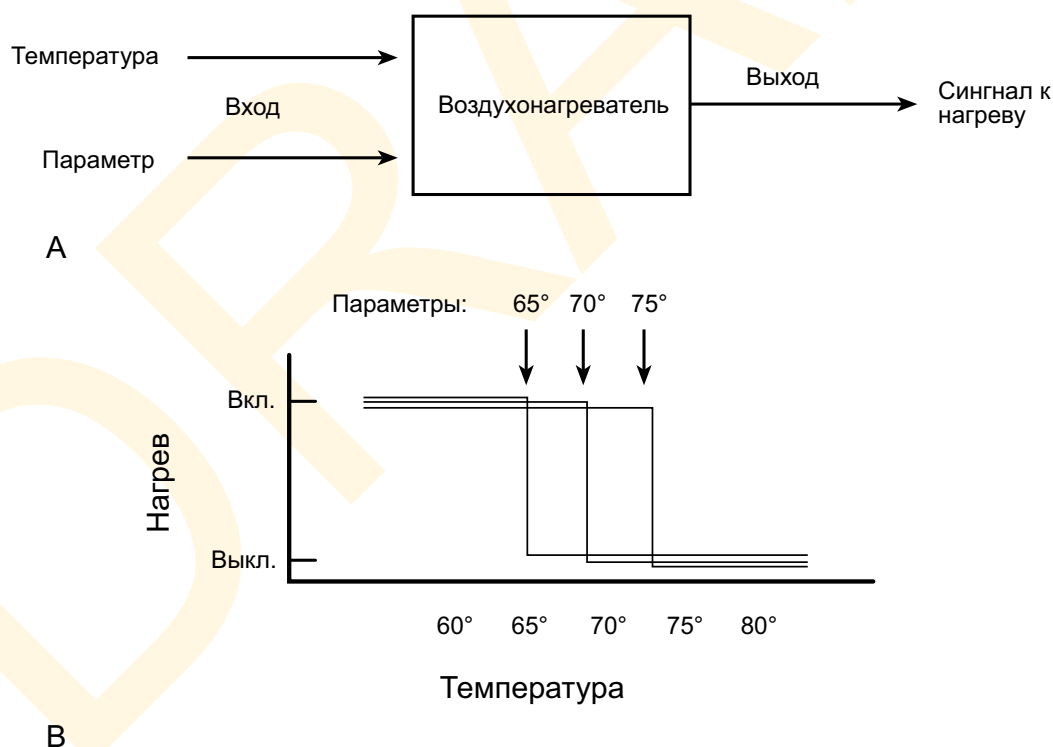
### 2.3.2.1 Вычислительный уровень

Наиболее абстрактное марровское описание — описание вычислительного уровня.

⇒ **Вычислительный уровень** информационной обработки — это уровень, на котором информационный поток ограничивается с целью мэппинга (проекции) входных данных на данные выходные.

Сей уровень позиционирует то, что вычисления имеют место и что информация не зависит от того, каким путем она была получена.

Чтобы проиллюстрировать сказанное, приведем в пример простейшую информационную систему — бытовой воздухонагреватель. «Вычисление» состоит в том, что текущее значение температуры воздуха в помещении и целевое значение, т.е. значение, заданное пользователем (входные данные), аппарат превращает в «да-нет»-сигнал нагревателя (выходные данные), — сигнал, зависящий от того выше или ниже целевого значения температуры воздуха в помещении и (см. рис. 2.3.1 А). Рис. 2.3.1 В демонстрирует сей двойственный мэппинг как функцию от температуры в трех различных точках: если температура воздуха оказывается ниже предустановленной, то включа-



### ВЫЧИСЛИТЕЛЬНЫЙ УРОВЕНЬ

**Рис. 2.3.1** Описание работы воздухонагревателя на вычислительном уровне. А — блок-схема, демонстрирующая вход и выход воздухонагревателя. В — график, демонстрирующий поведение устройства на входе и выходе (по трем установкам) как функцию от температуры.



ется нагреватель («да»-сигнал), но если она равна или выше — он выключен («нет»-сигнал). Теперь мы можем дать математическое описание вычислительного уровня в виде бинарной функции с двумя переменными:

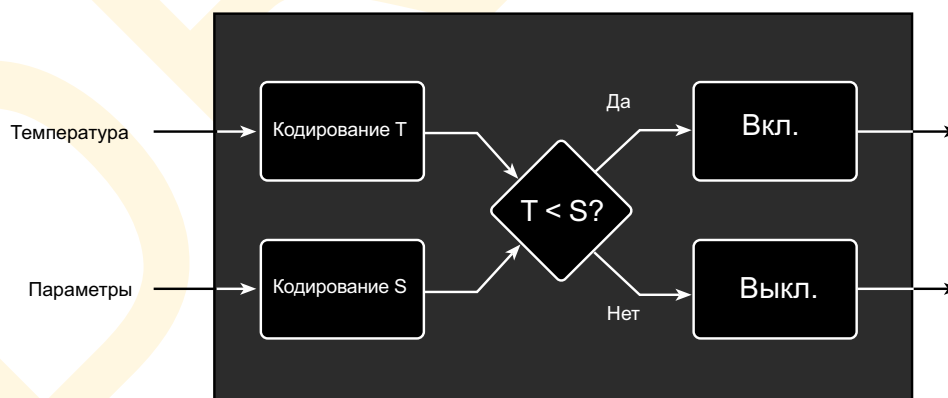
$$O(T, S) = \begin{cases} 1 & \text{если и только если } T < S \\ 0 & \text{если и только если } T \geq S \end{cases}$$

где  $O$  — выходная функция;  $T$  — значение измеренной температуры;  $S$  — целевое значение температуры. Отметим, что мы еще ничего не сказали о том, какая функция получена — мы лишь определили, каковы данные на входе и как они формально соотносятся с данными на выходе. Итак, по Марру, сие является описанием работы нагревателя на вычислительном уровне.

### 2.3.2.2 Алгоритмический уровень

Средний уровень в марровской иерархии — алгоритмический. Алгоритмическое описание процессов специфичнее описания вычислительного и, в частности, тем, что последнее рассуждает о вычислениях в информационной логике. В принципе, существует множество различных путей по которым можно выполнить вычислительный мэппинг входных данных на выходные, и точно также существует множество компьютерных программ, способных решить эту же вычислительную задачу. Таким образом, алгоритмический уровень более чем вычислительный согласуется с концепцией программирования, принятой в вычислительной технике.

Как говорят математики, чтобы построить алгоритм решения т.н. *численно определенной задачи*, мы должны договориться о том, чем представлена область входных данных, чем выходных, и четко позиционировать *алгоритм действий*, преобразующих первые во вторые. Разумеется, концепцию отображения и обработки мы разберем весьма поверхностно, и на сей момент будем говорить об отображении как о пути кодирования информации о чем-либо и об обработке как о пути преобразования одного отображения в другое.



АЛГОРИТМИЧЕСКИЙ УРОВЕНЬ

**Рис. 2.3.2** Описание работы воздухонагревателя на алгоритмическом уровне: черный ящик демонстрирует декомпозицию вычислительной схемы рис. 2.3.1 на этапы кодирования, сравнения и выходных операций.

В случае нашего воздухонагревателя наиболее очевидный алгоритм — это использование одной непрерывной переменной для кодирования температуры и второй — для кодирования целевого значения (см. рис. 2.3.2) с последующим выполнением операции сравнения между этими двумя величинами (дабы определить то, выше или ниже целевой находится фактическая температура воздуха в помещении).

Приведенный пример — стандартный алгоритм работы большинства бытовых воздухонагревателей, но встречаются и другие схемы, например, в которых температура представлена в виде серии бинарных переменных (скажем та, что выше  $60^\circ$  и та, что ниже; та, что выше  $61^\circ$  и та, что выше и т.д.). Отметим, что в отличие от предыдущего алгоритма отображение температуры и целевых ее значений должны быть дискретными. Стоит сказать также, что в сей простейшей системе возможны и другие алгоритмы обработки информации, но главным является то, что такой алгоритм не единственный.

Итак:

⇒ **Алгоритмический уровень** информационной обработки — это уровень, на котором действует алгоритм преобразования одного представления в другое.

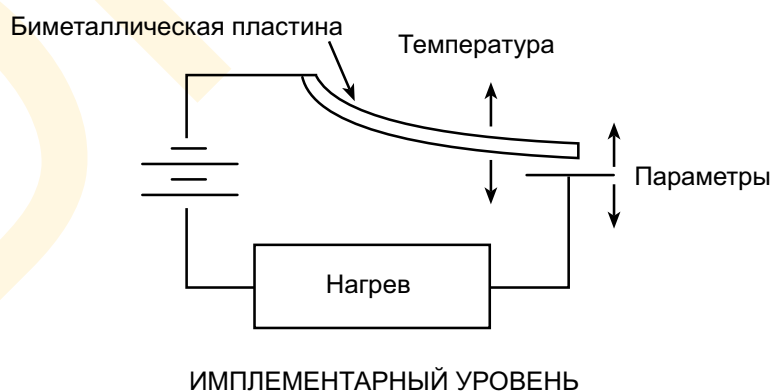
### 2.3.2.3 Имплементарный уровень

Низший уровень информационной обработки — имплементарный.

⇒ **Имплементарный уровень** информационной обработки — это уровень, демонстрирующий то, как алгоритм реализован в данной физической системе.

Одна и та же программа может работать на многих компьютерах, отличных друг от друга по своему физическому устройству, то есть один и тот же алгоритм реализуется физически разными устройствами. Легко догадаться, что под «физически разными устройствами» мы подразумеваем мозг разных особей одного биологического вида, способный реализовать тот или иной алгоритм, подобно тому как разные компьютеры способны реализовать одну и ту же программу.

Для иллюстрации сказанного обратимся к рис. 2.3.3, демонстрирующему один из вариантов конструкции воздухонагревателя, работающего по первому алгоритму: парная кривая — это т.н. биметаллическая пластина, состоящая из двух металличе-



**Рис. 2.3.3** Описание работы воздухонагревателя на имплементарном уровне: на схеме показано физическое устройство, выполняющее операции, позиционированные на рис. 2.3.2, демонстрирующем алгоритмический уровень.

ских полосок с разными коэффициентами теплового расширения. Различие в этих коэффициентах ведет к тому, что при изменении температуры окружающего воздуха кривизна пластины меняется. Одним из концов пластина замыкает электрическую цепь. Меняя вертикальное положение правого контакта, пользователь устанавливает целевое значение температуры в помещении. Замкнут контакт или нет зависит от двух факторов: от высоты свободного конца биметаллической пластины (что определяется фактической температурой воздуха в помещении) и пользовательским положением контакта.

Сие устройство реализует первый алгоритм воздухонагрева и за счет положения правого конца биметаллической пластины управляет непрерывной температурной переменной, тогда как непрерывная переменная целевой температуры управляется позиции пользовательского контакта. Сравнение с фактической температурой выполняется непосредственно, т.е. за счет относительных позиций конца пластины и пользовательского контакта: когда пластина и контакт соприкасаются — цепь замкнута и нагреватель включен; когда не соприкасаются — цепь разорвана и нагрева нет. Очевидно, что механизм рабочий, но далеко не единственный, способный реализовать алгоритмический уровень работы воздухонагревателя. Существует бесчисленное число вариантов регулировки непрерывной переменной температуры воздуха в помещении и вариантов установки пользовательского ее значения, каждый из которых будет работать по одному и тому же алгоритму, показанному на рис. 2.3.2.

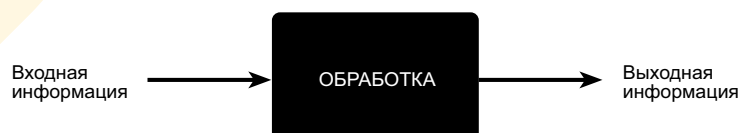
### 2.3.3 Три гипотезы информационной обработки

В 1986 г. Пальмер и Кимчи провели метатеоретический анализ парадигмы информационной обработки в физиологическом аспекте. И хотя анализ этот по началу кажется принципиально отличным от марровского, мы вскоре увидим, что на самом деле и тот, и другой очень близки друг к другу. Плюс к тому, исследователи классифицировали варианты неявных допущений, что лежат в основе теорий информационной обработки в когнитивной психологии. Рассмотрим три наиболее важных.

#### 2.3.3.1 Информационное описание

⇒ **Информационное описание** — ментальные события можно функционально описать как *события информационные*, каждое из которых состоит из трех частей: *ввод информации, ее обработка этой и выход*.

Т.е. суть первого допущения в том, что ментальные события, включая зрительное восприятие, можно описывать как процессы преобразования исходного множества входных данных в множество выходных, т.е. как *информационные события*. Информационное событие можно схематично изобразить в виде черного ящика на *схеме ин-*



**Рис. 2.3.4** Принцип черного ящика: ментальные события можно описать как события информационные, определяемые входной информацией, выходной информацией и обработкой, проецирующей входные данные на выходные.

*формационного потока* (рис. 2.3.4). Если мэппинг ввода-вывода четко определен, то сие позиционирует такие операции как получение информации на входе и определение данных на выходе. Когнитивная теория на данном уровне абстракции соответствует марровскому вычислительному уровню, поскольку задает параметры информационного мэппинга, но без указания на то, как именно должен выполняться этот мэппинг. Отметим, что выходные данные как функция от входных — это идеальный путь спецификации такой операции.

### 2.3.3.2 Рекурсивная декомпозиция

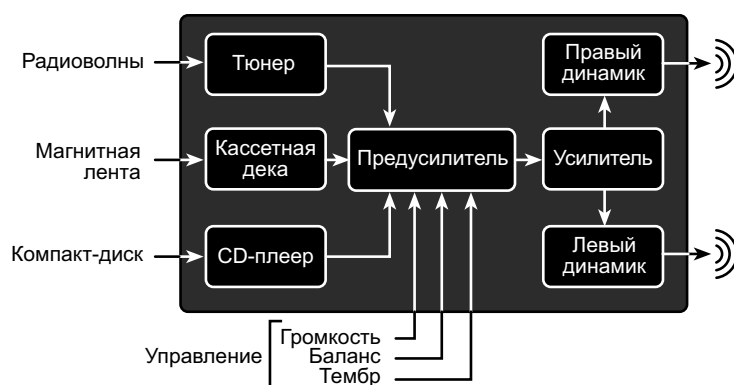
Несмотря на то, что информационное описание — это неотъемлемый атрибут информационной теории восприятия, описание это не является собственно информационной теорией. Гибсоновская концепция информационного захвата, к примеру, позиционирует информационные соответствия между входом (динамический внешний оптический массив) и выходом (результат восприятия окружающей обстановки) и, следовательно, удовлетворяет гипотезе информационного описания. Однако гибсоновская теория не является собственно информационной, поскольку категорически отрицает необходимость анализа внутренних процессов при выполнении мэппинга информации.

Пальмер и Кимчи (1986) позиционировали ключевое требование информационной теории как условие т.н. *рекурсивной декомпозиции*, необходимое для полнейшего описания того, что происходит в черном ящике.

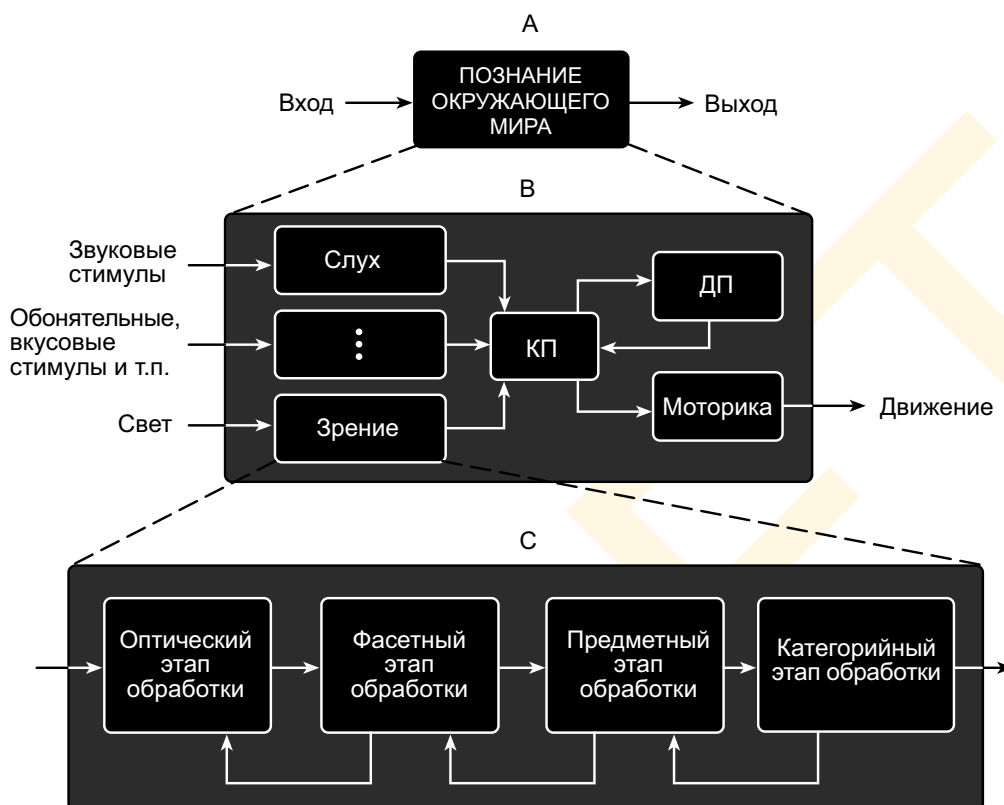
⇒ **Рекурсивная декомпозиция.** Любое сложное (непримитивное) информационное событие, текущее на том или ином уровне, можно полнее описать на подлежащем уровне путем разложения этого события на составляющие и руководствуясь при этом блок-схемой, описывающей временную последовательность взаимоотношений между компонентами.

Проще говоря, идея в том, что мы можем определить черный ящик как совокупность черных ящичков, сопровождаемую спецификацией взаимосвязей между ними. Черные ящички порой называют *фазами* или *стадиями*, в определенной степени независимыми друг от друга.

Простейшим примером декомпозиции является анализ бытового музыкального центра по его функциональным компонентам: тюнер, CD-плеер, кассетная дека, предусилитель, усилитель и колонки (рис. 2.3.5). В большинстве случаев перечисленные компоненты находятся в одном помещении,



**Рис. 2.3.5** Блок-схема стереосистемы. Функциональная декомпозиция музыкального центра раскладывает его на несколько составляющих (черных ящичков) и информационные взаимоотношения между ними (стрелки).



**Рис. 2.3.6** Рекурсивная декомпозиция познания окружающего мира. Акт познания можно описать как информационное событие с несколькими уровнями детализации. Каждая блок-схема деонстрирует функциональную декомпозицию вышестоящего черного ящика (соединены друг с другом пунктирными линиями) на серию простых операций и поток информации между ними.

ДП — долгосрочная память; КП — краткосрочная память.

но в некоторых ситуациях каждый компонент может существовать совершенно самостоятельно и соединяться с остальными с помощью проводов (стрелки на рис. 2.35).

Говоря, что декомпозиция является *рекурсивной*, мы имеем в виду то, что она выполняется в отношении результатов предыдущей декомпозиции, к примеру: компоненты музыкального центра тоже можно разложить на составляющие их внутренние электрические цепи и т.д. Такую процедуру мы назовем *итеративной* (повторяющейся). Итеративная декомпозиция ведет в итоге к тому, что на схеме появляются блоки, соответствующие функциям отдельных электронных компонентов: резисторов, конденсаторов, транзисторов и пр.

Схема декомпонированной информационной системы примерно соответствует марровскому алгоритмическому уровню, поскольку дробит систему на простые составляющие, каждая из которых несет в себе алгоритм, через который выполняется вычисление на более высоком уровне. Однако главное отличие позиции Марра в том, что ученый видел свой алгоритмический уровень как единичный, самостоятельный, тогда как Пальмер и Кимчи рассматривают его как субэлемент иных уровней. Компьютерную программу тоже можно рассматривать и как самостоятельную сущность, и как иерархический элемент первого порядка (второго, третьего и т.д.).



Пальмер и Кимчи (1986) настаивают на том, что рекурсивная декомпозиция — это базис, сердце информационного подхода в психологии, поскольку она определяет способ исследования: от общего к частному. На рис. 2.3.6 дан пример рекурсивной декомпозиции в приложении к процессу познания: разум обозначен черным ящиком (рис. 2.3.6 А), который психологи обычно раскладывают на относительно малое число базовых компонент (рис. 2.3.6 В) и рисуют связи между этими компонентами. Отметим, что предмет нашей книги — зрительное восприятие — представлен на схеме как некий единичный процесс. Поскольку декомпозиция рекурсивная, то каждую ступень можно и дальше декомпонировать на ступени низших уровней и соответствующие связи между этими ступенями. Один из декомпонатов относится к ступени зрительного восприятия. На рис. 2.3.6 С показан четырехкомпонентный процесс, который полнее мы опишем в конце данной главы. В дальнейшем мы покажем, что эти четыре ступени можно декомпонировать на серию более специфичных и простых процессов.

Итак, рекурсивная декомпозиция позволяет постепенно разобраться в сложных процессах информационной обработки: каждая следующая декомпозиция снимает часть сложности, *неявно (имплицитно)* присутствующей в черном ящике, делая ее *явной (эксплицитной)* через связи между операциями на нижележащих уровнях анализа. В результате процессы, идущие в черном ящике, на каждом следующем нижележащем уровне становятся все проще и понятнее. Однако важно помнить, что информационные события на нижележащих уровнях — декомпонатах, должны также располагать информационным описанием того, *что* происходит (в контексте мэппинга входных данных на выходные), но необязательного того, *как* это происходит. Ответ на вопрос «Как?» мы получим, спускаясь на следующие уровни по мере дальнейшей декомпозиции.

Отметим, что принцип рекурсивного декомпонирования системы на иерархические составляющие окажется успешным только в том случае, если система эта иерархична. Американский психолог, нобелевский лауреат Герберт Симон (1969) предложил ряд доказательств того, что человеческая система информационной обработки построена иерархично (по меньшей мере в общих чертах). Ученый охарактеризовал процесс познания как «полудекомпонлируемый», дабы отличать его от сугубо антропогенных систем, таких как, к примеру, стерео аудиосистема из нашего прошлого примера, которую можно полностью расщепить — *декомпонировать* — на составляющие. Симон отличает взаимодействия, возникающие внутри компонентов системы (к примеру, внутри черного ящика), с взаимодействиями между компонентами (к примеру, между двумя и более черными ящиками).

⇒ **Декомпонлируемая система** — это система, в которой взаимодействия между компонентами бесконечно менее значимы, чем взаимодействия внутри компонент.

В современной терминологии такие системы именуются *модулярными*, что указывает на возможность декомпонирования этих систем на отдельные независимо функционирующие компоненты. В своей *модулярной гипотезе* философ Джерри Фодор (1983) выдвинул предположение, что человеческое познание тоже модулярно.

⇒ **Полудекомпонлируемая система** — это система, в которой взаимодействия между компонентами невелики, но не настолько, чтобы пренебрегать ими.

⇒ **Недекомпозируемая система** — это система, в которой взаимодействия между компонентами важнее взаимодействий внутри компонент.

В отношении того, что когнитивный процесс (хотя бы отчасти) дробится на модули многие исследователи восприятия (в частности гештальтисты, коннекционисты и приверженцы экологической оптики) не столь оптимистичны как Симон (1969) и Фодор (1983): большинство гештальтистов, напротив, убеждены в том, что ментальные события следует понимать в логике комплексных, холистических, целостных взаимодействий, происходящих в мозге. Сия концепция никак не стыкуется с той идеей, что когнитивный процесс — это процесс, реализуемый полудекомпозируемой системой, которую можно описать с помощью рекурсивно декомпозируемой блок-схемы.

Много позже некоторые коннекционисты подняли вопрос о правомочности принципа рекурсивной декомпозиции, дав нейронное толкование восприятия и познания. Их аргументы состояли в том, что любое высокоуровневое описание, в частности такое, какое предлагают разного рода схемы — это лишь грубая аппроксимация описания нейронного (Smolensky, 1988). Гибсон и некоторые его последователи занимали противоположную позицию, заостряя свое внимание на информационном описании восприятия на высшем — вычислительном — уровне, но без попытки анализа внутренних процессов.

Та идея, что декомпозицию можно рекурсивно применить к информационным описаниям поднимает исключительно серьезный вопрос: «А когда же следует остановиться?». Общий ответ таков: «Останавливаться нужно тогда, когда декомпозиция дойдет до примитивного, элементарного события определенного вида». Но тогда встает следующий вопрос: «Каков критерий “элементарности”?». Пальмер и Кимчи (1986) выделяют два подхода к данной проблеме: первый, основанный на т.н. *программных элементах*; второй — на т.н. *элементах аппаратных*.

Суть программного подхода в том, что в качестве элемента выступает некий набор вычислений, достаточный для решения той или иной задачи. К примеру, Ньювелл и Симон (Newell, Simon, 1972) предложили некий набор элементов когнитивного процесса и доказали, что в целом набор этот эквивалентен ходам универсальной тьюринговой машины (компьютера). Она назвали эти элементы «программными» поскольку выполняемые ими операции инспирированы требованиями к «программам», написанным на «программном языке» мозга (но не вычислительных машин). Эффективность программных элементов в моделировании широкого диапазона высокоуровневых когнитивных процессов (например, мышления) несомненна, но она недостаточна для построения моделей восприятия.

Аппаратный подход в качестве элемента позиционирует такие операции, какие выполняются базовыми физическими компонентами системы. В контексте ментальных событий типичный теоретический посыл звучит так, что базовой единицей мозга является нейрон и, следовательно, декомпозицию следует остановить, когда операции окажутся функционально эквивалентны информационной обработке, исполняемой на уровне отдельного нейрона. Нейронное моделирование — это общий подход к элементам теории восприятия, в частности, на ранних стадиях обработки (сетчатка), о чем сегодня известно уже довольно многое. Отметим, что идея моделирования, основанного на нейроноподобных элементах, горячо отстаивается коннекционистами.

Важно показать и то, что когда если мы выполним декомпонирование до уровня аппаратных элементов, мы не дойдем до физической системы как таковой. Декомпонированное описание остается внутри информационного домена, не отвечая на вопрос о том, насколько мелко выполнено дробление на простейшие черные ящички. Так про-

исходит потому, что информация принципиально отлична от физических тел или энергии (несмотря на то, что своим существованием она обязана физическому телу и энергии). Поэтому информацию невозможно расчленить на физические составляющие.

Возможно наилучшим путем иллюстрации данной точки зрения является декомпонированная нами на аппаратные элементы бытовая стереосистема (рис. 2.3.5): каждый черный ящик представляет собой функциональный элемент системы — резистор, транзистор, конденсатор и пр. Но даже такое описание не дает данных о том, как функционирует фактическое физическое устройство, состоящее из этих элементов. Сие применимо как к стандартным электрическим резисторам и транзисторам, так и к некоторым высокотехнологическим оптическим элементам или иным устройствам, задействующим аналогичную функцию ввода-вывода. Данное наблюдение вполне соотносимо с марровским разделением на алгоритмический и имплементарный уровни. В обоих случаях мы видим, что физическая и информационная сферы принципиально отличны друг от друга.

### 2.3.3.3 Физическое воплощение

Пальмер и Кимчи (1986) описали связь между информационным и физическим уровнями в третьей гипотезе — гипотезе *физического воплощения*.

⇒ **Физическое воплощение:** в физической системе, чье поведение можно описать как цепь информационных событий, информация заключена в состояниях этой системы (*представлениях*), тогда как операции, использующие эту информацию, — плод изменений в состоянии (*процессов*).

Третья гипотеза перекидывает понятийный мостик между функциональным уровнем «бесплотной», абстрактной информации (и операций с ней) и фактической работой реальной физической системы (или *воплощения* в марровской терминологии). Согласно такой точке зрения информация и операции с ней, говоря техническим языком, существуют в абстрактном домене информационной обработки, тогда как представления и процессы существуют в физическом мире, являясь физическим воплощением информации и операций. Это очень тонкое отличие, не описанное в научной литературе. Отметим, что многие ученые применяют термин «представление» в отношении информационных сущностей и термин «процессы» в отношении изменений в информационном содержимом.

Три обсуждаемых нами уровня — вычислительный, алгоритмический и имплементарный — это основа нашего дальнейшего разговора о теории зрения на страницах этой книги.

Стоит отметить, что львиная доля разработок, касаемых вычислительного уровня выполнена исследователями компьютерного зрения. Они попытались идентифицировать информацию, заключенную в оптических изображениях на сетчатке и позволяющих воспринимать окружающий мир. Выполняя сие, ученые неявно следовали гиббсовскому принципу экологической оптики, изучающему восприятие путем тщательного анализа математических взаимоотношений между проксимальным и дистальным стимулами. На алгоритмическом уровне и программисты, и психологи активно исследуют то, как сложные вычислительные проблемы можно разложить на простейшие составляющие и информационные потоки между ними. Программисты в основном интересуются определением того, насколько хорошо *работают* алгоритмы, реа-

лизирующие восприятие; психологи озабочены тем, насколько хорошо алгоритм *моделирует* человеческое восприятие (при условии того, что эксперименты поставлены верно). На имплементарном уровне программисты иногда воплощают свои алгоритмы в фактические электронные устройства; физиологи и психологи пытаются определить, как мозг фактически обрабатывает визуальную информацию.

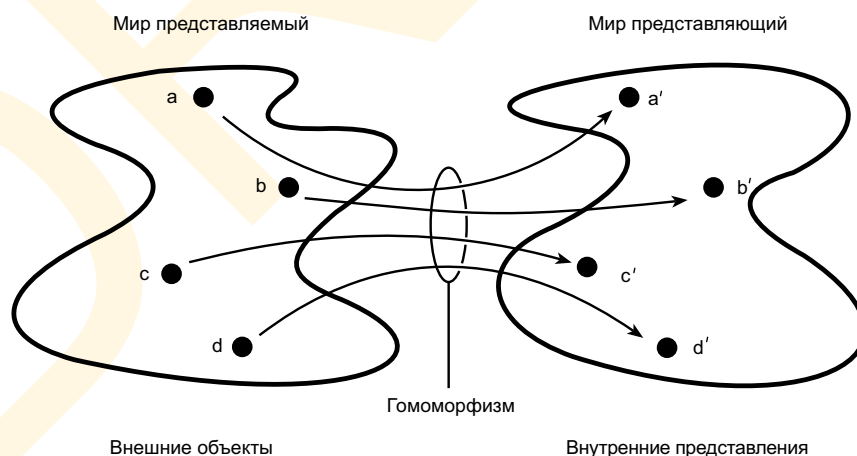
Итак, одной из центральных позиций современного междисциплинарного подхода является тезис о том, что исследование зрения должно идти на всех трех уровнях одновременно.

### 2.3.4 Представления

В предыдущих разделах мы показали, что фундаментальными компонентами информационной системы являются *представления* и *процессы*. Мы и далее будем определять представление как физическую сущность, несущую информацию о предметах и процессах, а последние — как физические преобразования, переводящие одно представление в другое. Однако каков характер информации, дающей зрительное представление и как эта информация организуется? Каков тип процессов, исполняемых системой обработки зрительной информации? Сие основополагающие вопросы, звучащие в контексте информационных теорий, нами обсуждаемых.

Как термин, используемый на протяжении всей нашей книги, *представление* относится к состоянию зрительной системы, воспринимающей внешнюю обстановку — ее объекты и события: сие есть *модель* того, что представляется (Palmer, 1978). Рассуждая таким образом, представление толкуется только как часть большой репрезентативной системы, включающей в себя два взаимосвязанных, но одновременно различных мира: *мир представляемый*, т.е. расположенный вне системы информационной обработки (часто называемый *внешним миром* или *окружающей средой*) и *мир представляющий*, (часто именуемый *внутренним представлением* или просто — *представлением*), расположенный внутри системы информационной обработки.

Каким образом внутренний мир представляет мир внешний? Один из механизмов



**Рис. 2.3.7** Представление гомоморфического мэппинга: внешние (представляемые) объекты проецируются на объекты внутренние (представляемые), в результате чего взаимоотношения между внешними объектами отражаются в соответствующих отношениях между объектами внутренними.

таков, что внутреннее представление несет информацию о структуре внешнего мира за счет схожести структур первого и второго. Формально репрезентативную систему можно проанализировать с позиций *гомоморфизма*: мэппинга объектов одного домена (внешнего мира) на объекты другого домена (внутреннее представление), в результате чего *отношения* между объектами внешнего мира отражаются в взаимоотношениях между соответствующими объектами внутреннего представления (Tarski, 1954). Сей гомоморфический мэппинг схематично показан на рис. 2.3.7: ежели такой мэппинг имеет место, то объекты внутреннего мира выступают в роли точного представления объектов мира внешнего.

Ранее в разговоре о воздухонагревателях мы дали примеры представлений (например, температура была представлена высотой свободного конца биметаллической пластины). Сейчас нам предстоит разобраться в вопросе детально.

Итак, когда температура воздуха в помещении растет, вертикальная позиция конца пластины становится выше. В результате отношения между внешними температурами («холоднее, чем» и «теплее, чем») согласуются с отношениями между высотой конца пластины («ниже, чем» и «выше, чем»). Таким образом, высота свободного конца пластины несет информацию о внешней температуре за счет т.н. *причинно ведомого гомоморфизма*, то есть представляет внешнюю температуру.

Причинный фактор, лежащий в основе гомоморфизма, интересен в двух аспектах:

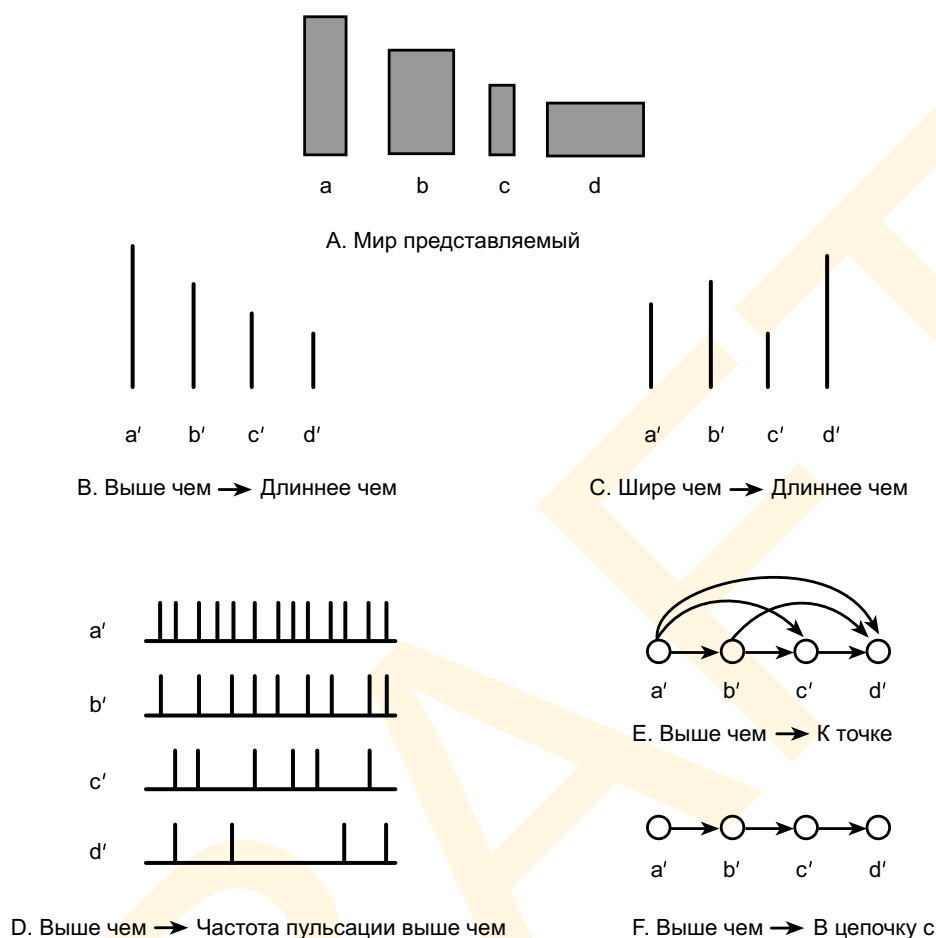
- во-первых, чтобы текущее представление было перцепционным, требуется постоянное обновление. Причинная цепь от событий внешнего мира к событиям внутреннего представления — это идеальный (но не единственный) путь в достижении сего;

- во-вторых, чтобы представление было аутентичным, но не случайным, должна существовать некая связь с представляемым миром. Вновь мы видим, что логика причинных связей оптимальна.

Чтобы нечто работало как представление чего-то, причинно ведомая структурная схожесть между моделью и ее объектом является необходимым, но недостаточным условием. Углубленное требование состоит в том, что должны существовать процессы, дающие представление как суррогат внешнего мира, которому они соответствуют. К примеру, у температуры воздуха множество причинно-следственных связей, которые можно использовать для ее представления — концентрация водяных паров, температура поверхности предметов и пр. — но если бы эти свойства не использовались системой (такой как воздухонагреватель) для контроля поведения этой системы (например, управления нагревом), нельзя было бы говорить о представлении.

На рис. 2.3.8 даны несколько примеров, иллюстрирующих некоторые аспекты сказанного о представлении (Palmer, 1978). При этом мы не даем никаких указаний на то, что эти представления реалистичны: они нацелены лишь на дальнейшую иллюстрацию того, каким должно быть представление, чтобы оно являлось гомоморфичной проекцией внешнего мира. Во всех случаях «внешний мир» представлен набором из четырех прямоугольников, показанных на рис. 2.3.8 А. Несмотря на кажущуюся простоту этих предметов, они несут в себе множество различных свойств, которые можно смоделировать или закодировать в представление, к примеру: высота, ширина и площадь. На рис. 2.3.8 В и 2.3.8 С показано, как два различных параметра данного миниатюрного мира можно представить за счет внутреннего отношения «длиннее, чем». На рис. 2.3.8 В длина линий в представлении отражает относительную высоту прямоугольников во внешнем мире, то есть то, что  $a'$  длиннее, чем  $b'$  в мире В отражает тот факт (сохраняет информацию), что  $a$  выше, чем  $b$  в мире А. Те же рассуждения применимы к





**Рис. 2.3.8** Примеры репрезентативных систем. Каждая система (с B по F) путем различных отношений между соответствующими внутренними объектами представляет специфические отношения между свойствами прямоугольников (A).

любой паре прямоугольников в мире A и соответствующим им линиям в мире B. Приняв такое положение дел, мы можем сказать, что относительная длина линий в B-мире сохраняет информацию об относительной высоте прямоугольников в мире A. Любой вопрос о высоте прямоугольников в A-мире найдет ответ в мире B, если оценить относительную длину отрезков в последнем. То есть, мы говорим, что относительная длина отрезков в B-мире *представляет* относительную высоту прямоугольников в мире A.

На рис. 2.3.8 C длина отрезков представляет параметры объектов A-мира, но уже с целью передачи информации об относительной ширине этих объектов. К этой ситуации применим точно такой же анализ, что был дан в отношении рис. 2.3.8 B.

На рис. 2.3.8 D дано физиологически достоверное представление высоты прямоугольников в показателях частоты пульсации нейронов: каждая вертикальная полоска указывает на спайк; горизонтальная — на паузу. Частота пульсации сих гипотетических нейронов пропорциональна высоте соответствующих прямоугольников: чем выше прямоугольник, тем выше частота. Отметим, что ничто не мешает нам закодировать высоту прямоугольников и обратным образом, то есть, поставив частоту пульса-

ции в обратно пропорциональную зависимость от высоты предметов. В этом случае частота пульсации будет тем выше, чем ниже прямоугольник. Не принимая во внимание очевидные «технические» отличия между частотно-импульсным представлением и кодированием по длине полосок, оба варианта являются *информационным эквивалентом* друг друга, поскольку оба отражают один и тот же факт внешнего мира (Palmer, 1978) и, следовательно, несут идентичную информацию о прямоугольниках.

Принципиально отличным путем кодирования отношений «выше, чем» являются варианты, показанные на рис. 2.3.8 Е и 2.3.8 F: каждый прямоугольник согласуется со специфическим узелком (кружочком) в сети и в этом случае отношения по высоте кодируются стрелками между узелками. На рис. 2.3.8 Е стрелки непосредственно интерпретируют «выше, чем»-отношения: факт, что *a выше, чем b* представлен тем, что *a'* направлена на *b'* с помощью одной из стрелок рис 2.3.8 Е. В данном случае все требуемые стрелки участвуют в представлении, и ни одно из отношений не требует дополнительных действий.

В противоположность рис. 2.3.8 Е рис. 2.3.8 F иллюстрирует много более экономичное представление той же информации. Отношения «выше, чем» кодированы за счет «сцепления с», где *x сцепливается с y*, указывая на то, что существует цепь стрелок, которая начинается на *x* и идет к *y*. Сие определение «сцепления с» является весьма экономичным в том смысле, что для выражения всех потенциальных отношений требуется меньшее число стрелок. Однако эта экономия выходит дороже, потому что для извлечения «ненапрямую» закодированной информации (например такой, как *a'* сцепливается с *c'*) требуется дополнительная обработка. Требуемая информация логически выводится исходя из того, что *a'* сцепливается с *b'*, *a b'* сцепливается с *c'*. Сие стандартный компромисс в представлении информации: больше информации хранится непосредственно — меньше логически выводится с помощью дополнительной обработки; меньше информации хранится непосредственно — больше логически выводится.

Стоит сказать, что существует множество противоречивых толкований характера зрительных представлений: одни считают, что такое представление *локализовано* в некоем специфическом представляющем отделе мозга — другие, что представление это *диссеминировано* (рассеяно) на многие отделы; одни считают, что зрительное представление является *аналоговым* (т.е. непрерывным и «картиноподобным») — другие, что *пропозициональным* (т.е. дискретным и «языкоподобным»); одни исследователи говорят, что достоверный факт представлен эксплицитно или имплицитно — другие, что все зрительные представления можно сократить до некоего конечного набора элементарных атомов, либо что представление это суть *открытая система*. Эти (и многие другие) вопросы мы будем поднимать на страницах нашей книги по мере того как мы будем касаться тех или иных видов представлений, лежащих в основе различных зрительных способностей.

Прежде чем начать разговор о представлении, следует отметить, что хотя мы и говорили о представляемых и представляющих сущностях как об объектах, корректнее говорить о них как о *событиях*. Внутренние представления в головном мозге — суть временное повышение активности нейронов, активности определенного характера, которая скорее событие, нежели объект. События, представляющие содержимое внешнего мира, следует рассматривать как протяженные во времени, но не как вневременные объекты: даже неподвижный предмет может (и должен) рассматриваться как событие, хотя и неизменное. Однако же мы будем уходить от временной структуры событий, поскольку таковая заметно усложняет ситуацию и поскольку не зависи-

мо от того, учитывается временной фактор или нет, цели восприятия остаются прежними.

### 2.3.5 Процессы

Мы уже говорили о том, что представления — это физические сущности в системе обработки информации и что они несут эту информацию. Но откуда берутся эти представления? И как одно представление образуется из другого? Чтобы ответить на эти вопросы мы должны разобрать вторую часть системы обработки информации — *процессы*.

⇒ **Процессы** — это активные компоненты информационной системы, преобразующие информацию из одного представления в другое.

Иными словами, процессы — это динамический компонент информационной системы.

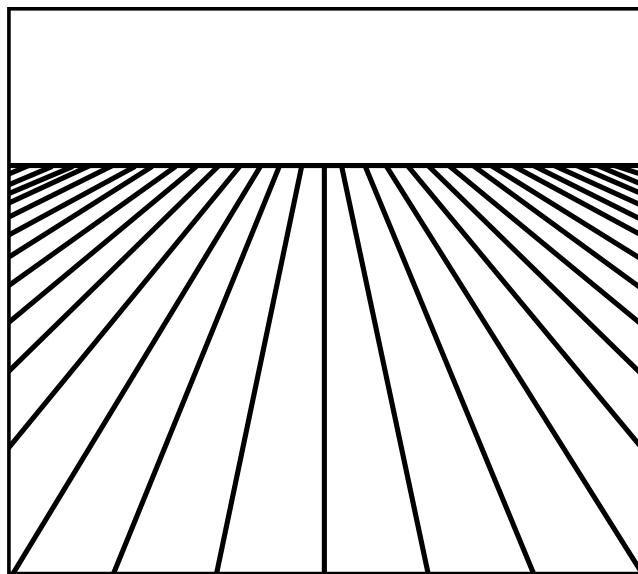
Теперь становится понятно, почему и в представлениях, и в процессах существует необходимость: без представлений процессам не с чем работать; без процессов нет никакой работы с представлениями. Итак, в текущем разделе мы обсудим *что* и *как* выполняется процессами.

#### 2.3.5.1 ИмPLICITная и эксплицитная информация

Одним из важнейших аспектов информационной обработки является *эксплицирование* (извлечение) на выходное представление той информации, что *имплицирована* («запрятана») во входном ее представлении (и наоборот). Процессы не могут создавать информацию о внешнем мире «из воздуха»: они могут лишь преобразовывать эту информацию, перемещая ее извне и изменяя ее форму. Вся информация, с которой работает зрительная система, доступна особи либо из оптической структуры изображения на сетчатке (что есть проекция внешнего мира), либо из внутренних источников (т.е. находящихся в самом наблюдателе). Процессы различными путями собирают и объединяют информацию, дабы построить новые представления, из которых следующие процессы легко извлекут нужные данные, если/когда возникнет необходимость.

Исходно эксплицитная (явная) зрительная информация — это относительная интенсивность света, регистрируемая каждым фоторецептором двухмерной ретиальной мозаики. К сожалению, в своей исходной форме эта информация биологически почти бесполезна, поскольку дает очень мало эксплицитных (явных) сведений об окружающем мире. Гораздо важнее много более сложные вещи, в частности: «Где проходят границы между участками сетчатки, принявшими на себя проекции различных поверхностей?»; «Где в сцене расположены эти поверхности?»; «Как эти поверхности организуются в значимые для особи предметы?»; «Что можно делать с этими предметами?». Ответы на эти вопросы никак не представлены эксплицитно в отклике сетчатки на проекционное оптическое изображение сцены.

Однако существует точка зрения, согласно которой все перечисленное неявно (имплицитно) заключено в ретиальных изображениях, и к чему добавлено некое внутреннее знание. Задача зрительного восприятия состоит в объединении внешней и внутренней информации с целью обретения значимых сведений об окружающей обстановке, воздействующей на организм. Комплексные процессы, способные выполнить сие — главный предмет нашей книги.



**Рис. 2.3.9** Пример конвергенции (схождения) параллельных линий: линии на плоскости, уходящей в глубину пространства сближаются в направлении т.н. точки схода согласно закону перспективной проекции.

### 2.3.5.2 Обработка как инференция

Согласно информационной точке зрения имплицитную (неявную) информацию можно превратить в эксплицитную (явную) путем мэппинга одного представления в другое. Процессы, выполняющие такой мэппинг, можно толковать прямо по Гельмгольцу, т.е. как *умозаключения (инференции)*, хоть и бессознательные. Чтобы продемонстрировать характер инференций обратимся к классическому силлогизму.<sup>1</sup> Исходно эксплицитная информация дается в форме посылов (к примеру, «Все люди смертны» и «Джон — человек»), а логика предлагает правила, по которым имплицитную информацию, заключенную в посылах, можно эксплицитировать в вывод («Следовательно, Джон тоже смертен»). Есть даже мнение, что посылы исходно несут в себе выводы, но только имплицитно

— неявно, и нужно применить некие логические операции, чтобы сделать эти выводы эксплицитными — явными.

Итак, перенеся сию «инферентную» позицию на сферу зрительного восприятия, мы можем получить «посылы», содержащие как собственно ретинальные изображения, так и некие накопленные ранее знания или априорные *версии*, на которые опирается наблюдатель в контексте перцепционной обработки. К примеру, взяв ретинальное изображение, в котором несколько линий сходятся в одной точке (см. рис. 2.3.9) и прибавив к нему версию о линейной перспективе (т.е. о том, что линии, сходящиеся в некоей точке горизонта на проекционном ретинальном изображении, фактически параллельны и уходят вдаль в реальном мире), зрительная система делает вывод о том, что сходящиеся линии на изображении в действительности параллельны и удаляются от наблюдателя. Сей простейший пример позволяет нам толковать перцепционную обработку как некую форму инференции (умозаключения), то есть то, о чем еще в конце XIX в. говорил Гельмгольц (1867/1925) и что исследовали Грегори (1970) и Рок (1983).

Однако же, несмотря на формальное подобие зрительной обработки логическому умозаключению между первым и вторым существует ряд серьезных отличий. Обычные логические построения, выполняемые человеком при решении силлогизмов, как правило весьма взвешены, неторопливы, вербальны (словесны) и сознательны. Тогда как зрительные инференции в целом легки, быстры, невербальны и бессознательны. Действительно, немногим из нас известно, что линии, параллельные в реальном мире,

<sup>1</sup> *Силлогизм* (гр.) — одна из форм дедуктивного умозаключения в формальной логике, когда из двух данных суждений (посылов) вытекает третье (заключение). — *Прим. пер.*

проецируются на сетчатку в линии сходящиеся. Но даже те, кто в курсе, например, художники и фотографы, не осознают сей факт в своем повседневном восприятии мира.

В каком же смысле можно говорить о том, что в зрительной обработке присутствуют умозаключения? Чтобы получить ответ, давайте тщательнее проговорим то, на что направлена перцепционная инференция (умозаключение) и как она возникает в рамках информационной системы.

Существует два основных вида инференций: *дедуктивные* и *индуктивные*. Оба вида — это пути *объединения* информации и прихода к заключениям. Но эти пути — разные.

**Дедуктивные инференции.** Мы уже рассматривали классический пример дедуктивной инференции в силлогизме «Джон тоже смертен». Прочие формы символической логики также попадают в класс дедуктивных инференций, являясь также стандартными математическими операциями (если у нас было два яблока, но нам дали еще два, то логический вывод таков, что у нас теперь их четыре).

Одним из ключевых моментов дедуктивной инференции является то, что ее выводы *точны* до тех пор, пока исходные посылы истинны. Если последние окажутся ложны, то любым дедуктивным выводам, на этих посылах основанным, верить нельзя: если некто утверждает, что все люди бессмертны, то получается, что Джон тоже бессмертен (хотя это не так).

**Индуктивные инференции.** В противоположность дедуктивным индуктивные инференции по определению *неточны* и *вероятностны*, даже если исходные посылы истинны. Классическим примером индуктивной инференции является вывод о том, что все люди смертны. Читатель вряд ли воспринимал данный посыл как вывод поскольку мы использовали его именно как посыл. Но на чем основано наше доверие к последнему? Тезис о смертности человека невозможно вывести из того, что некоторые люди умерли (к примеру, г-н А умер, г-н В умер, г-н С умер и т.д.), поскольку, сугубо умозрительно, может существовать человек, ныне здравствующий или еще не родившийся, который никогда не умрет. Следовательно, весь прошлый опыт человечества не гарантирует истинности глобального вывода о смертности всех людей. Итак, мы говорим, что индуктивные умозаключения неточны поскольку основываются на *неполных* и *вероятностных* данных.

Большинство инференций в зрительной обработке индуктивны, поскольку не гарантируют истинности (и по большей части от того, что пытаются решить проблему инверсии, носящую неточный и вероятностный характер). К примеру, версия, что линии, которые сходятся в сетчаточном изображении, в реальном мире параллельны, не всегда верно, но верно лишь с определенной долей вероятности (достаточно большой): множество сходящихся линий суть проекция линий параллельных, уходящих вдаль, но при этом и непараллельных сторон трапеций и прочих многогранников, которые, разумеется, тоже в изобилии встречаются во внешнем мире. Поэтому зрительная система никогда не даст гарантии того, что сходящиеся линии рис. 2.3.9 — это проекционное изображение параллельных линий, уходящих вдаль, а не аналогичного рисунка.

### 2.3.5.3. Скрытые версии

Предыдущий разговор определил генеральную стратегию зрительных инференций. Хотя множество ключевых процессов в зрительном акте скорее индуктивны нежели дедуктивны, мы все же можем говорить о них как о дедуктивных, но работающих при поддержке т.н. *скрытых версий* (Cutting, 1991). Фактически такая работа —



это то, что мы делаем в нашем примере с параллельными линиями. Эксплицитно посыл, что линии сходятся в изображении, сам по себе не позволяет сделать вывод об их параллелизме в реальности, за исключением лишь некоторой вероятности *наведения* или, как еще говорят, *индукции* (т.е. влияния того, что в предыдущем опыте сходящиеся линии в реальном мире были параллельными). Однако, допустив истинность дополнительного посыла о предположительном параллелизме, вывод об этом параллелизме можно сделать дедуктивно на основании ретинального изображения в связке со скрытым допущением о линейной перспективе.

Ежели дополнительные версии оказываются ложными, выводы могут оказаться неверными. Фактически, это случай самого рис. 2.3.9, поскольку линии действительно непараллельны и действительно сходятся. Сие как раз предмет толкования теориями зрения факта существования и природы зрительных иллюзий, которые детально мы рассмотрим в 7-й главе. К примеру, многие ученые полагают, что т.н. *иллюзии Понцо* (рис. 2.3.10) рассчитаны на ошибочную версию о линейной перспективе, (Gregory, 1970; Rock, 1975, 1983). Если линии, сходящиеся в ретинальном изображении, в реальности параллельны и горизонтальные линии лежат на этой же плоскости, то верхняя горизонтальная линия действительно будет длиннее нижней. Поскольку на самом деле линии лежат на бумаге, разница в длинах иллюзорна.

Рассматривая вопрос зрительной обработки в логике дедуктивных/индуктивных инференций, профессор корнельского университета Джеймс Каттинг (1991) обобщил спор между сторонниками непосредственной (гибсоновской) и опосредованной (гельмгольцевской) теориями восприятия. Ученый назвал гибсоновскую позицию «непосредственной» в том плане, что зрительное восприятие дедуктивно, т.е. базируется исключительно на информации, извлеченной из оптического изображения на сетчатке (следовательно — точной). Гельмгольцевскую позицию бессознательных инференций Каттинг истолковал как т.н. *индикативное восприятие*, которое хоть и основано на сетчаточном изображении, но работает только при наличии определенных допущений



**Рис. 2.3.10** Иллюзия Понцо как бессознательная инференция. Поскольку сходящиеся линии (рельсы) бессознательно интерпретируются как параллельные, уходящие в глубину пространства, то верхняя (дальняя) горизонтальная линия выглядит длиннее, чем нижняя (ближняя), несмотря на то, что на рисунке их длины совершенно одинаковы.

— версий, часть из которых окажутся ложными в нестандартных ситуациях, т.е. породят зрительные иллюзии.

По мнению таких «индирективистов»<sup>1</sup> как Рок (1983, 1997) и Грегори (1970) важной функцией перцепционного обучения является внедрение наиболее подходящих скрытых предположений, благодаря которым восприятие оказывается веридикальным настолько часто, насколько это возможно.

Многие исследователи, изучавшие вопросы обработки информации рассматривают зрительные процессы (т.е. смену одного представления другим) как построенные на индуктивных инференциях, т.е. умозаклчениях, в которых используются скрытые версии. Однако сие необязательно указывает на участие в работе дискретных (языкоподобных) логических правил, поскольку множество зрительных процессов работают, вероятнее всего, с непрерывными представлениями, но не с дискретными, и процессы, через которые реализуются инференции, более похожи на арифметические вычисления, нежели работу символической логики. Все же у нас нет оснований сводить процессы к работе только лишь с символическо-логическими представлениями, поскольку математические операции также являются примерами дедуктивных инференций. Фактически, количественные процессы, преобразующие аналоговые представления, являются общими для всех современных вычислительных теорий зрения (что мы покажем в следующих главах).

В акте зрительного восприятия сцены задействуется множество скрытых версий. Хорошо, если все они ведут к одному выводу (что чаще всего и бывает). А если нет? В случае, показанном на рис. 2.3.9, к примеру, можно выдвинуть «равнодистантную версию»: коль скоро проекции двух объектов на сетчатку равновелики, то объекты эти равноудалены от наблюдателя (Cogel, 1965, 1978). Если сие верно, то верен и вывод о том, что линии в реальном мире непараллельны, но при этом сходятся в сетчаточном изображении (что и есть на самом деле). Каким образом зрительная система определяет, какая инференция верна в ситуации конфликтующих версий?

**Софт-ограничения.** Если, к примеру, две разные версии ведут к разным выводам, мы не можем сделать никакого заключения из обеих, либо же одна версия (или непрерывная цепь версий) должна быть отобрана в качестве предполагаемо верной. Такое решение явится, как говорят, *неспецифическим*, бо потребует алгоритма отбора доминирующей версии. Однако же, по-видимому, существуют альтернативные схемы, более подходящие вероятностному характеру зрительных инференций.

Одна из альтернативных схем состоит в моделировании скрытых версий как *софт-ограничений*, т.е. ограничений информационных, принимаемых в той или иной ситуации, но которые можно проигнорировать в иных ситуациях. Софт-ограничения варьируют от слабых до очень сильных, из-за чего зрительные инференции оказываются порой весьма противоречивыми.

Версия перспективного схождения линий, к примеру, выглядит строже, чем версия равнодистантная, поскольку преодолевает равнодистантную тенденцию в восприятии линий рис. 2.3.9 как параллельных. Конкретные коннекционистские модели (см. Приложение В) можно рассматривать как механизмы, улаживающие множество конфликтных софт-ограничений.

**Нечеткая логика.** Еще одним вариантом является использование в моделях зрительных инференций т.н. *нечеткой логики*. В отличие от двух стандартных показате-

<sup>1</sup> Т.е. сторонников концепции опосредованного восприятия.

лей «истинно» и «ложно», нечеткая логика использует промежуточные уровни истинности. Принцип нечеткой логики был сформулирован американским математиком азербайджанского происхождения Лотфи А. Заде (1965) и применен им в нескольких весьма интересных моделях зрительной обработки (Kay & McDaniel, 1978; Massaro, 1987).

**Вероятностные инференции.** Еще одним подходом является подача вопроса в логике *вероятностных инференций*, основанных на теореме Байеса (Geman & Geman, 1984; Yuille & Bulthoff, 1995). Преимуществом таких рассуждений является то, что инференции по определению вероятностны и что разрозненные фрагменты фактов можно интегрировать в единую математическую схему.

Наконец отметим, что модели, построенные на нечеткой логике, математически эквивалентны байесовским правилам принятия решений (Massaro & Friedman, 1990), так что и первые, и вторые не противоречат друг другу.

### 2.3.5.4 Эвристические процессы

Безотносительно к тому, как зрение строит свои инференции, процессы, которые решают ту или иную проблему, задействуя неточную и вероятностную информацию, именуются *эвристическими*. Как было показано ранее, эвристические пути обычно (но не всегда) ведут к верному решению.

Для понимания эвристического характера зрительных процессов вновь обратимся к сходящимся/параллельным линиям рис. 2.3.9. Напомним при этом, что исходная проблема состояла в интерпретации конвергирующих линий на сетчатке, в трехмерную пространственную картину. Поскольку в результате проекции трехмерной сцены в двухмерное ретинальное изображение часть информации утрачивается, вопрос невозможно решить однозначно без дальнейших ограничений. То есть версия, что линии, конвергирующие на сетчатке — плод линейной перспективной проекции линий параллельных являлась *эвристической версией*, которая чаще всего (но все же не всегда) истинна. Инферентный процесс предполагает дальнейшее объединение этой версии с информацией, полученной из ретинального изображения, а также последующий вывод о том (или «интерпретацию»), что сходящиеся линии фактически параллельны и лежат на плоскости, протяженной в глубину.

Видимо, вопреки гибсоновской идее непосредственного восприятия, введение ограничений, это единственно возможное толкование пути зрительного восприятия в целом. Однако дополнительной информации, представленной в виде логических правил (как, скажем в силлогизмах) не требуется: информация эта, к примеру, может внедриться в паттерн связей внутри сложной нейронной сети. Используя термин «инференция» для описания таких процессов, можно показать их некоторую метафоричность и тем ослабить пафос утверждения, что восприятие работает с помощью бессознательных инференций. Однако, как мы показали в начале, бессознательные инференции все же должны быть слегка метафоричными, поскольку обычное умозаключение очевидно — инертно, трудоемко и осознанно, тогда как восприятие — стремительно, легко и бессознательно. Важным моментом является то, что восприятие опирается на процессы, которые вполне можно рассматривать как инференции, требующие эвристических версий. Далее мы покажем, что многие биологические процессы зачастую оказываются процессами, если угодно, инферентоподобными.

### 2.3.5.5 Скрытые версии против экологического обоснования

Как было показано в разделе 2.1.3, Гибсон отчаянно сопротивлялся идее инференций в восприятии. Ученый проводил встречную мысль о восприятии непосредственном. Стоит однако разобрать то, как именно Гибсон парировал тезис своих оппонентов о том, что феномены зрительного восприятия опираются на инференции, основанные на эвристических версиях (Gregory, 1966, 1970). Как мы уже отмечали, большая часть доказательств в пользу концепции эвристической инференции идет от зрительных иллюзий, в которых скрытые версии оказываются ошибочными (как в случае сходящихся линий понцовской иллюзии, воспринимаемой как параллельные линии, уходящие вдаль).

Главной гибсоновской тактикой был поиск оппонентных доказательств, задающих т.н. *граничные условия* непосредственного восприятия. Гибсон утверждал, что иллюзии появляются только в *экологически некорректных условиях*, то есть ситуациях, которые редко или вообще никогда не возникают в повседневности.<sup>1</sup> Ученый настаивал на том, что его гипотеза прямого восприятия верна, но что это касается только ничем не стесненных наблюдателей, способных активно исследовать окружающий мир. Он доказывал (и весьма убедительно), что когда психолог рассматривает объекты сцены с одной статичной точки на светящемся экране, надев старые давно негодные ему очки — он бесконечно далек от тех условий, в которых эволюционировало зрение человека. Поэтому, полагал Гибсон, результаты зрительного восприятия в таких сугубо искусственных обстоятельствах никак нельзя использовать в целях опровержения его теории. Вооружившись собственной терминологией, Гибсон яростно отстаивал свою позицию как единственно верную: действительно, условия, в которых ученые изучают зрение подчас совершенно ненатуральны, но при этом феномены все-таки существуют и требуют объяснения. Мы же полагаем, что интереснее всего то, что имеется тесная связь между граничными условиями экологического восприятия и эвристическими версиями вычислительных теорий.

Экологическими Гибсон называл такие условия, в которых эвристические версии устойчиво верны, к примеру: наблюдатели не всегда разглядывают сцены с выгодных точек; левый и правый глаза направлены на одну и ту же сцену; в окружающем нас трехмерном мире предметов больше, чем в его двухмерной проекции, построенной из светлых и темных пятен и т.д. Отметим, что ежели граничные условия экологического восприятия — это строго те же ситуации, в которых эвристические версии инферентных теорий истинны — экологическое восприятие всегда окажется веридикальным. Таким образом, Гибсон, в противовес тезису о неверидикальности иллюзий, доказывал верность своей теории прямого восприятия, апеллируя к тому, что «иллюзогенные» ситуации были попросту не экологичны. Именно поэтому Гибсон столь убедителен в своем противостоянии критике, основанной на факте существования зрительных иллюзий. Ученый не отрицал при этом, что иллюзии существуют, но отвергал их появление в естественных экологических условиях.<sup>2</sup>

<sup>1</sup> Ранее экологические рассуждения предлагал Эгон Брунsvик (1956), пытавшийся умалить ценность экспериментальных данных, полученных в условиях маловероятной стимуляции. Гибсон и его последователи развили идеи Брунsvика, точно определив т.н. «ежедневные условия», в которых искусственный стимул (например, 2D-фотография или кинолента) отображает сцены и события. Даже допустив, что мы часто сталкиваемся с такими условиями в нашей современной высокотехнологичной жизни, условия эти далеки от тех, в которых эволюционировало зрение человека.

Точка зрения, принятая в нашей книге, такова, что перцепционная теория должна объяснять все феномены зрительного восприятия, вне зависимости от того согласуются они с экологической концепцией или нет. В этой связи подход, основанный на эвристических вычислительных инференциях, предпочтительнее такого, основанного на непосредственном восприятии, поскольку потенциально концепция эвристическая дает объяснение как экологическим, так и неэкологическим феноменам. Следовательно мы примем то положение, что восприятие эксплуатирует некоторые виды наведенных (индуктивных) инференций и что сие реализуется путем вычислений в нейронных сетях разного рода. В процессе изучения зрительного восприятия мы попытаемся выявить подлежащие ему эвристические версии, процессы им исполняемые и нейронные механизмы, через которые это восприятие реализуется.

### 2.3.5.6 Нисходящие и восходящие процессы

Еще одной важной особенностью обработки перцепционной информации являются ее «направления»: либо «сверху вниз» — *нисходящие*, либо «снизу вверх» — *восходящие*. Пространственная метафора позволяет построить условную схему зрительных процессов, в которой ретинальное изображение расположено внизу, а все промежуточные интерпретации — на более высоких уровнях (см. рис. 2.3.11). «Восходящая» обработка суть *информационно-управляющая*, дающая информацию «низкого уровня» на входе и создающая (или видоизменяющая) «высокоуровневое» представление на выходе. «Нисходящая»-обработка, именуемая также *гипотезо-управляющей* (или *управляющей ожиданием*), относится к процессам, работающим в противоположном направлении и берущими «высокоуровневое» представление на входе и создающими (или видоизменяющими) «низкоуровневое» представление на выходе.

Многие наивно полагают, что зрение — это сугубо восходящий процесс, берущий начало в сенсорной стимуляции, в информации ретинальных изображений и идущий «вверх» к восприятию и концептуальным интерпретациям. Действительно, большинство исследователей единодушны в том, что на ран-

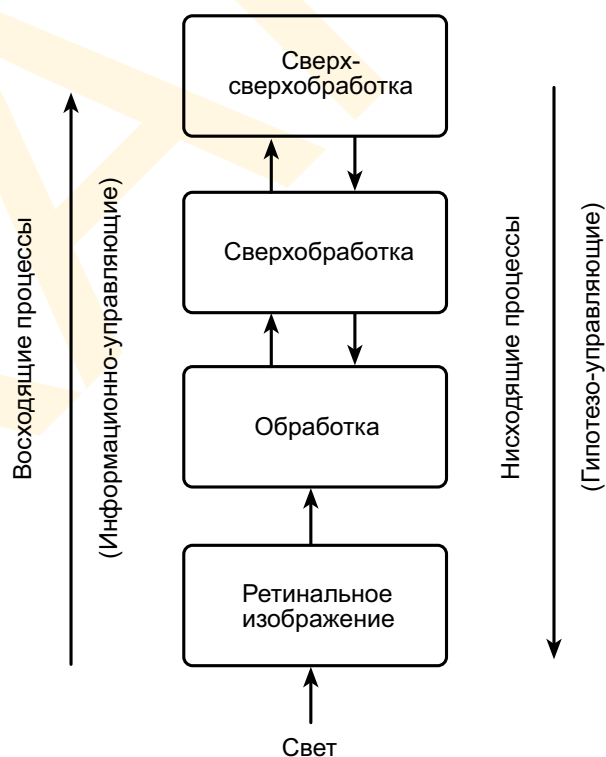


Рис. 2.3.11 Восходящие (информационно-управляющие) и нисходящие (гипотезо-управляющие) процессы.

<sup>2</sup> Множество простых геометрических иллюзий имеют место даже в естественных, ничем не ограниченных условиях просмотра. Де-Люция и Хочберг (1991) показали, что иллюзия Мюллера-Лиера (см. рис. 1.1.4 А) продолжает существовать в самых разнообразных условиях.



них этапах зрительная обработка в целом восходяща, однако есть основания полагать, что сие не может относиться ко всем процессам зрительного акта. В разделе 1.1 мы доказывали, например, что восприятие в текущем состоянии строит версии будущих событий. Эти версии содержат в себе нисходящие компоненты зрительной обработки, свидетельствующие о том, что предшествующие высокоуровневые интерпретации оказывают влияние на текущую обработку на нижних уровнях. К примеру, при чтении мы идентифицируем литеры в строгой зависимости от того, являются ли они фрагментами знакомых слов или бессмысленной цепью символов. Когда мы начнем разбор теорий перцепционной категоризации (гл. 9), мы столкнемся с примерами того, что нисходящая обработка при получении сенсорных данных задействует т.н. *сцепленные модели* близких друг другу объектов и сцен.

Точка, в которой нисходящие процессы начинают доминировать над восходящими — ныне предмет горячих споров: некоторые исследователи уверены в том, что эта точка находится в самом начале зрительной обработки, другие, что много позже.

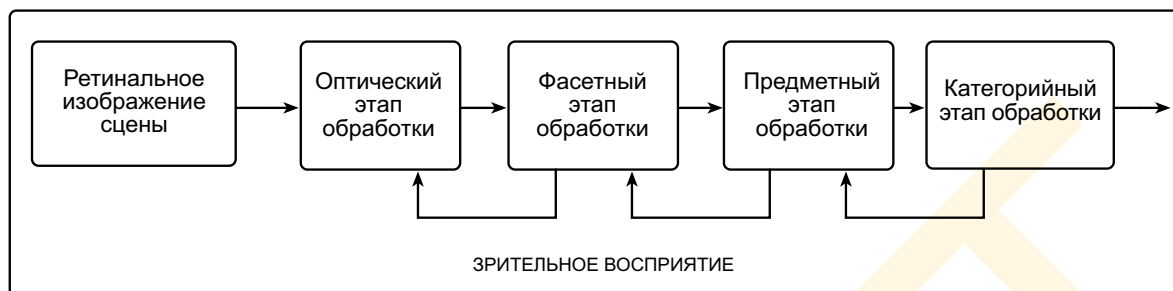
## 2.4 ЧЕТЫРЕ ЭТАПА ЗРИТЕЛЬНОЙ ОБРАБОТКИ

Итак, разобрав общие положения теории информационной обработки, часть из них мы применим к работе зрения, которую на алгоритмическом уровне разложим на четыре больших этапа, начав с ретинального (см. рис. 2.4.1). Поскольку на данную схему мы будем опираться на протяжении всей книги, имеет смысл разобрать ее в самом начале.

Каждый этап зрительной обработки определен выходными представлениями различных видов и процессами, необходимыми для получения этих представлений из представлений входных. Хотя разные ученые по-разному именуют эти стадии, мы будем пользоваться некоей условной искусственной схемой именования, в которой каждый этап зрительной обработки поименован согласно тому виду информации, который она представляет эксплицитно (т.е. в явном виде): *оптический, фасетный, предметный и категорийный* этапы восприятия. Отметим, что предлагаемая схема в целом основана на выдающихся работах Дэвида Марра (1982) и его коллег по МТИ. При этом, конечно, существуют и другие достойные внимания схемы, однако предлагаемая — это отличная основа понимания зрительной функции как вычислительного процесса.

### 2.4.1 Ретинальное изображение

Проксимальный стимул — это пара изображений, представляющих собой двумерную проекцию сцены в точку пространства, в которой находятся в данный момент глаза наблюдателя. На рис. 2.4.2 показано одно из таких изображений — изображение простейшей сцены — чашки, стоящей на белой скатерти на фоне темной стены. Оптическое изображение, воздействующее на сетчатку, пространственно и тонально непрерывно, но его регистрация мозаикой ретинальных рецепторов — дискретна. Следовательно, полный комплект всех частот пульсации всех рецепторов сетчатки обоих глаз составляет *первичное представление* оптической информации в зрительной системе. Сразу отметим, что ретинальное представление усложнено характером заселенности сетчатки рецепторами: они плотнее упакованы в центральной ямке, нежели на периферии (см. рис. 1.3.12) и, вдобавок к этому, у четырех различных видов фоторецепторов (три типа колбочек и палочки) разное пространственное распределение по сетчатке (см. раздел 1.3).



**Рис. 2.4.1** Четыре этапа зрительного восприятия. Зрительную обработку, следующую за ретинальной стимуляцией, можно разделить на четыре основных этапа — *оптический, фасетный, предметный и категорийный*.

В формальной и вычислительной теориях зрения сетчаточное представление всегда упрощенно и подается как гомогенная двухмерная матрица сенсоров. Пространственное распределение сих идеализированных рецепторов можно однозначно идентифицировать по их координатам  $(x, y)$  на плоскости с центром в середине фовеа и осями  $x$  и  $y$ , расположенными горизонтально и вертикально в пространстве сетчатки.<sup>1</sup>

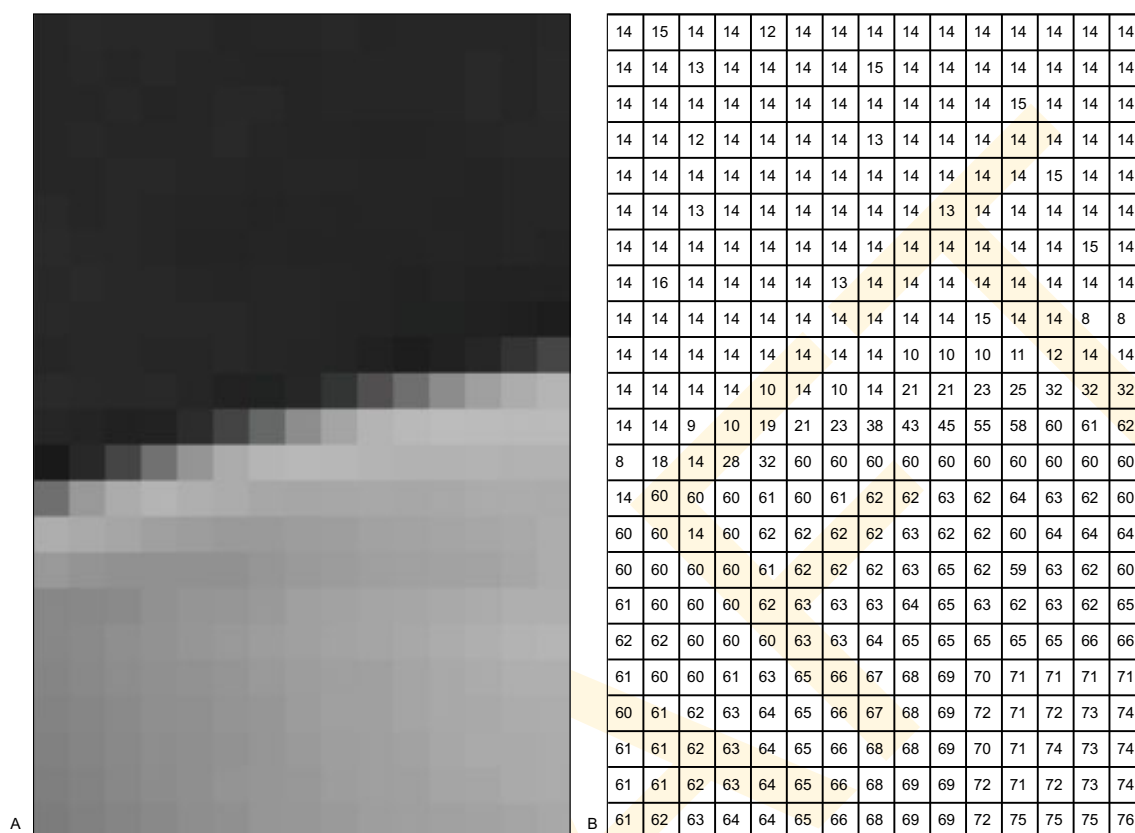
Квадратный *пиксел* как мельчайший элемент изображения, мы рассматриваем в качестве примитивной, невидимой, эксплицитно представленной единицы информации во входящем изображении. Показатель, относимый к пикселу — это фотометрическая яркость в данной локале сетчатки:  $I(x, y)$ .<sup>2</sup> Пиксельная структура маленького фрагмента изображения сцены рис. 2.4.2 продемонстрирована на рис. 2.4.3 А — показан увеличенный фрагмент изображения рис. 2.4.2, отмеченный белым квадратиком. На рис. 2.4.3 В дано числовое представление фотометрической яркости каждого пиксела в диапазоне от 8 до 76. Безотносительно к тому, делались ли описанные выше упрощающие допущения, предполагается, что координатная система ретинального изображения эксплицитно связана с



**Рис. 2.4.2** Изображение простейшей сцены: мы без всякого труда воспринимаем керамическую чашку, стоящую на столе, но при этом всё, что доступно зрительной системе — это световой массив, интенсивности которого плавно меняются в пространстве.

<sup>1</sup> В прежних теориях зрения информация о пространственном расположении рецептора именовалась «локалью».

<sup>2</sup> Для удобства хроматический компонент не учитывается.



**Рис. 2.4.3** Фрагмент изображения чашки (А) и соответствующий набор чисел, отражающий фотометрические яркости по каждому пикселу этого изображения (В).

внутренней структурой сетчатки. Центр ретинальной системы координат привязан к середине центральной ямки, а ее оси, как уже было сказано выше, располагаются по ретинально заданным горизонталям и вертикалям. Позиции рецепторов определяются согласно предложенной системе отсчета.

Чтобы оценить то, насколько сложна проблема интерпретации сырых данных на выходе матрицы сенсоров (рецепторов), вновь обратимся к рис. 2.4.3 В: мы видим, что интенсивность света, воздействующего на каждый рецептор представлена цифрами, и что наша зрительная система никак не в состоянии интерпретировать эти цифры в края, поверхности, предметы и т.д. Эти данные кажутся совершенно бесполезными и неинтерпретируемыми. Но все наоборот, когда мы глядим на рис. 2.4.2, на котором дано изображение — источник этих данных: мы моментально воспринимаем края, участки, поверхности и предметы, все значимые объекты сцены, которых мы не обнаруживали в цифровом их представлении.

Трудность в интерпретации цифрового кода возникает из-за того, что наша зрительная система очень тонко настроена, или как теперь говорят «заточена», на обработку информации, закодированной в яркостях, но не в цифрах. Следовательно, все наши процессы детекции краев, поиска участков, локализации поверхностей и интерпретации объектов были совершенно бесполезны (за исключением чтения цифр как таковых) и мы пытались отыскать смысл в каждом фрагменте информации, тщетно

надеясь найти хоть намек на то, что же изображено. Когда мы озадачены поиском ответа на вопрос «Какие информационные процессы требуются для возникновения зрительного восприятия?», мы всегда должны помнить о том, как нам было трудно интерпретировать числовое изображение и о том, в чем все теории зрительного восприятия единодушны: видимые объекты трехмерного мира — это базис двухмерного массива чисел. И да будет так.

### 2.4.2 Оптический этап зрительного восприятия

Сегодня большинство исследователей зрения согласны с тем, что исходная регистрация изображений сетчаткой обоих глаз — это представление, основанное не только на двухмерной ретинальной организации, но и на неких дополнительных представлениях и процессах. Эти дополнительные представления и процессы мы отнесем к *оптическому этапу*, включающему в себя обработку оптической информации, а также детекцию краев и линий, соединение их в общие контуры, сличение соответствующих изображений на сетчатке левого и правого глаз, определение двухмерных участков в изображении и детекцию прочих оптических элементов, таких как концы линий и «пятна». Сии двухмерные признаки характеризуют структуру изображений и их организацию *до того*, как зрение выполнит интерпретацию этих изображений в модель трехмерной сцены.

К примеру, рис. 2.4.4 А демонстрирует местоположение отдельных краев, составляющих часть оптического представления изображения чашки на рис. 2.4.2. Отметим, что фотометрическая яркость краев, которые мы видим на рис. 2.4.4 А, не строго та же, что легко читаемые контуры на рис. 2.4.4 В: многие из яркостных границ, представленных в части «А», игнорируются зрением — либо потому, что они нечетки, либо потому, что образованы границей светов и теней, но не собственно краем поверхности. Не менее интересен тот факт, что некоторые из наиболее заметных краев, которые каждый из нас воспринимает в этом изображении, фактически отсутствуют в карте краев рис. 2.4.4 А. Сей простейший пример — эффектнейшее доказательство того, что набор яркостных краев в изображении (рис. 2.4.4 А) — это не то же самое, что воспринимаемые контуры предметов (рис. 2.4.4 В).

Марр (1982) показал, что оптические представления — суть т.н. *опорные скетчи*, и, что их два: первый — это *простой опорный скетч*,<sup>1</sup> являющийся результатом процессов элементарной детекции краев, штрихов, точек и концов (линий); второй — это *полный опорный скетч*, образующий т.н. *глобальные группы*, т.е. организующий связи между элементами простого опорного скетча.

Марр утверждал также, что процессы оптического этапа могут оказаться некорректными — действительно, в конце 4-й главы мы продемонстрируем принципиально иную концепцию — но общая идея построений чего-либо на базе оптических представлений — остается весьма практичной.

Итак, не считая разных тонкостей и деталей, в целом оптические представления характеризуются тем, что:

— состоят из т.н. *оптических элементов* — *примитивов*, несущих информацию о двухмерной структуре яркостного изображения, т.е. о краях и линиях, образованных отличием в фотометрической яркости соседних областей, но не о физических

<sup>1</sup> Скетч (англ.) — эскиз, абрис, набросок. — Прим. пер.



**Рис. 2.4.4** Края в изображении чашки. Яркостные края можно вполне успешно регистрировать с помощью компьютерных алгоритмов яркостной краевой детекции (А). Однако контуры предмета, идентифицируемые зрительной системой, это не просто линии, прочерченные по явным яркостным границам (В).

объектах внешнего мира, создавших это изображение — краях поверхностей или теневых краях;

— *двухмерны* — геометрия пространственной информации, заключенной в оптических представлениях по сути своей двухмерна, т.е. представлена в аналоговом формате двухмерных массивов;

— базируются на *фовеоцентрической системе координат*, представляющей собой координатную систему внутри которой 2D-элементы расположены относительно сетчатки, то есть оси этой системы привязаны к глазу, но не к телу или сцене.

В 4-й главе мы будем подробно говорить о том, как образуются оптические представления.

### 2.4.3 Фасетный этап зрительного восприятия

На втором этапе зрительной обработки, который мы назовем «фасетным», происходит восстановление свойств поверхностей объектов внешнего мира, давших те признаки, что были выявлены на этапе оптическом. Принципиальное отличие фасетного этапа от предыдущего — оптического — состоит в том, что фасетный этап представляет информацию о внешнем мире в логике расположения видимых поверхностей в трех пространственных измерениях, тогда как на оптическом этапе зрение оперирует информацией, извлеченной из двухмерного светового паттерна, воздействовавшего на сетчатку.

Тезис о том, что одной из основных задач зрительной системы является восприятие т.н. *планировки поверхностей* — пространственного распределения видимых поверхностей в трехмерном мире — впервые была озвучена и горячо отстаивалась Гибсоном



(1950). В то время почти все остальные исследователи говорили о восприятии трехмерных объектов и только Гибсон заявил, что восприятие поверхностей — биологически важнее. Однако мысль ученого не находила широкой поддержки до тех пор пока, много позже, исследователи компьютерного зрения не встали на защиту позиции Гибсона (Marr, 1978; Marr & Nishihara, 1978). Непопулярность гибсоновской теории частично объясняется тем, что ученый никогда не давал ни определения планировки поверхностей, ни набора процессов, которые могли бы вывести такую планировку из ретинальных изображений. Сие неудивительно, потому что Гибсон не принимал информационной позиции в восприятии и слышать ничего не хотел ни о представлениях, ни об их обработке. Но, вместе с тем, он лучше чем кто бы то ни было понимал, что восприятие поверхностей критично важный аспект зрения.

Концепция эксплицитного представления поверхностей как промежуточного этапа в зрительном акте, стала популярной, когда исследователи компьютерного зрения сформулировали ее количественно и интегрировали в работу компьютерных имитаторов. Марр (1978), а также Барроу и Тенненбаум (1978) стали говорить о поверхностных представлениях примерно в одно время и сошлись в мысли, что существуют скетч-алгоритмы, способные построить эти представления из фактических яркостно-модулированных изображений на сетчатке. Марр (1978) назвал свои фасетные представления *2.5D-скетчами*, дабы подчеркнуть тот факт, что они лежат где-то между собственно 2D структурой оптических представлений и собственно 3D-структурой представлений предметных (см. ниже). Барроу и Тенненбаум (1978) назвали свои поверхностные представления «внутренними изображениями», дабы подчеркнуть, что они отражают внутренние свойства поверхностей внешнего мира, нежели свойства входного изображения на сетчатке.

Создание фасетного представления — это первый шаг на пути восстановления третьей пространственной размерности из двумерных изображений. Фасетное представление не несет в себе информации обо всех поверхностях окружающего мира, но только сцены, то есть тех поверхностях, что видимы из данной точки наблюдения. Как мы покажем в гл. 5, данные об удалении поверхностей от наблюдателя и их уклоне невозможно вывести из ретинальных оптических изображений без дополнительных версий, поскольку выполнение сего — суть проблема инверсии. Однако дополнительных версий, что требуются для выявления свойств видимых поверхностей, относительно мало и они почти всегда верные (особенно в сравнении с теми, что требуются для выявления свойств тех поверхностей, что не видны).

Поскольку фасетное представление включает в себя не только видимые участки поверхностей, его можно уподобить единичному, исключительно гибкому резиновому листу, который плотно «сел» на те поверхности, что отражают свет в глаза наблюдателя. Большинство сегодняшних теорий зрения толкуют поверхности в этом представлении как собранные из множества мелких, локально плоских кусочков. Сие возможно потому что даже сильно искривленная поверхность состоит из более или менее плоских мелких фрагментов, подобно тому как поверхность Земли выглядит ровной с высоты человеческого роста. Такое упрощение позволяет дать полное фасетное представление с помощью информации об окраске, уклоне и удалении от точки наблюдения по каждому локальному фрагменту поверхности в каждом из направлений, радиально расходящихся из точки наблюдения.

Рис. 2.4.5 проиллюстрирует фасетное представление керамической чашки из рис. 2.4.2: кружками обозначены локальные фрагменты поверхности, а перпендику-

лярными им отрезками (подобно воткнутым иголкам) — конкретное местоположение этих фрагментов и их уклон.

Итак, несмотря на то, что марровская идея 2.5D-скетчей представляется несколько сомнительной, следует сказать что фасетные представления характеризуются тем, что:

- состоят из *фасетных элементов*
- примитивов фасетного представления, соответствующих локальным участкам 2D-поверхностей того или иного уклона и того или иного удаления от наблюдателя в 3D-пространстве. Каждый фасетный элемент можно дополнительно охарактеризовать по его окраске и форме.

- *трехмерны* — несмотря на то, что поверхности сами по себе двухмерны, расположены они в трехмерном пространстве.

- базируются на *антропоцентрической системе координат* — координатной системе, в которой позиции поверхностей представлены в показателях направления и удаленности от точки наблюдения и не привязаны к сетчатке.

На схеме рис. 2.4.6 показано, что фасетное представление строится из нескольких компонент: из *стереоданных* (легком отличии в положении одних и тех же объектов в ретинальном изображении левого и правого глаз), из *параллакса движения* (отличии в скорости движения точки на разных ее удалениях от наблюдателя), из полутонов и теней, а также прочих пикториальных свойств поверхностей, таких как текстура, раз-

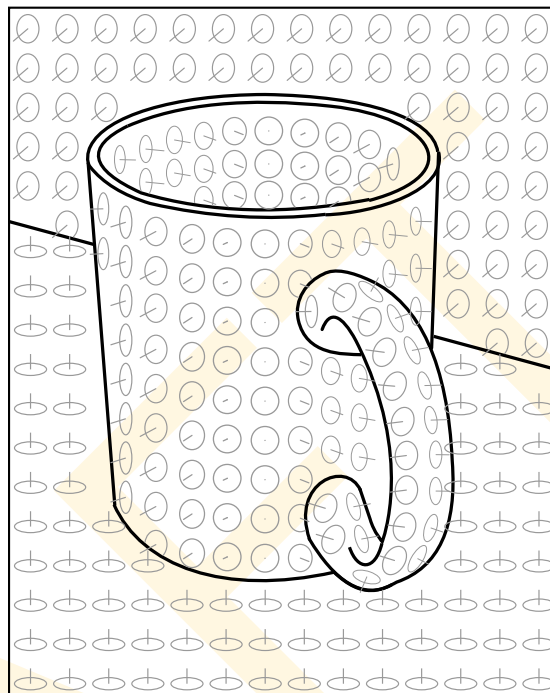
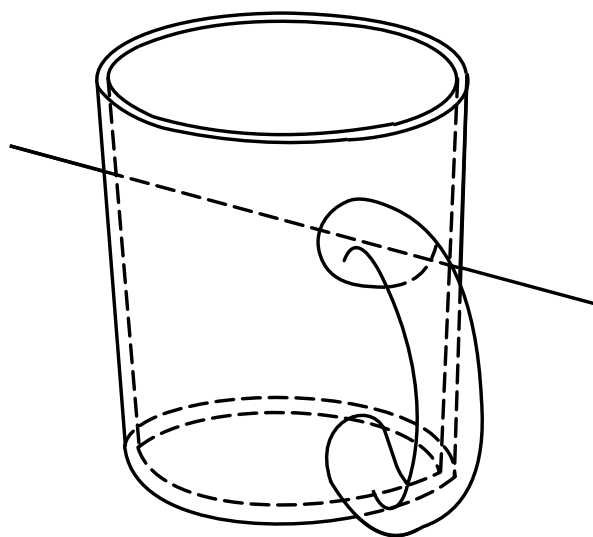


Рис. 2.4.5 Фасетное представление чашки: поверхности, видимые на рис. 2.4.2 представлены в виде набора определенным образом пространственно ориентированных круглых фасеток (наклоненных и повернутых) с перпендикулярными им отрезками, указывающими на пространственную позицию каждой фасетки в изображении.



Рис. 2.4.6 Блок-схема, демонстрирующая то, как фасетное представление можно получить из представления оптического. Источники информации об удалении и ориентации поверхностей детально обсуждаются в гл. 5.



**Рис. 2.4.7** Пример предметного представления: для демонстрации того, что человек обычно воспринимает подобные сцены как собранные из 3D-составляющих, пунктиром показаны скрытые края чашки рис. 2.4.2.

включающие в себя скрытые поверхности сцены. Сие значит, что на данном этапе восприятия, который мы назовем *предметным*, зрительные представления содержат точную трехмерную информацию.

Реализация предметного этапа требует от зрительной системы дополнительных версий об объектах сцены, поскольку теперь инференции несут в себе информацию о невидимых поверхностях, либо о невидимых фрагментах видимых поверхностей. Мы назвали сей этап «предметным», поскольку на нем зрительная система оперирует данными о невидимых поверхностях, и подразумевали при этом, что в деле участвуют эксплицитные представления цельных предметов сцены. На рис. 2.4.7 пунктиром показаны скрытые края, которые каждый с легкостью различает на рис. 2.4.2. Край стола закрыт чашкой, задние и внутренние поверхности, а также низ чашки прикрыты теми ее фрагментами, которые мы видим непосредственно.

Итак:

⇒ восстановление 3D-структуры предметов сцены — это цель и смысл информационной обработки на предметном этапе.

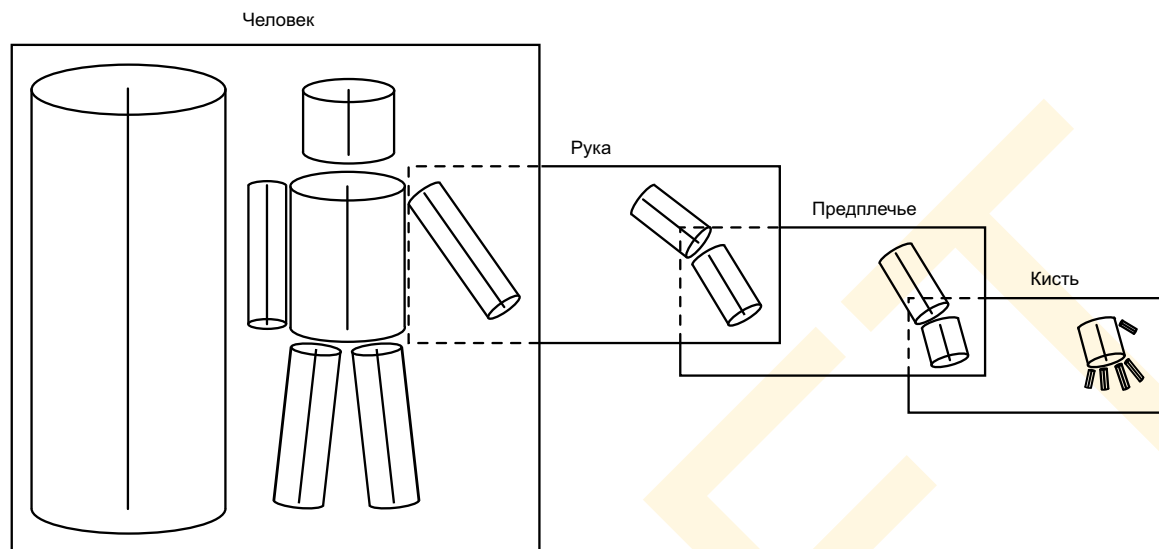
Существует как минимум два пути, по которым может строиться предметное представление.

Первый путь — это простое расширение фасетного представления до уровня информации о скрытых поверхностях, расположенных в полноценном трехмерном пространстве. Сие можно назвать *граничным подходом* к предметному представлению.

мер, форма и наложения. Мы будем подробно говорить об этих факторах в 5-й главе, посвященной тому, из чего и как строятся фасетные представления.

#### 2.4.4 Предметный этап зрительного восприятия

Очевидно, что зрительное восприятие не заканчивается фасетным представлением, то есть представлением видимых поверхностей. Если бы восприятие исчерпывалось анализом структуры поверхностей, мы не удивились бы тому, что у чашки (рис. 2.4.2) вместо задней поверхности — пустое место или что поверхность эта радикально отлична по форме от гладкого цилиндра (который напрашивается сам собой). Как было показано в 1-й главе, такие «открытия» оказываются весьма и весьма неожиданными. Коль скоро зрительная система строит версии того, как по ее «мнению» должны выглядеть частично и полностью скрытые поверхности, то очевидно, что существуют трехмерные представления,



**Рис. 2.4.8** Волюметрические примитивы в предметном представлении. Тело человека как трехмерный объект грубо представлено набором цилиндров соответствующего размера, пропорций, ориентации и соединений. Каждая секция демонстрирует объект на грубом, общем уровне как единичный цилиндр, но по отношению к предыдущей — на более тонком, локальном уровне как конфигурацию нескольких цилиндров. (Marr & Nishihara, 1978).

Второй — это толкование предметов как трехмерных сущностей, представленных в виде некоего набора примитивных 3D-форм. Сие можно назвать *волюметрическим подходом*, поскольку он эксплицитно представляет объекты в виде совокупности трехмерных фигур специфической формы.

На рис. 2.4.8 показано, как человеческое тело можно разложить на составляющие, каждая из которых представлена элементарной цилиндрической фигурой. Выдающиеся труды Эйгина и Бинфорда (1976), а также Марра и Нишихары (1978), посвященные 3D-элементам в контексте компьютерного зрения, на многие годы утвердили волюметрический подход в качестве доминирующей теории предметной обработки. Однако возможно, что некоторое заполнение закрытых поверхностей имеет место и до образования волюметрического представления. В главах 6, 7 и 8 мы будем говорить о том, откуда и как возникают предметные представления.

Еще раз напомним, что очень важно отделять марровское толкование предметного представления, данное им в логике цилиндрических элементов, от общей теоретической концепции волюметрического описания. Последнее даже в общем виде далеко от ясности. В добавок к вопросу о том, что является исходным элементом — поверхность или фигура, активно обсуждается система отсчета и геометрия.

Однако в целом, предметное представление характеризуется тем, что:

- состоит из *волюметрических элементов* — примитивов предметного представления, которые можно описать подлинно трехмерными фигурами и за счет подключения информации о невидимых поверхностях предмета.

- *трехмерно* — пространство, в котором расположены волюметрические элементы, — трехмерно.

— базируются на *предметной системе координат* — координатной системе, в которой пространственные отношения между волюметрическими элементами можно определить в логике внутренней структуры самих фигур.<sup>1</sup>

### 2.4.5 Категорийный этап зрительного восприятия

В гл. 1 мы говорили о том, что главной целью восприятия, в частности зрительного, является получение особую точной информации о внешнем мире, что необходимо для выживания и размножения. Сие однозначно предполагает, что финальный этап восприятия призван идентифицировать функциональные свойства наблюдаемых объектов, то есть их способность или неспособность вызывать доверие к себе, рождать желания, формировать цели и мотивы поведения. Мы назовем сию обработку визуальной информации *категорийным этапом* восприятия, поскольку принято считать, что функциональные свойства предметов внешнего мира доступны особи через процесс *категоризации*. К примеру, восприятие содержимого рис. 2.4.2 как фотографии чашки есть результат некоей категорийной обработки. Какой?

#### 2.4.5.1 Категоризация

Категоризация при восприятии (или *распознавание образов*) — это эволюционно значимая функция, состоящая из двух этапов:

— на первом происходит классификация зрительной системой объекта как представителя одной из большого числа известных категорий (исходя при этом из таких его визуальных свойств как форма, размер, окраска и местоположение);

— на втором этапе выполненная идентификация открывает доступ к хранилищу информации об объектах данного типа, включая их функцию и различные виды предположений об их возможном поведении в дальнейшем. К примеру, известно, что предмет, изображенный на рис. 2.4.2 подходит для того, чтобы налить в него некую жидкость, а потом выпить ее.

Сия двухэтапная схема хороша тем, что любое функциональное свойство можно ассоциировать с любыми объектом, поскольку отношения между формой объекта и хранимой информацией о его функции, происхождении и применении могут оказаться совершенно произвольными и следствием процесса категоризации.

#### 2.4.5.2 Аффорданс

Существует принципиально иной путь, по которому зрительная система распознает функцию предмета и который реализуется за счет регистрации функциональных свойств объектов более или менее непосредственно из их визуальных характеристик и без предварительной категоризации. Гештальтисты первыми сформулировали эту версию, назвав ее *физиогномической*. Коффка высказался об этом так: «первобытному человеку вещи сами говорили о том, что он хочет с ними делать: яблоко сказало: “ешь меня”; вода сказала “пей меня”; гром сказал: “бойся меня” и женщина сказала: “люби меня”» (Koffka, 1935). Спустя четыре десятилетия Гибсон (1979) провозгласил аналогичный подход, распространив свою идею прямого восприятия и на функции предметов внешнего мира. Он назвал зрительные функции объекта его *аффордансами*.<sup>2</sup>

<sup>1</sup> Концепция внутренних предметных координатных систем весьма сложна и детально будет рассмотрена нами в 8-й главе.



Примером аффорданса может служить то, как стул «приглашает» сесть на него. Согласно такой точке зрения мы, прежде чем сесть на стул, не выполняем классификации этого предмета, как относящегося к категории «стул», поскольку его аффорданс воспринимаем непосредственно и без какой-либо категоризации.

Действительно, сие возможно, поскольку восприятие использует как прямые, так и не прямые процессы. Функциональные свойства некоторых предметов, таких как стул или чашка, тесно связаны с их видимой структурой, отчего у нас не возникает необходимости категоризировать эти предметы, дабы узнать, как ими пользоваться. Иные предметы, такие, скажем, как компьютеры и телефоны, обладают функциями, которые не читаются непосредственно из их визуальных характеристик и поэтому целиком требуют предварительной категоризации. Мера применения нами каждой из описанных тактик сегодня неизвестна. Подробнее мы будем говорить об этом и некоторых других вопросах в 9 главе.

### 2.4.6 Заключение

Итак, четыре этапа зрительной обработки информации — оптический, фасетный, предметный и категорийный — представляют современный взгляд на общую структуру зрительного восприятия. Мы расположили их в функциональной последовательности, но при этом не сказали, что каждый предыдущий этап должен непременно закончиться, прежде чем начнется следующий: стрелки, направленные назад на рис. 2.4.1, указывают на то, что последующие процессы оказывают определенное влияние на процессы предыдущие.

В предстоящих главах — с целью структурирования нашего разговора по главным темам: цветовому восприятию (гл. 3), восприятию пространственных свойств предметов (гл. 4-9) и восприятию движения (гл. 10) — мы будем активно пользоваться описанной нами четырехэтапной системой. В последних трех главах мы рассмотрим три смежных вопроса, окружающих ядро зрительного восприятия: зрительную фиксацию, реализуемую за счет глазных движений и внимания (гл. 11), зрительную память, воображение (гл. 12) и зрительную осведомленность (гл. 13).

---

<sup>2</sup> *Аффорданс* (англ. *affordance* от *afford* — иметь или предоставлять возможность) — в теории восприятия Дж. Гибсона, «приглашающее» (манящее) качество воспринимаемых предметов и событий. Оригинальность этого понятия иногда сильно преувеличивается, поскольку значительно раньше К. Левин глубоко и детально разработал идею о *Aufforderungscharakter* — о побудительном (побудительном) характере вещей. — *Прим. пер.*

DRAFT

Одним из самых загадочных и обособленных аспектов зрения является цветовое восприятие окружающего мира, вызывающее у индивида различные ощущения, качественно отличные от любых других ощущений и недоступных ни в одной другой сенсорной модальности, то есть: физические объекты и источники света мы воспринимаем в той или иной мере *окрашенными*. При этом совершенно неочевидно, что так должно быть. Но оно так, и большинство из нас совершенно уверены в том, что предметы выглядят окрашенными, потому что таково их физическое свойство: небо синее — потому что синее, трава зеленая — потому что зеленая, а кровь красная — потому что красная.

Как мы покажем далее, сия уверенность в корне ошибочна: любой предмет «окрашен» в зависимости от того при каких обстоятельствах мы наблюдаем его, то есть:

⇒ Цвет — это *физиологическое свойство* зрительного восприятия предметов и источников света, но отнюдь не физический атрибут этих предметов и источников.

Безусловно, цветовое ощущение зависит от физических свойств предметов и источников, но свойства эти принципиально отличны от цветовых ощущений, возникающих у наблюдателя. Точнее всего цвет толкуем как результат сложных взаимодействий между световыми потоками внешней среды и нервной системой человека, и для понимания сути и характера этих взаимодействий нам потребуется долгий и серьезный разговор. Он будет очень непрост, потому что вопросы цветового восприятия мира лежат где-то между хорошо изученными аспектами науки о зрении (да и всей когнитологии в целом). Однако сегодня мы уже многое знаем о том, как физика световых потоков влияет на цветовое восприятие этих потоков, как ретинальные рецепторы отвечают на эти потоки, и как представители разных культур именуют цветовые ощущения, возникающие при рассматривании поверхностей. Разумеется, мы не сказали, что знаем о цветовом восприятии всё — отнюдь, но при этом все же мы можем утверждать, что знаем в этой сфере больше, чем в прочих аспектах науки о зрении и поэтому будем использовать это знание в качестве модели применения междисциплинарного подхода к изучению зрения в целом.

Итак, принято считать, что высокоразвитое цветовое зрение — это эволюционное достижение наших предков и их преимущество перед другими видами животных. И хотя с точки зрения современной науки, сие не критично в плане выживания, тем не менее колоссально обогащает зрительную картину окружающего мира. Некоторые исследователи полагают также, что высокоразвитое цветовое зрение — это эволюционное достижение обезьян, позволившее им находить и различать тропические фрукты и съедобную зелень (Osorio & Vorobyev, 1996).

Важнейшие понятийные составляющие сущности цвета пришли к нам из физики, психологии, физиологии, компьютерной науки, лингвистики, генетики и антропологии. Коль скоро цветовое восприятие и именование цветов дают добротную модель понимания зрения в когнитивном ключе, то очевидно, что их изучение крайне необходимо.

Рассказ о том, как именно исследователи пришли к пониманию механизмов цветового зрения, разумно построить в логике исторической ретроспекции, поскольку сие

лучше всего проиллюстрирует то, как возникла и сформировалась более или менее стройная теория цветового восприятия.

Вначале о тех фактах, что требовали объяснения. Сперва предложена некая теория, объясняющая эти факты, а затем — альтернативная теория, объясняющая прочие; между представителями обеих разгорелась весьма бурная дискуссия, но спортаки разрешился и разрешился путем т.н. *теоретической интеграции*, в результате чего все последующие эксперименты стали проводиться в рамках новой объединенной концепции.

Текущую главу мы построим согласно метатеоретическим принципам, изложенным во 2-й главе и начнем с анализа количественных аспектов цветового восприятия: во-первых, мы рассмотрим природу и характер входной информации, которую несет свет (физика); затем обратимся к выходной информации, которая суть цветовое ощущение (психология); на следующем этапе разберем взаимоотношения между первым и вторым, то есть т.н. *мэппинг* физических данных (свет) на данные психологические (цветовое ощущение); после чего перейдем на алгоритмический уровень и ответим на вопрос: «Каким образом мэппинг физических характеристик излучения на цветовое ощущение можно истолковать в контексте информации и операций с ней?»; наконец мы определим три главных этапа цветовой обработки, которые обсудим в оставшейся части главы: *оптический этап*, *фасетный* и *категорийный*. Сразу отметим, что в обсуждении четвертого — предметного этапа — нет необходимости, поскольку цвет — суть свойство зрительных перцептов поверхностей, но не волуметрических компонент перцептов объекта.

В процессе разбора этапов мы проанализируем алгоритмы цветовой обработки с позиции рекурсивной декомпозиции (см. раздел 2.3.3.2), т.е. пытаюсь отыскать низкоразрядные описания процессов. В некоторых случаях мы достигнем той точки, с которой четко видно, как данный функциональный анализ соотносится с исполнением алгоритма — подлежащей физиологией цветовой обработки в нейронных структурах глаза и мозга (в целом сие касается оптической обработки, описанной в разделе 3.2). Однако к тому моменту, как в разделах 3.3 и 3.4 мы обсудим фасетную и категориальную обработку, наше понимание работы нервного компонента зрительной системы по-прежнему останется недостаточным для установки связи с более абстрактным анализом визуальной информации: установка такой связи требует намного более глубокого исследования.

### 3.1 КОЛИЧЕСТВЕННЫЙ ЭТАП ЦВЕТОВОГО ВОСПРИЯТИЯ

Дэвид Марр (1982) предложил начинать разбор устройства и работы любой сложной информационной системы с количественного уровня, описывающего вход и выход, а также принципы, положенные в основу мэппинга входных данных на выходные. В контексте разговора о цветовом восприятии на входе имеем свет, отраженный от поверхностей и попавший в глаз наблюдателя (рассматривающего эти поверхности), на выходе — цветовое ощущение. Мэппинг же — суть психофизическое соответствие второго первому. В текущем разделе мы дадим упрощенное количественное описание данных аспектов цветового восприятия, но позднее доберемся до его алгоритмических и имплементарных аспектов.

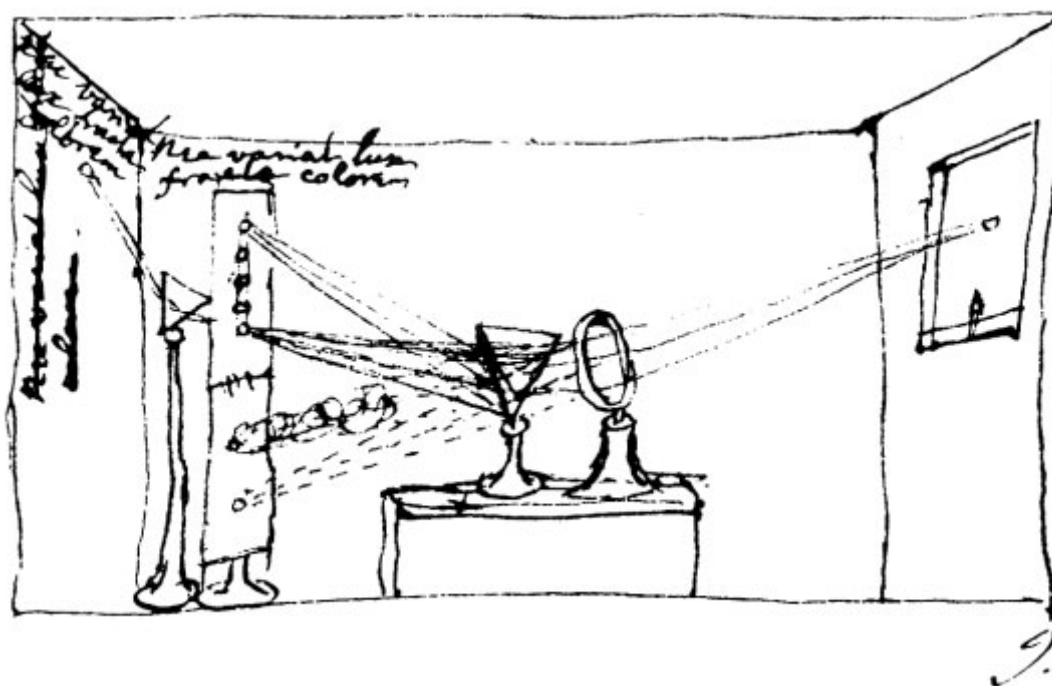


Рис. 3.1.1 Рисунок Ньютона, иллюстрирующий его опыт по дифракции дневного света. С помощью призмы ученый разложил солнечный свет на составляющие. Сведя их воедино второй призмой, он показала, что декомпозиция обратима.

### 3.1.1 Физическое описание света

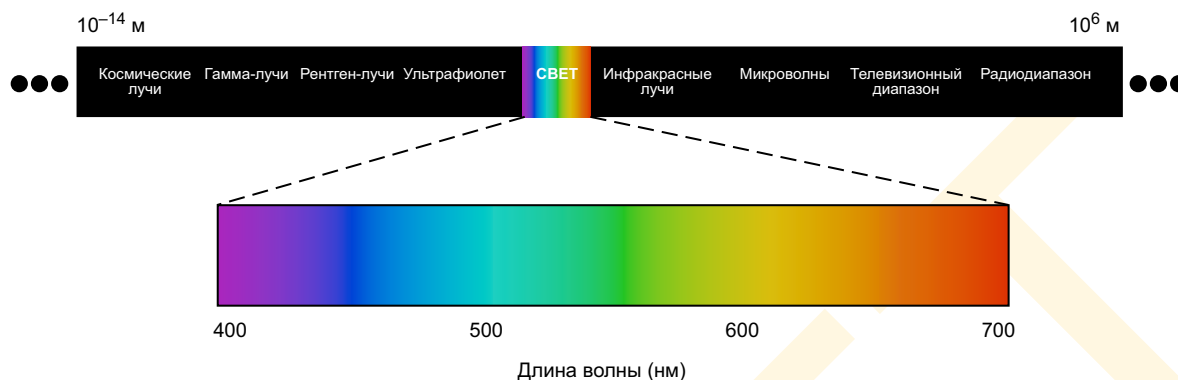
Первый серьезный шаг к пониманию цветового зрения был сделан в 1666 г. сэром Исааком Ньютоном, обнаружившим, что «белый» (неокрашенный) солнечный свет состоит из множества компонент, каждая из которых вызывает самостоятельное цветовое ощущение. Ньютон проделал крошечную дырку в ставне окна и на пути светового потока установил трехгранную призму. Когда луч прошедший через призму, упал на лист бумаги, ученый увидел радугу, т.е. набор цветовых полос (рис. 3.1.1). Тогда ученый добавил на пути света собирающие призмы и наблюдал то, что сборка оказалась вновь белой. Теоретическим обоснованием опыта (и ряда других) стало то, что солнечный свет, по мнению ученого, фактически «собран» из множества отдельных «цветов». Ньютон понял также, что «цвета» — это не собственно свет, но результат *воздействия* этого света на зрение, о чем красноречиво свидетельствует цитата:

«Сами по себе лучи отнюдь не окрашены. В них нет собственно ничего, кроме энергии и характера, возбуждающих то или иное цветное ощущение... Словом, у предметов нет цвета, но лишь способность отражать одни лучи больше других».

Работа Ньютона — это одно из первых фундаментальных открытий в сфере изучения физических свойств света, лежащих в основе цветового зрения, но с тех пор, конечно же, наши знания стали намного глубже. Как мы уже говорили в первой главе, свет состоит из сверхмелких неделимых частиц — *фотонов*, обладающих как свойствами частиц, так и волн. В большинстве аспектов науки о зрении корпускулярные свойства фотонов важнее волновых, но в контексте цветового зрения их волновое поведение представляет наипервейший интерес.

Итак, фотон — это сверхмелкий пакет пульсирующей электромагнитной энергии,





**Рис. 3.1.2** Электромагнитный спектр. Энергия излучения характеризуется длиной ее волны, плавно меняющейся от очень маленькой до огромной. Видимый спектр занимает весьма скромный диапазон от 400 до 700 нанометров ( $10^{-9}$  м) и это единственный вид электромагнитной энергии, который может восприниматься зрительной системой человека.

характеризуемый одним из волновых свойств, а именно — *длиной волны колебания*. Длины волн фотонов измеряются в единицах расстояния, варьирующих от сверхмалых ( $10^{-14}$  м) до сверхбольших (тысячи км). Длины волн квантов, способных оказать воздействие на зрительную систему человека и вызвать зрительное ощущение — т.е. фотонов — это крошечная порция общего электромагнитного спектра, показанного на рис. 3.1.2. Фактически зрение человека чувствительно к фотонам с длинами волн, лежащими в диапазоне от 400 до 700 нм ( $1 \text{ нм} = 10^{-9} \text{ м}$ ). Рентгеновские лучи, СВЧ-лучи, радио- и телеволны также занимают определенные участки спектра, но наши органы чувств не дают отклика на них, т.е. на кванты, лежащие вне означенного диапазона. В результате мы не можем воспринимать их без применения специальных сложных устройств — рентгеновских аппаратов, телевизоров и радиоприемников — конвертирующих эти волны в видимый свет или звуковые колебания воздуха.

Вооруженные знанием этих фактов, мы можем дать исчерпывающее физическое описание любого равномерного светового пятна в логике числа фотонов в нем (на единицу времени) по каждой из длин волн в диапазоне от 400 до 700 нм. Сие описание, именуемое *спектральным распределением энергии светового потока* или коротко — *спектром света*, можно представить графически (рис. 3.1.3).

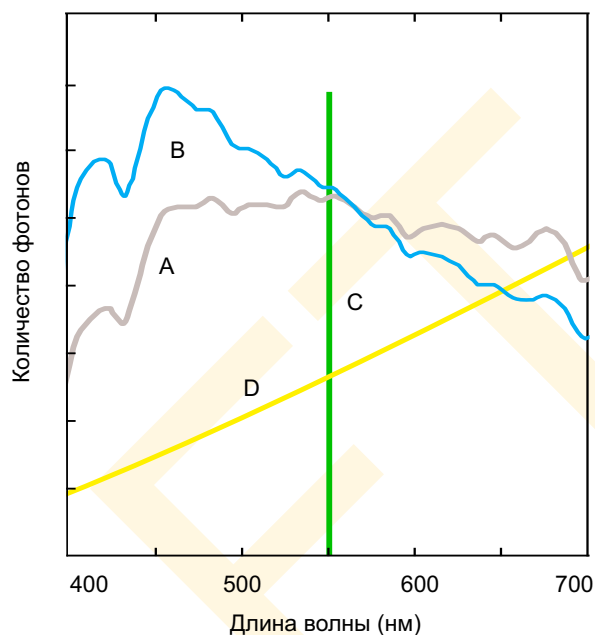
Свет, спектр которого являет собой узкую полоску по одной из длин волн (рис. 3.1.3 С), именуется *монохроматическим* светом (от греч. «mono» — «один» и «chroma» — «окраска»). Все прочие варианты спектрального распределения (к примеру: А, В, D) относятся к *полихроматическим спектрам* (от греч. «поли» — «много»): их распределение таково, что фотоны заполняют весь видимый спектральный диапазон. Естественный (солнечный) свет в грубом приближении несет в себе равное число фотонов по всем длинам волн (рис. 3.1.3 А и В), то есть как и предполагал великий Ньютон.

### 3.1.2 Физиологическое описание цвета

Отметим, что в физическом описании света нигде не встречается термин «цвет». Сие благодаря ньютоновскому «сами по себе лучи отнюдь не окрашены». Напомним, что цвет рождается только тогда, когда свет оказывает воздействие на зрительную сис-

тему наблюдателя (необязательно человека), нервная система которого в состоянии воспринимать этот свет окрашенным. Ситуация напоминает старую загадку о том, возникает ли звук, когда в пустом лесу падает дерево? Излучение любых длин волн существует самостоятельно, т.е. независимо от наблюдателя, но цвета вне наблюдателя не существует, поскольку цвет — суть физиологический феномен, рождающийся только внутри живого существа.

Несмотря на то, что порой на наше зрение воздействует свет, *излучаемый* поверхностями, скажем экраном телевизора или компьютера, обычно мы интерпретируем цветное ощущение как свойство отражающих поверхностей. Так происходит от того, что свет, испущенный источником (солнцем или электролампой), попадает в наши глаза *отраженным* от различных предметов (см. раздел 1.2). Это тот самый отраженный свет, что позволяет нам видеть несамосветящиеся поверхности и воспринимать их свойства. Поверхности эти, из-за того, что по разным длинам волн отражают свет в разных пропорциях, воспринимаются по-разному окрашенными.



**Рис. 3.1.3** Спектральные распределения энергии света различных источников. А — смесевой свет от неба и Солнца; В — свет неба (без Солнца); С — монохромат с длиной волны 550 нм; D — лампа накаливания.

### 3.1.2.1 Цветовое пространство

Ощущение цвета поверхности по сути своей принципиально отлично от физики света. Все цветовые ощущения здорового наблюдателя, вызванные светом, отраженным от поверхностей, описываются в логике трех чувственных размерностей: *цветового тона (hue)*, насыщенности (*chroma*) и светлоты (*lightness*). Вместе размерности эти задают базис *цветового пространства*: трехмерной координатной системы, внутри которой любое возможное цветовое ощущение можно представить единичной точкой уникальной позиции. Полный комплект возможных цветовых ощущений человека с нормальным цветовым зрением образует подмножество цветового пространства, именуемое *цветовым телом*, несколько напоминающим искривленную детскую юлу (см. рис. 3.1.4 А). Каждое цветовое ощущение можно однозначно локализовать в определенной точке цветового тела.

Отметим, что переход от физического описания света к психологическому описанию цвета — суть колоссальное упрощение ситуации: физику для характеристики спектра данного светового потока нужно бесконечное число значений по каждой из длин волн в диапазоне от 400 до 700 нм, тогда как психологу лишь три переменных: одна для цветового тона, другая для насыщенности и третья — для светлоты. Иными словами, цветовые ощущения несут в себе гораздо меньше информации о световом

потоке, чем его полный спектр, и сие допускает возможность того, что множество физически отличных друг от друга световых потоков (то есть потоков разного спектра) могут вызывать одно и то же цветовое ощущение. Последнее — суть важнейшее свойство цветового зрения, которое мы обсудим, когда поведем разговор о т.н. *метамерных смесях*.

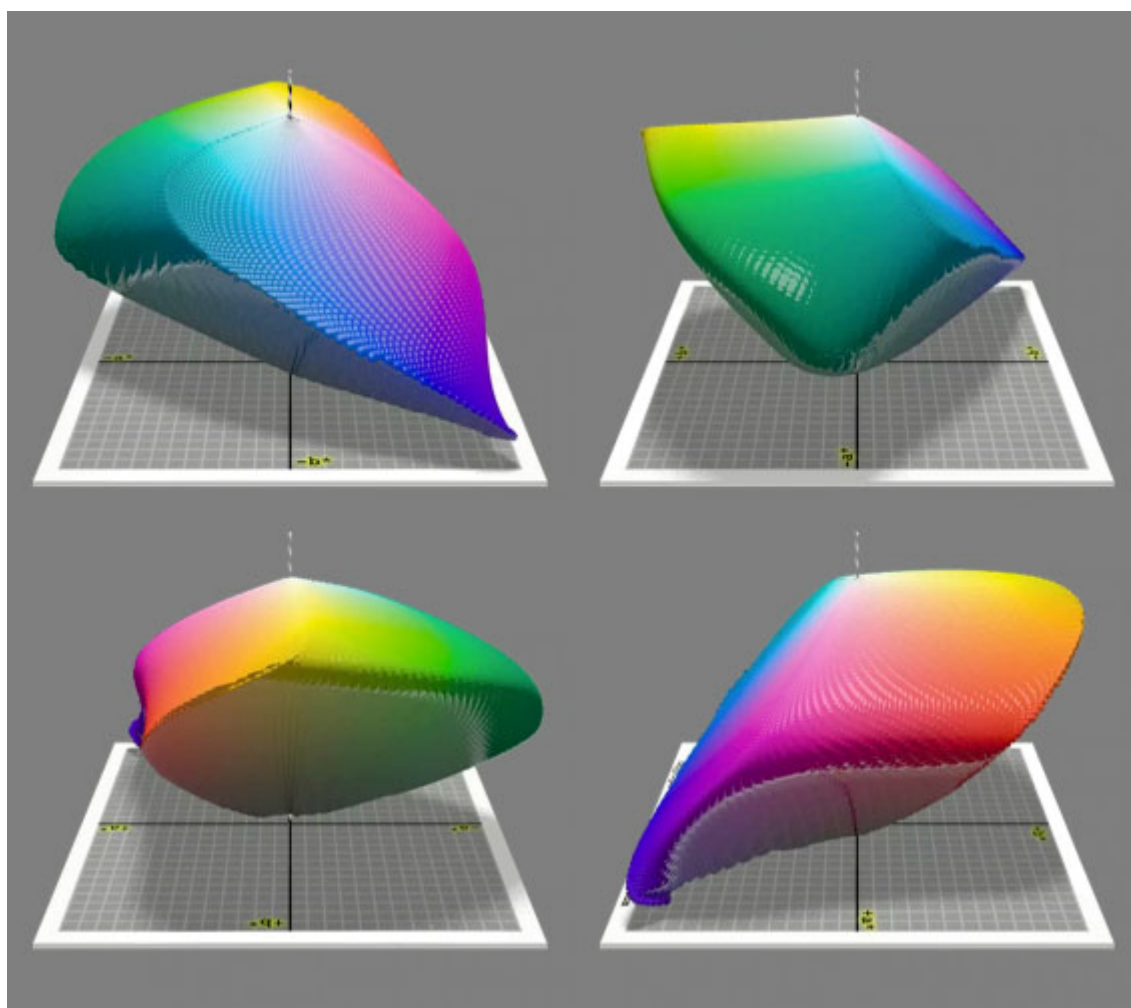
**Манселловская система.** Цветовое пространство, в котором размерности цветового тона, насыщенности и светлоты разбиты на чувственно-равные интервалы, предложил в 1905 г. художник-педагог Альберт Манселл. Стандартизированный атлас цветных образцов, изданный (и издаваемый по сей день) на основе изобретения Манселла, именуется Манселловским атласом цветов и состоит из двух томов: 1600 глянце-вых и 1600 матовых образцов, окраска которых проконтролирована наитщательнейшим образом (рис. 3.1.4 В). Любое цветовое ощущение можно специфицировать, найдя максимально близкий образец в манселловском атласе и зафиксировав значения его координат в манселловской системе.

Условия просмотра манселловского атласа также определены очень жестко: сравнение тестируемой поверхности с образцами атласа должно проводиться при строго регулируемом освещении, а также на равномерном средне-сером фоне, либо через равномерно нейтральную маску. В противном случае отличия в цвете фона и/или составе освещения могут сильно повлиять на результат цветового восприятия тестового образца: вспомним, что многие из нас сильно расстраиваются, когда новенький костюм, который так замечательно смотрелся на витрине, радикально меняется в цвете на дневном свете или при свете ламп накаливания. Обсуждение сих сложностей весьма важно в понимании цветового восприятия, но мы отложим на время их разбор, дабы упростить общую задачу.

### 3.1.2.2 Цветовой тон

Чувственная размерность, которую мы в первую очередь ассоциируем с «цветом» поверхности, именуется *цветовым тоном*. В цветовом пространстве цветовой тон соответствует азимуту, взятому от центральной оси к точке, представляющей данный цвет. Цветовые тона как атрибуты ощущений, вызываемых монохроматическими световыми потоками (т.е. состоящими из фотонов какой-либо одной долины волны), организуются от коротких к длинным волнам точно так же как в радуге: плавно переходят от фиолетового, через синий, зеленый, желтый и оранжевый к красному. То есть, психологическая размерность «цветовой тон» сопряжена с физической размерностью «длина волны». Однако между первым и вторым существуют принципиальные отличия.

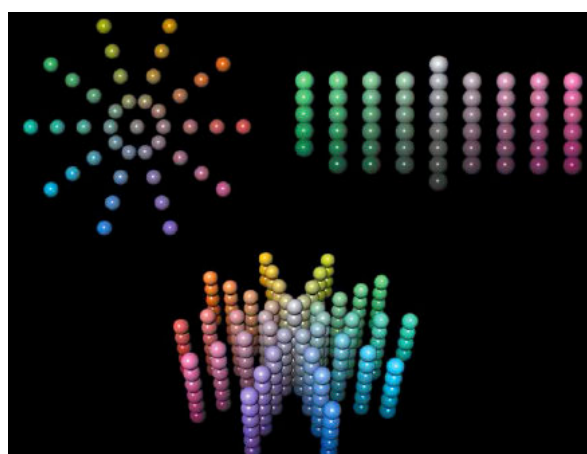
Коль скоро световые потоки организованы согласно их физическому подобию в логике длин волн, то они оказываются лежащими на одной прямой (см. 3.1.2). То есть коротковолновые и длинноволновые потоки схожи друг с другом как минимум по физическим показателям. Но ежели монохроматические световые потоки расположить согласно их чувственному подобию, то они расположатся по периферии т.н. *цветового круга* (см. 3.1.5). Мы видим, что размерность «цветовой тон» — размерность сильно искривленная и ощущения, вызванные ультракоротким светом (воспринимается как фиолетовый) и ультрадлинным (воспринимается как оранжево-красный) весьма близки друг у другу, по общей красноватости. Итак, еще раз повторим, что психологическая размерность цветового восприятия — «цветовой тон» и физическая размерность



А



В



**Рис. 3.1.4** А — Цветовое тело человека с нормальным цветовым зрением внутри цветового пространства CIELAB в нескольких боковых проекциях. В — Манселловский атлас цветов (слева) и графическое трехмерное представление манселловской системы (справа).

монохроматического излучения — «длина волны», хоть и тесно взаимосвязаны, но далеко не одно и то же.

Наконец отметим, что между фиолетовыми тонами и оранжево-красными лежат т.н. *неспектральные цветовые тона*: от пурпурных через малиновые к глубоко красным. «Неспектральными» они названы потому, что в видимом спектре нет монохроматов, способных вызвать ощущения в означенном диапазоне. Пурпурные, малиновые и глубоко красные цветовые тона — плод воздействия *пар* монохроматических световых потоков.

### 3.1.2.3 Насыщенность

Вторая чувственная размерность цвета — *насыщенность*, характеризующая «чистоту» или «силу» цветового ощущения. В цветовом пространстве — это расстояние между центральной осью и точкой представляющей данный цвет. К примеру, все насыщенные цвета, вызванные монохроматическими стимулами (цвета радуги), лежат по наружному краю цветового тела, т.е. на самых высоких величинах насыщенности. Все нейтральные (серые) лежат по центральной оси и их насыщенность равна нулю. Приглушенные, пастельные тона занимают по параметру насыщенности промежуточное положение.

### 3.1.2.4 Светлота

Третья размерность цветовых ощущений, вызываемых потоками, отраженными от несамосветящихся поверхностей — *светлота*. В цветовой координатной системе (цветовом пространстве), светлота — это «высота» позиции цветового ощущения (см. рис. 3.1.4 А). У всех цветовых ощущений, вызванных светом, отраженным от поверхностей, имеется некое значение светлоты, хотя, возможно, наиболее выражен этот компонент у *ахроматических цветов* (серых), т.е. тех, что лежат на центральной оси (с белым наверху и черным внизу).

На рис. 3.1.5 показана проекция цветового тела человека на плоскость, перпендикулярную оси светлоты, что позволит нам лучше понимать то, как соотносятся друг с другом цветовой тон и насыщенность, и как они образуют цветовой «круг». Серый находится в центре, а большинство «ярких» (высоконасыщенных) цветов располагаются по краям.

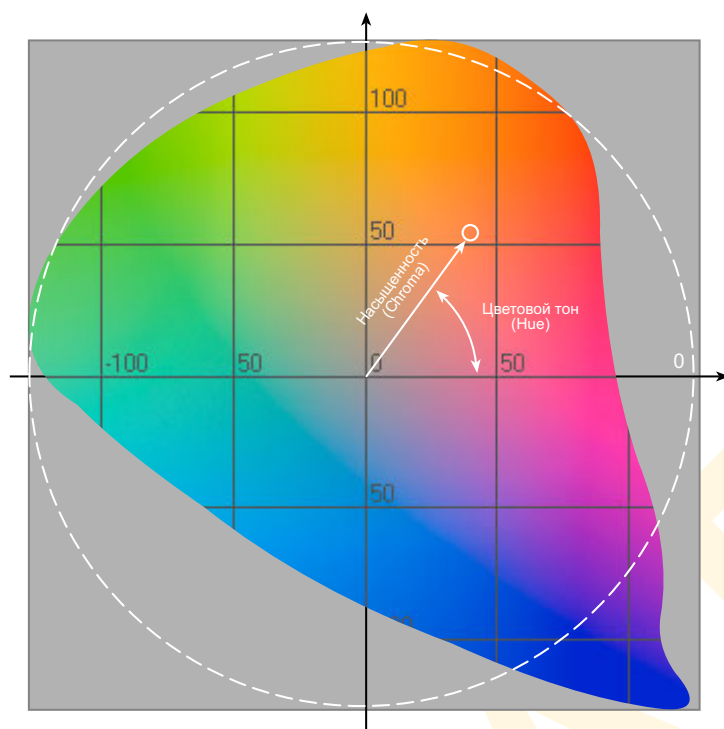
Цветовые тона, расположенные на противоположных сторонах цветового круга, называют *комплементарными*. Если световые потоки, вызывающие ощущения комплементарных тонов, смешать в определенной пропорции, то результирующая смесь окажется ахроматической (неокрашенной) и восприниматься будет как некая градация серого.

Особо отметим, что те цвета, которые мы называем «коричневыми» — это фактически желтые и оранжевые низкой светлоты.

### 3.1.2.5 Светлота и субъективная яркость

Размерность «светлота» относится к цветовому ощущению, возникающему при восприятии отражающих поверхностей. Сие обычно, поскольку лишь немногие объекты, к примеру Солнце, звезды, электролампы, экраны телевизоров и мониторов испускают свет самостоятельно, но не переотражают его. За исключением телевизоров и мониторов, свет, испускаемый его источниками, в целом хроматически нейтрален.





**Рис. 3.1.5** Проекция цветового тела, образующая плоскость цветовых тонов (цветовой «круг») пространства CIELAB(CH). Четко видно, что тона низкой насыщенности расположены ближе к центру; высокой — к периферии. Условность термина «цветовой круг» означена пунктирной линией. Серые поля — проекция неостребованных участков цветового пространства на плоскость цветовых тонов.

В сказанном легко убедиться на примере телевизора: когда он выключен (то есть, когда он лишь отражает свет от других источников, но сам не испускает ничего), он выглядит средне-серым. После включения, однако, мы видим яркие белые и глубокие черные цвета. Что есть белые, понять легко — это участки, испускающие фотоны. Но почему темные участки экрана выглядят теперь темнее? Кажалось бы, такого не должно быть, потому что выключенный телевизор не может излучать, а включенный не может испускать «отрицательные» фотоны, понижающие количество света. Следова-

В анализе цвета самосветящихся объектов светлоту заменяет размерность, именуемая субъективной яркостью (brightness).<sup>1</sup> Так происходит потому, что самосветящиеся поверхности чаще всего воспринимаются самостоятельно, т.е. либо ярче, либо тусклее, нежели светлее или темнее окрашенными (как отражающие). С другой стороны, форма цветового пространства у цветовых ощущений, вызванных самосветящимися поверхностями и отражающими ровно одна и та же.

Перцепционная размерность «светлота» относится скорее к контрасту между областью и ее окружением: если две поверхности отражают (или испускают) одно и то же количество фотонов в единицу времени, их субъективные яркости остаются одинаковыми, но поверхность с темным окружением выглядит светлее, а окруженная светлым — темнее.

<sup>1</sup> К сожалению, термин «субъективная яркость» весьма неопределенный, поскольку отражающие поверхности наряду со светлотой обладают и неким уровнем субъективной яркости. Обычно светлоту относят к исходному ощущению, возникающему до учета зрительной системой вариаций в уровне освещенности этих поверхностей (см. раздел 3.3). К примеру, светлота равномерно окрашенной белой поверхности, лежащей частично на солнце, частично в тени, останется неизменной после того, как зрительная система распознает разницу в освещенностях, но будет иметь при этом много большую субъективную яркость на солнечном свете, нежели в тени. В логике, принятой нами в разделе 2.4.1, субъективная яркость — это термин, чувственно характеризующий интенсивность излучения всех поверхностей поля зрения на оптическом этапе восприятия, тогда как на фасетном этапе перцепты отражающих поверхностей обретают светлоту, а самосветящиеся так и остаются с одной лишь с субъективной яркостью.

тельно, усиленная чернота при включенном телевизоре суть результат *контраста* между высокой субъективной яркостью белых участков и прочих элементов изображения.

Понятие контраста относится не только к черным и белым, но ко всем темным и светлым цветам. Коричневый, однако, стоит особняком — как мы уже говорили, это темновая версия оранжевого или желтого. Например коричневый участок на экране телевизора или монитора окажется оранжевым или желтым, если мы посмотрим на него через черную картонку, свернутую в трубочку.

В текущей главе мы еще вернемся к обсуждению эффектов контраста.

### 3.1.3 Психофизическое соответствие

Теперь, когда определены вход и выход, рассмотрим то, как именно соотносятся между собой входные и выходные данные. Спецификация мэппинга физических параметров на психологические параметры перцептов именуется *психофизическим соответствием* и, в целом, суть описание на количественном уровне. Было бы очень удобно, если бы существовала ясная и простая связь между стимулом (светом) и цветовым ощущением, им вызываемым, но таковой, увы, нет. Понимание того, как световые потоки, отраженные поверхностями, преобразуются в цветовые ощущения, требует радикально больших знаний, чем те, что мы приобрели к данному моменту, потому что процесс такого мэппинга невероятно сложен. Поэтому, чтобы идти дальше, нам потребуются определенные упрощения.

Ранее мы уже сделали ряд упрощений, ограничив рассматривание окрашенных образцов стандартизированными условиями: постоянным нейтральным источником света и расположением образца на нейтральном фоне. Дальнейшее упрощение будет состоять в том, что рассматривать мы станем только те поверхности, которые отражают свет с т.н. *нормальным распределением энергии* (колоколообразным спектром) — см. рис. 3.1.6 А, В, С. Сразу отметим, что таких поверхностей в целом очень мало и еще меньше их в дикой природе.

Примечательным свойством нормальных распределений является то, что они характеризуются лишь тремя параметрами: *высотой пика, площадью и шириной*. Эти параметры позволяют упрощенно проиллюстрировать связь спектрального распределения энергии светового потока с тремя размерностями цветового ощущения. Физика стимулов нормального спектрального распределения согласуется с психологией цветового восприятия по трем направлениям, речь о которых ниже.

#### 3.1.3.1 Доминирующая длина волны → цветовой тон

Напомним, что коротко мы уже рассматривали связь между длиной волны монохроматического излучения и цветовым тоном, индуцируемого этим излучением цвета (цветового ощущения). У стимулов нормально распределенного полихроматического спектра цветовой тон их восприятия определяется точно так же как и у монохроматов, но только позицией пиковой (доминирующей, модальной) длины волны (см. рис. 3.1.6А). Когда мы говорим, что цветовой тон монохромата тот же, что и у полихромата, мы отнюдь не предполагаем идентичности цветовых ощущений, поскольку вполне могут иметь место отличия по двум другим перцепционным размерностям: светлоте (или субъективной яркости) и насыщенности. Также важно понимать и то, что стимулы с нормальным распределением способны вызывать ощущения далеко не

всех возможных цветовых тонов, в частности — неспектральных пурпурных (о чем мы поговорим позднее).

Отметим, что способность зрительной системы человека различать монохроматы разных длин волн неравномерна по спектру (Wright, 1941; Pokorny & Smith, 1970), что легко продемонстрировать простым экспериментом: левая половина поля сравнения излучает тестовый монохромат с длиной волны  $\lambda_c$ ; правая половина — исходно тот же монохромат, но затем длина его волны плавно меняется экспериментатором до тех пор, пока между половинами поля наблюдатель не обнаружит некоторую разницу. Приращение по длине волны —  $\Delta\lambda$  — суть *порог различимости по длине волны*. Перцепционным же критерием отличия явится изменение цветового тона правой половины поля сравнения.

Итогом эксперимента является W-образная кривая (рис. 3.1.6 АА). В районе наивысшей различимости — порядка 495 — 590 нм два стимула окажутся отличными друг от друга по цветовому тону, даже если разница по длине волны между ними составит всего лишь несколько нм. К примеру, зрение человека выдает разницу по цветовому тону между монохроматами в 490 и 495 нм. При этом монохроматы в 415 нм и в 410 нм — неразличимы поскольку в коротковолновой части спектра селективная способность зрения намного ниже.

Почему наша способность различать стимулы по длинам их волн варьирует по спектру? Окончательного ответа до сих пор нет, но, вероятно, максимумы различительной способности приходится на те участки, где наклон кривых спектров поглощения света фотопигментами меняется быстрее всего.

### 3.1.3.2 Спектральная площадь → светлота

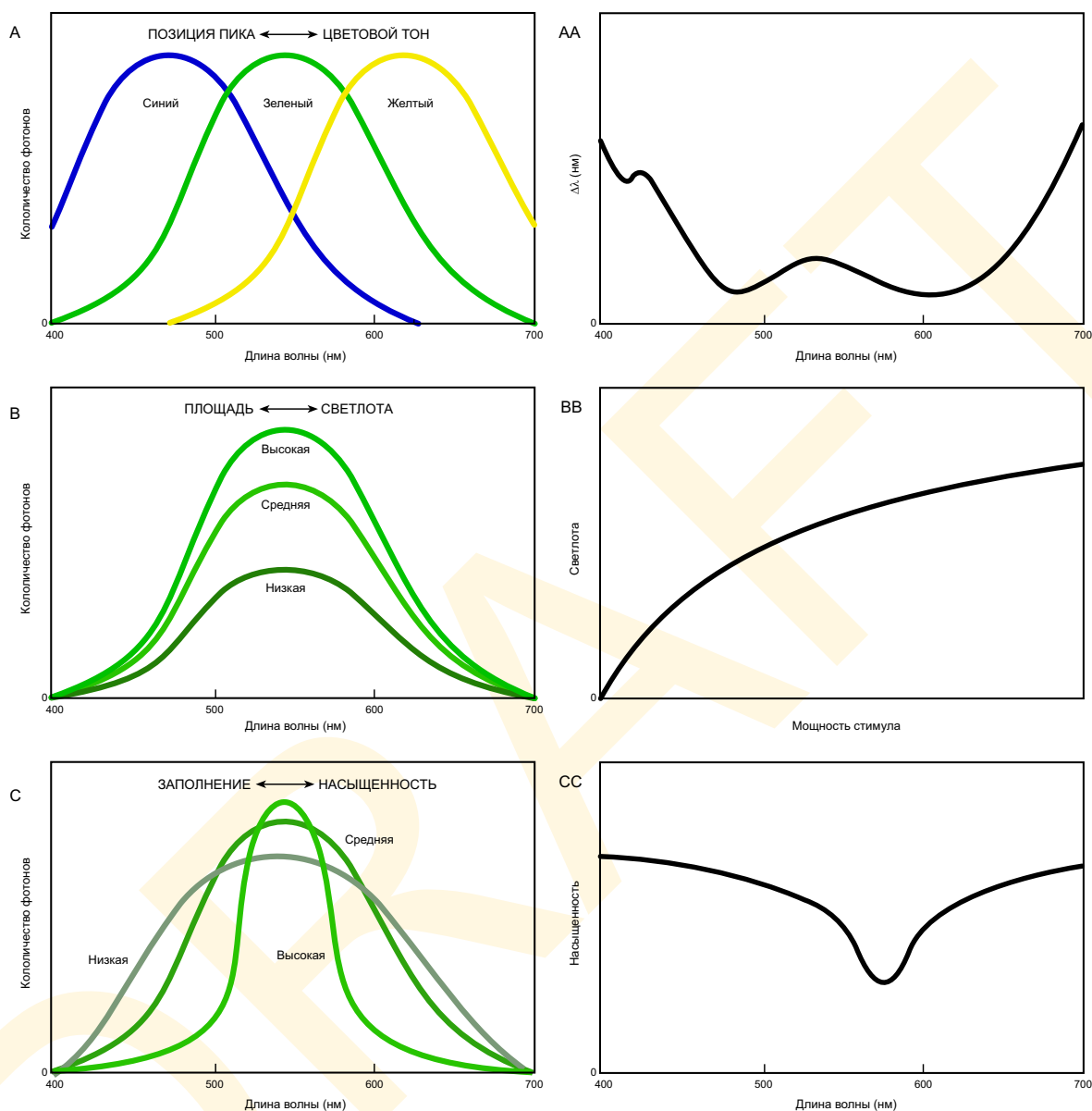
Связь светлоты с площадью под кривой спектрального распределения энергии стимула вполне закономерна, потому что площадь эта соответствует общему числу фотонов (в единицу времени), бомбардирующих сетчатку. Бóльшая площадь под спектральной кривой означает большее количество света, отраженного, соответственно, от светлой поверхности; малая — малому количеству света, отраженному, от темной поверхности (см. рис. 3.1.6 В).<sup>1</sup>

Каковы же взаимоотношения между величиной ощущения светлоты и мощностью стимула? Ответ на этот вопрос дает закон, открытый в 1957 г. американским исследователем Стенли Смитом Стивенсом и который гласит, что величина ощущения находится в степенной зависимости от величины интенсивности стимула (см. рис. 3.1.6 ВВ):

$$S = I^c$$

где  $S$  — величина ощущения,  $I$  — интенсивность стимуляции,  $c$  — постоянная. В контексте разговора о светлоте (или субъективной яркости)  $c = 1/3$ . Подробнее вопросов психофизики мы коснемся в Приложении А.

<sup>1</sup> Фактически нас интересует не абсолютное число фотонов, но относительное, т.е. взятое в сравнении с количеством фотонов, прилетевших от окружения исследуемого стимула. Сие потому, что, в конечном итоге, важен *контраст*, задающий размерность светлоты (субъективной яркости). Но когда образцы рассматриваются на общем фоне, то поверхности, отражающие меньшее количество света, разумеется выглядят темнее, а большее — светлее.



**Рис. 3.1.6** Психофизические соответствия по спектрам нормального распределения. А — отличия в положении спектрального пика стимулов дают отличия по цветовому тону их перцептов; В — отличия в спектральной площади и высоте пика дают отличия по светлоте; С — отличия в заполнении спектра (ширине пика) дают отличия по насыщенности; АА — различие монохроматов по длинам волн как функция от длины волны (Pokorny & Smith, 1970); ВВ — графическое представление в линейных осях степенного закона Стивенса (сплошная линия); СС — насыщенность как функция от длины волны; демонстрирует то, что перцепт монохромата на 570 нм обладает наименьшей насыщенностью в сравнении с прочими (Wright, 1946).

### 3.1.3.3 Заполнение спектрального диапазона → насыщенность

Давеча мы говорили, что монохроматические стимулы, у которых, напомним, нулевое спектральное заполнение, вызывают цветовые ощущения максимально возможной насыщенности. Напомним также, что «белый» солнечный свет обладает максимальным спектральным заполнением (почти горизонтальный спектр) и, соответственно, минимальной насыщенностью. У света, отраженного от поверхностей, разумеется, точно та же взаимосвязь между спектральным заполнением и насыщенностью цветового ощущения.

Отметим, что существует показатель обратный спектральному заполнению — это т.н. *колориметрическая чистота* стимула:

⇒ *Колориметрическая чистота* тестового стимула — это величина, определяемая отношением яркости монохроматического стимула к сумме яркостей данного монохроматического и стандартного ахроматического стимулов, которые при аддитивном смешивании уравнивают по ощущению насыщенности тестовый цветовой стимул.

То есть, колориметрическая чистота выражается уравнением:

$$p = \frac{L_{\lambda}}{L_{\lambda} + L_w},$$

где  $L_{\lambda}$  — фотометрическая яркость монохромата данной длины волны;  $L_w$  — фотометрическая яркость ахроматического стимула, подмешиваемого к монохромату для уравнивания с тестируемым стимулом по критерию насыщенности.

Понятно, что колориметрическая чистота монохроматов равна единице, но это не означает, что все монохроматы спектра воспринимаются равнонасыщенными.

\* \* \*

Итак, приняв, что у света, отраженного от поверхности, спектральное распределение энергии близко к нормальному, физиологическое соответствие прослеживается весьма четко и просто. Но стоит только отступить от нормального распределения, как сложности начинают расти как снежный ком, и для прогнозирования цветового тона, насыщенности и светлоты стимула, отраженного от поверхности, но имеющего произвольный спектр, требуется намного более сложный анализ, чем тот, что мы описали выше.

Еще труднее говорить о цветовом восприятии реальных объектов в реальных условиях. Как мы покажем в разделе 3.3, невозможно предсказать результат цветового восприятия поверхности только на основе знания спектрального распределения света, от нее отраженного: в игру вступает ряд прочих факторов — характер поверхностей, окружающих исследуемую, условия освещения, ориентация поверхности относительно источника света, воспринимаемый уклон поверхности вглубь пространства, а также тени от других предметов. Все из перечисленного оказывает серьезное влияние на результат цветового восприятия. Однако же мы отложим этот разговор до раздела 3.3.



## 3.2 ОПТИЧЕСКИЙ ЭТАП ЦВЕТОВОГО ВОСПРИЯТИЯ

Говоря об оптическом этапе в цветовом восприятии, мы будем исследовать представление цвета на уровне двухмерного ретинального изображения. По правде сказать, не совсем ясно, как это делать, потому что на оптическом этапе восприятия цветового ощущения может и не возникнуть. Вероятнее всего, осознанное цветовое ощущение появляется только на последних этапах, когда воспринимается 3D-структура поверхностей. Если так, то о цветовом представлении на оптическом этапе можно говорить, лишь сделав ряд допущений. К примеру, в целом предполагается, что ежели стимул однозначно и веридикально воспринят двухмерным, перпендикулярным прямой видимости и равномерно освещенным, то масса сложных факторов, что внедряются на фасетном этапе, не окажут влияния на результат цветового восприятия этого стимула. Ежели все перечисленные условия будут выполнены, то цветовое ощущение, возникшее от такой жестко двухмерной стимуляции, продемонстрирует нам суть и характер обработки на оптическом этапе цветового восприятия.

Оптимальными условиями достижения восприятия такого рода является просмотр равномерно окрашенных образцов через нейтрально-серый *редукционный экран* (*редукционную трубку*), отсекающий из поля зрения всё, кроме тестового стимула (или стимулов). Сии условия рождают такой психофизический феномен, как *апертурный* (*нелокализованный, пленочный*) *цвет*, поскольку стимул воспринимается нелокализованным по глубине, в виде пленки позади отверстия редукционного экрана. Редукционный экран устраняет всю усложняющую контекстную информацию о стимуле, в частности о его удалении от наблюдателя. Исходя из целей данного разговора мы будем полагать, что столь стесненные условия позволят нам отследить сугубо оптический этап цветового восприятия.

### 3.2.1 Феноменология

Наше понимание оптического этапа цветового зрения зиждется на неких базовых фактах. Многие из них были открыты более века назад исследователями, применявшими т.н. *поведенческие методы*, в частности тесты на различение и интроспективные описания. Мотивом к проведению исследований явилась необходимость в отработке эффективных приемов краскосмещения, понимании причин цветовой слепоты, природы хроматических постобразов, наведенных оттенков и хроматической адаптации.

#### 3.2.1.1. Метамерия

Как было показано ранее, лишь малая область в цветовом пространстве соответствует ощущениям, вызванным монохроматическими стимулами. Чтобы у наблюдателя возникло ощущение пурпурного цвета, либо какого-то ненасыщенного, скажем пастельного или серого, необходим стимул, представляющий собой смесь двух и более монохроматов. Сие поднимает наиважнейший вопрос о смешении цветовых стимулов, ответ на который во многом зависит от того, что смешивается: световые потоки или краски. Смешение световых потоков легче понять в логике цветового пространства, о котором мы будем говорить в первую очередь. Смешение красок (что ближе нам с сугубо практической точки зрения) много сложнее и описано в соответствующей литературе.

Итак, существует нехитрое правило, позволяющее средствами цветового пространства предсказать цветовое ощущение, вызванное смесью стимулов А и В: результирующее ощущение будет лежать на линии, соединяющей позиции ощущений А и В

(см. рис. 3.2.1 А).<sup>1</sup> Ежели два световых потока смешаны в равных пропорциях, то позиция, соответствующая результирующему ощущению окажется на середине расстояния между А и В. Если в смеси доминирует стимул А, то результирующий перцепт окажется ближе к позиции А в цветовом пространстве, но все равно на линии АВ. Понятно, что комплементарные смеси дают ахроматический (белый) свет: они лежат на противоположных сторонах цветового круга, и ахроматическая центральная точка оказывается на линии их соединяющей.

Если в смеси появится третий стимул (С), результирующее ощущение окажется где-то внутри треугольника с вершинами в позициях А, В и С. Коль скоро позиции ощущений, вызванных смесью А и В располагаются по отрезку АВ (в зависимости от соотношения количеств А- и В-стимулов), то ежели к АВ-смесям добавить какое-то количество стимула С, то позиция результирующего перцепта окажется где-то на линии, соединяющей позицию ощущения по АВ-смеси с позицией по ощущению от стимула С (см. рис. 3.2.1 В). Точно рассчитать позицию смеси можно, зная относительные пропорции стимулов А, В, и С.

Из сказанного на сей момент следует, что ежели три стимула А, В и С обладают максимальной насыщенностью вызываемых ими ощущений и выбраны так, что сильно удалены друг от друга по цветовому кругу, скажем насыщенный красный, насыщенный синий и насыщенный зеленый, то, манипулируя лишь относительными количествами стимулов А, В и С, мы можем воспроизводить вторичные смесевые стимулы практически всех цветов в круге.<sup>2</sup> Сие значит, что любое цветовое ощущение, находящееся внутри треугольника АВС можно получить, воспроизведя смесевой стимул, состоящий из различных долей исходных стимулов А, В и С. Однако по данной схеме невозпроизводимы высоконасыщенные желтые, пурпурные и голубые, хотя вполне воспроизводимы их менее насыщенные варианты.

Манипулируя абсолютными количествами стимулов А, В и С, то есть, меняя их фотометрическую яркость, можно воспроизвести стимулы, дающие ощущения, позиции которых заполнят почти всё цветовое тело человека. Сей факт в 1855 г. обнаружил выдающийся шотландский физик Джеймс Клерк Максвелл — тот самый Максвелл, что вывел уравнения электромагнитных полей.

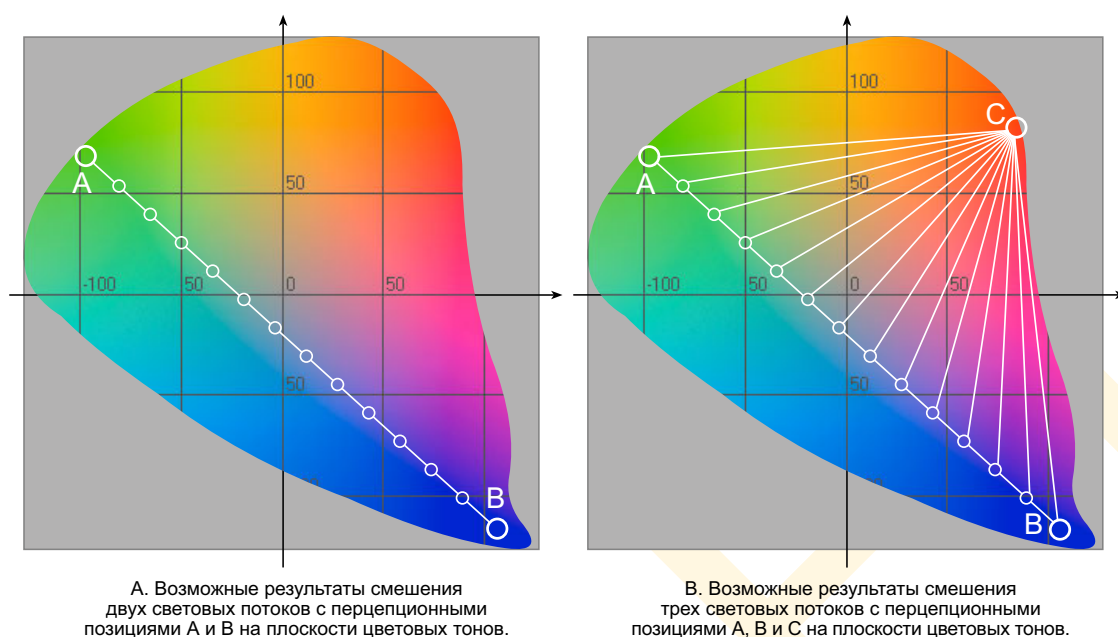
Итак, сказанное выше можно свести к общему правилу смешения световых потоков:

⇒ В отношении наблюдателей с нормальным цветовым зрением должная комбинация первичных световых потоков может воспроизвести стимул любого цвета.

«Должная комбинация» значит, что позиции ощущений, вызываемых *первичными световыми потоками* (или как еще говорят — *кардинальными стимулами* или *уравнивающими*), не лежат на одной прямой в цветовом треугольнике.

<sup>1</sup> Идеально прямыми линии окажутся только на т.н. *диаграммах цветности*, предложенных Международной комиссией по освещению (CIE). Однако диаграммы цветности характеризуют хроматическое *воздействие* стимулов на орган зрения, но не ощущения ими вызываемые. Поэтому в пространствах, характеризующих цветовые ощущения, линии, соединяющие позиции ощущений, возникающих от собственно смешиваемых стимулов, не обязательно оказываются прямыми.

<sup>2</sup> Говоря об относительных количествах стимулов в смеси, мы имеем в виду то, что суммарная фотометрическая яркость этой смеси остается неизменной и поэтому позиции цветовых ощущений от смесевых стимулов расположены в одной плоскости цветового пространства. Проще говоря, речь идет о балансе трех световых потоков.

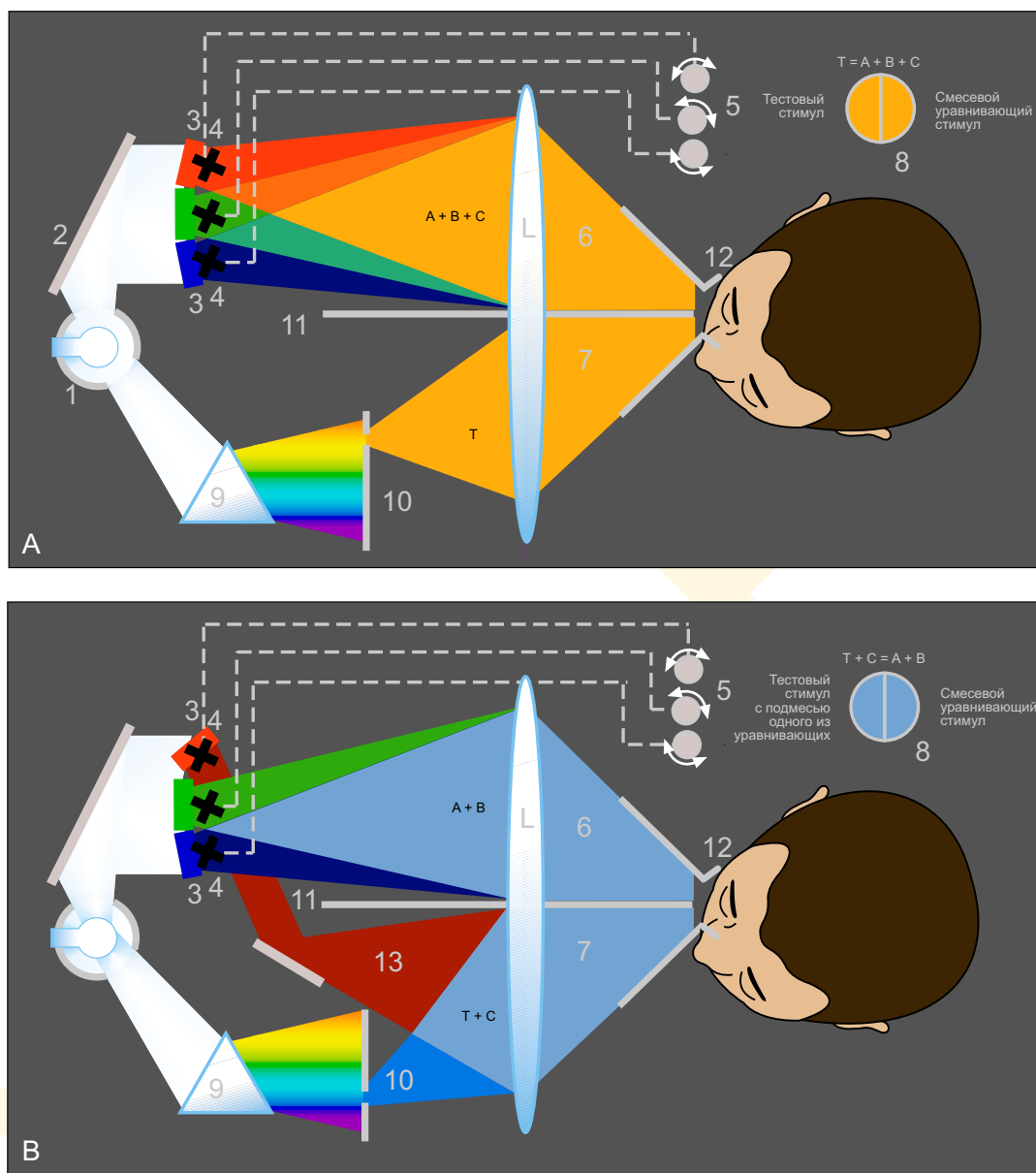


**Рис. 3.2.1** Результаты восприятия световых смесей, представленные в цветовом пространстве. А — смешивая окрашенные световые потоки в различных пропорциях, можно получать смеси, ощущения от восприятия которых расположатся в цветовом пространстве на одной линии. В — то же в отношении трех потоков образует треугольник, внутри которого располагаются позиции перцептов смесей.

Закон смешения положен в основу экспериментов по визуальному цветовому уравниванию, проведенных в 20-е годы прошлого века в Великобритании молодыми исследователями Дж. Гилдом и У. Райтом: с помощью специальной установки наблюдателю предъявлялось т.н. *поле сравнения*, левая половина которого излучала некий стимул (тестовый стимул), а в правой смешивались три *первичных световых потока* (*кардинальных стимула*), формирующих *смесевой уравнивающий стимул* (рис. 3.2.2 А). Наблюдатель, вращая ручки — регуляторы мощности первичных потоков, подбирал их соотношение в смеси так, чтобы ощущение, вызванное этой смесью, уравнивало бы исследуемый стимул, и граница между полями исчезла. Особо отметим, что спектрально световые потоки правого и левого полей были совершенно разными. Итак:

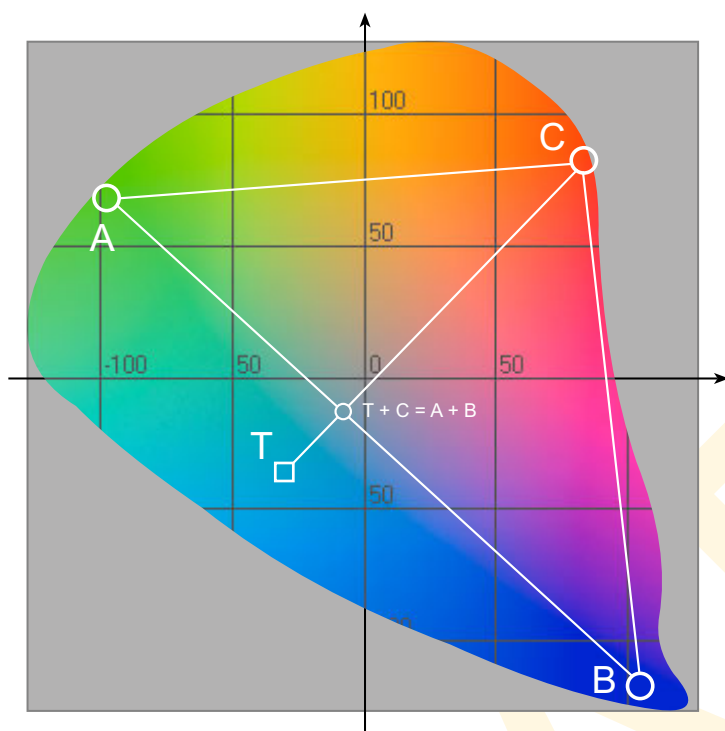
⇒ Стимулы, имеющие разный спектральный состав, но оказывающие на зрительную систему человека одно и то же воздействие и, следовательно, в одних и тех же условиях наблюдения, вызывающие одно и то же цветовое ощущение — именуются *метамерными стимулами* (*метамерами*).

Однако же истинность закона смешения может вызвать у читателя сомнение, потому что, как мы уже убедились, невозможно уравнивать некоторые стимулы, вызывающие ощущение высокой насыщенности, например, монохроматы. То есть, координаты цветовых ощущений, вызываемых такими стимулами, лежат вне цветового треугольника. Причина того, что простая смесь трех кардинальных стимулов не может покрыть всего цветового пространства в том, что процедура смешения, описанная выше, не позволяет вычитать, а, вернее, добавлять отрицательные световые потоки (подобно тому как мы можем добавлять векторы обратного направления в обычном



**Рис. 3.2.2** Трихроматическое уравнивание. А — стандартный вариант — тестовый монохроматический световой поток (Т) с длиной волны 610 нм (левая половина поля сравнения) уравнен смесью разных количеств трех монохроматических потоков с длинами волн 480 нм, 560 нм и 690 нм. В — когда позиция тестового стимула оказывается вне треугольника, образованного позициями уравнивающих потоков, то «отрицательное» количество одного из них (или даже двух, если необходимо) можно получить, добавив его определенное «положительное» количество к исследуемому стимулу.

Установка, позволяющая выполнять описанную процедуру, называется *трехстимульным (трихроматическим) колориметром* и состоит из следующих функциональных узлов: 1 — источник опорного белого света; 2 — зеркало; 3 — фильтры; 4 — obtюраторы, меняющие интенсивность (удельное количество) уравнивающих потоков; 5 — регуляторы угла раскрытия obtюраторов; 6 — уравнивающая смесь; 7 — уравниваемый стимул; 8 — поле сравнения; 9 — призма; 10 — шторка с узкой щелью; 11 — разделительная перегородка; 12 — визир; 13 — перенаправленный поток; L — фокусирующая линза.



**Рис. 3.2.3** Уравнивание с использованием «отрицательных» количеств кардинальных стимулов. Поскольку ощущение, вызываемое тестовым стимулом (Т) лежит вне треугольника АВС, уравнивать такой стимул можно только, используя «отрицательное» количество одного (или даже двух) кардиналов, т.е. подмешав его к стимулу тестовому. Результирующий перцепт окажется в точке пересечения отрезков АВ и ТС.

геометрическом пространстве). Однако же существует некий методологический трюк, срабатывающий аналогично вычитанию: «отрицательное» количество компонента смеси «добавляют», попросту подмешивая один из кардиналов к стимулу исследуемому — рис. 3.2.2 В.

На рис. 3.2.3 показано, как ощущение, координаты которого лежат в цветовом пространстве вне цветового треугольника данной тройки кардинальных стимулов, можно уравнивать с помощью отрицательных количеств одного из первичных световых потоков. Как видим, внутри треугольника нет позиции идентичной позиции ощущения от тестового стимула, и в данном случае стимул С добавлен не к смеси, а к стимулу тестовому, что и есть «вычитание» стимула С из уравнивающей смеси. Такая комбинация позволяет воспроизводить любые стимулы, координаты ощущений от которых лежат на отрезке ТС.

Поскольку одна из точек при этом неизбежно окажется на линии, соединяющей позиции А и В, можно говорить, что стимул, дающий ощущение Т, уравнивается положительными количествами А- и В-стимулов с добавкой отрицательного количества стимула С. Таким образом закон смешения, хоть и с натяжкой, но распространяется на все возможные тестовые стимулы.

Тот факт, что три световых потока — три кардинальных стимула — могут воспроизвести стимул любого цветового тона имеет огромное практическое значение, поскольку это базис индустрии изображений. К примеру, всё, что показывают нам цветной телевизор и экран компьютера — суть смеси всего лишь трех световых потоков: высоконасыщенных красного, зеленого и синего в разных комбинациях (в этом легко убедиться, осторожно и быстро взглянув на экран телевизора через лупу).

### 3.2.1.2 Законы константности метамерных равенств (законы Грассмана)

Законы, открытые немецким исследователем Германом Гюнтером Грассманом, дают три базовых характеристики трихроматического зрения (Grassman, 1854) и разговор о них непосредственно следует из предыдущей беседы о трихроматической теории.



⇒ **Закон аддитивности:** если к каждому из двух метамеров одним и тем же образом добавить одно и то же количество световой энергии, то они по-прежнему останутся метамерами.

То есть, добавка света к каждому из стимулов метамерной пары сохраняет метамерное равенство.

Рассмотрим пример. Предположим, что имеются два образца одного и того же желтого цвета, т.е. воспринимаемые одинаково желтыми. Один образец испускает монохроматический стимул с длиной волны 570 нм, а второй — является собой смесь трех монохроматов на 530, 560 и 600 нм. Если к каждому из образцов мы добавим по монохромату в 630 нм, оба образца по-прежнему останутся одинаковыми, но при этом синхронно станут оранжевыми. То есть, метамерное равенство сохранится.

⇒ **Закон скалярности (пропорциональности):** синхронное повышение или понижение интенсивности двух метамеров не влияет на метамерное равенство.

Возьмем желтые метамеры из предыдущего примера. Каждый инициирует одинаковое поглощение фотонов колбочками. Если интенсивность излучения обоими образцами повысить на одинаковую величину, то и уровень поглощений в обоих случаях изменится на одинаковую величину: стимулы станут восприниматься субъективно более яркими, но их метамерность сохранится.

⇒ **Закон ассоциативности:** замена одного метамера другим метамером сохраняет метамерное равенство.

Если некий стимул оказывает на зрительную систему точно такое же воздействие, что и один из стимулов метамерной пары, мы можем свободно заменить второй первым без утраты метамерного равенства.

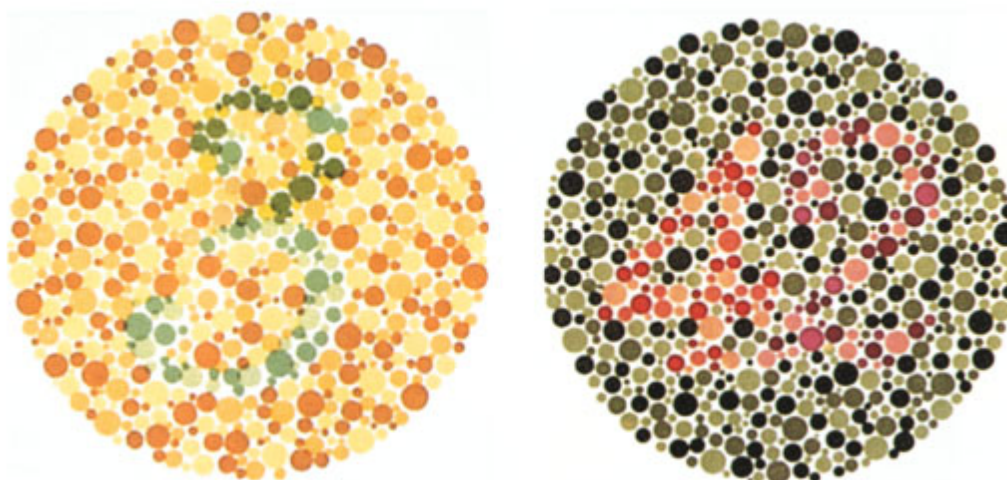
### 3.2.1.3 Цветовая слепота

Утверждая, что к любому стимулу можно подобрать уравнивающую смесь из трех кардиналов, мы обязательно делаем оговорку, что сие касается наблюдателей с нормальным цветовым зрением. Однако далеко не у каждого из нас цветовое зрение отвечает норме у 8% мужчин и у 1% женщин имеет место та или иная форма цветовой аномалии — эти люди не в состоянии испытать все ощущения, представленные в цветовом теле. Разница в заболеваемости у мужчин и женщин говорит о наследственном характере патологии и ее сцеплении с полом.

Дабы исключить у себя грубую цветовую аномалию, нужно всмотреться в рис. 3.2.4: если среди множества точек удастся легко различить цифры 3 и 42, цветовое зрение в норме. Если одна из цифр неразличима, значит имеет место как минимум частичная цветовая слепота.

**Нормальные трихроматы.** Лиц с нормальным цветовым зрением именуют *нормальными трихроматами*: это значит, что они могут визуальнo уравнивать любой стимул цветового пространства с помощью трех кардинальных стимулов и что они в состоянии испытать все цветовые ощущения, представленные в цветовом теле.

**Дихроматы.** Существует несколько видов цветовой слепоты и несколько степеней у каждой. Большинство тех, кого мы называем «цветослепыми» — это представители



**Рис. 3.2.4** Тест на цветовую слепоту. Лица с нормальным цветовым зрением легко различают цифру «3» в левом круге и цифру «42» в правом. Протанопы не видят «3» в левом круге и «2» в правом. Дейтеранопы не различают «3» в левом круге и «4» в правом.

С любезного разрешения Graham-Field Surgical Company.

той или иной группы дихроматов, потому что им для визуального уравнивания обычно требуются лишь два кардинала.

Существует три вида дихромазии: в двух случаях у больных отсутствует способность различать стимулы, в норме (т.е. у трихроматов) вызывающие ощущения красных и зеленых цветовых тонов. Эти стимулы воспринимаются, вероятнее всего, как нейтральные. В результате эти люди окружающий мир видят в градациях синих, желтых и серых тонов, поскольку красно-зеленая хроматическая размерность выпадает из их цветового пространства (см. рис. 3.2.5 В и С). Данные два вида патологии относят к категории *красно-зеленой слепоты*, а по отдельности они именуются как *протанопия* («pro» по-гречески означает «первый» + «анопия» — «слепота») и *дейтеранопия* («deuter» — «второй»).

Отличить один вид патологии от другого можно только с помощью тщательно выполненного теста на поиск т.н. *монохроматической нейтральной точки* — специфической длины волны монохроматического стимула, который воспринимается больным как серый. Стимулы коротких длин волн воспринимаются синими, а длинные — желтыми, но где-то между ними есть длина волны, стимул которой воспримется нейтральным, то есть, ни синим и ни желтым. (У нормальных трихроматов нет нейтральной монохроматической точки, поскольку все монохроматические стимулы воспринимаются ими как высоконасыщенные определенного цветового тона). У протанопов нейтральная точка лежит на 492 нм, а у дейтеранопов — на 498 нм (см. рис. 3.2.6). Несмотря на очевидное сходство двух вариантов патологии, различие между ними имеет важное физиологическое значение, которое, однако, мы рассмотрим позже и очень коротко. Наконец отметим, что красно-зеленая слепота (в обоих вариантах) встречается не столь уж часто — всего лишь у 1% мужчин и 0.02% женщин.

Третий вариант дихроматизма — *тританопия* («tri» с греческого значит «третий») — крайне редкая патология и встречается только у 0.02% мужчин и 0.01% женщин. Тританопы не в состоянии различать стимулы, в норме вызывающие ощущения



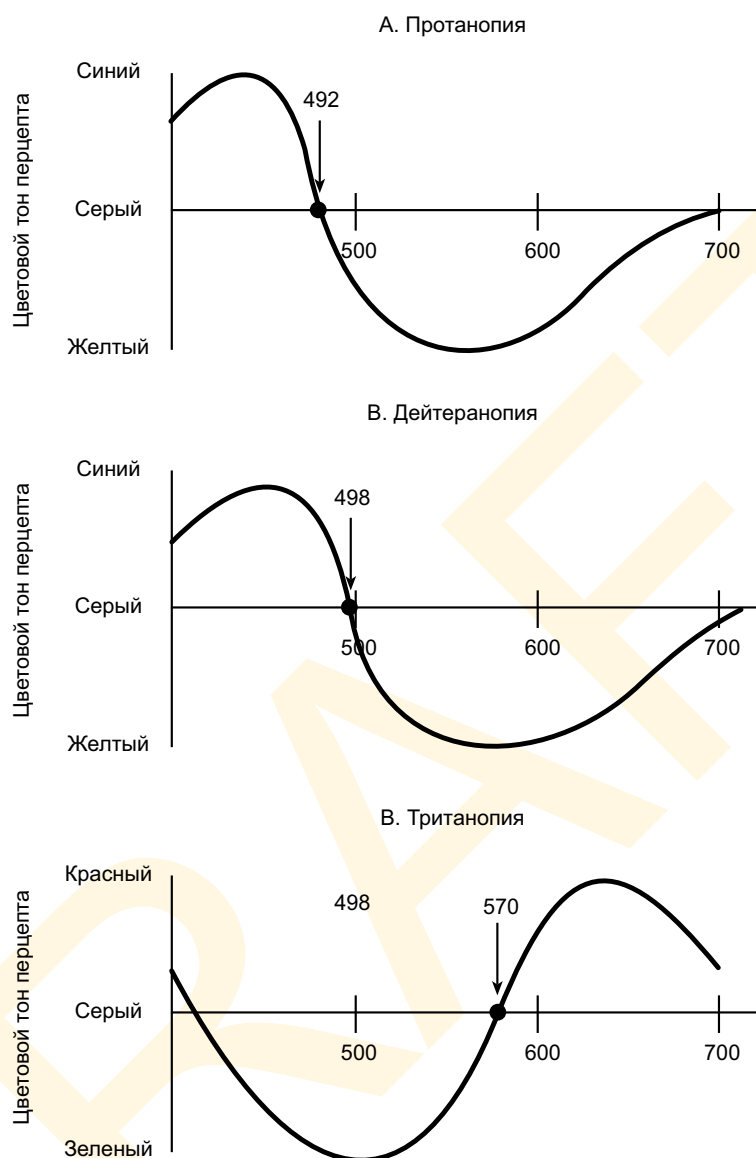
**Рис. 3.2.5** Имитация восприятия сцены при трех вариантах дихромазии. А — полноцветный перцепт нормального трихромата. В — вероятный перцепт протанопы. С — вероятный перцепт дейтеранопы. D — вероятный перцепт тританопы.

синих и желтых цветовых тонов, поэтому окружающий мир в их восприятии — это градации красных, зеленых и, вероятно, серых (см. рис. 3.2.5 D), поскольку выпадает сине-желтая размерность цветового тела. Из рис. 3.2.6 ясно, что стимулы коротких длин волн воспринимаются тританопами как зеленые, а длинных — как красные, с нейтральной точкой на 570 нм.

**Монохроматы.** Зрительный опыт некоторых лиц — это вариации светлоты предметов и только, т.е. фактически черно-белое зрение. Таковых мы назовем *монохроматами*, и понятно, что им для уравнивания любого стимула понадобится лишь один кардинал. Их цветовое пространство — это проекция цветового тела на ось светлоты.

#### 3.2.1.4 Цветовая недостаточность

Существует несколько видов т.н. *цветовой недостаточности*, при которой цветовое тело страдает не столь жестко. Лиц с такого рода патологией относят к *аномальным*



**Рис. 3.2.6** Нейтральная точка при цветовой слепоте. В спектре существует одна длина волны, монохромат которой воспринимается дихроматами неокрашенным. Конкретное значение характеризует вид цветовой слепоты.

*трихроматам* и на их долю приходится оставшаяся часть цветовых аномалий. Эти люди способны уравнивать любой стимул трихроматической смесью, но пропорции кардиналов в смеси окажутся совсем иными, нежели у нормальных трихроматов. У аномальных трихроматов, страдающих красно-зеленой недостаточностью, цветовое тело сплюснуто вдоль красно-зеленой оси цветового пространства, но все равно при этом остается трехмерным; у лиц с сине-желтой недостаточностью цветовое тело, соответственно, сплюснуто по сине-желтой оси. Отличие аномальных трихроматов от трихроматов нормальных не выражено резко, но имеет ряд градаций отклонения от нормы.

Теперь настало время обратиться к физиологической подоплеке цветовых аномалий, дабы понять, почему одни виды патологии встречаются, а другие — нет. Стоит



сказать, что о цветовых аномалиях человечеству было известно задолго до того, как был выявлен патогенез этих расстройств, и одной из главных задач первых теорий зрения являлось именно истолкование механизма цветовых аномалий. Как мы увидим в дальнейшем, феномены цветовой слепоты играют исключительно важную роль в истории науки о цветовом восприятии человеком окружающего мира.

### 3.2.1.5 Дихроматические и монохроматические перцепты

Два слова о том, откуда нормальные трихроматы знают — или по крайней мере думают, что знают — как видят мир дихроматы или монохроматы. Проблема уходит корнями в древнюю философию характера осознанного зрительного опыта: как думает уважаемый читатель, скажем опыт ощущения красноты у автора тот же, что и у читателя? Тот факт, что одни и те же вещи мы называем «красными» — еще не доказательство равенства ощущений, потому что ощущение красного у одного индивида может оказаться сходным с ощущением зеленого у другого, только первый его называет «красным». Сие на полном серьезе: мы называем красным один и тот же объект, но наши ощущения могут оказаться взаимнопротивоположными. В самом деле, логически вполне возможно, что у двух наблюдателей цветовые ощущения взаимнообратны, как если бы радуга была задом наперед (подробнее см. раздел 13.1.2).

Итак, учитывая, что никто не в силах заглянуть в сознание к другому, как же трихроматы узнают о цветовых ощущениях дихроматов?

Существует два возможных ответа. Первый дают нам исключительно редко встречающиеся цветослепые на один глаз (MacLeod & Lennie, 1976; Sloan & Wollach, 1948). Эти люди могут поведать нам о том, каков окружающий мир правым глазом и каков он левым. Однако удовлетворительным этот ответ кажется лишь до тех пор, пока мы не зададим себе вопрос: «А не влияет ли наличие трихроматического глаза на восприятие информации, поступающей от его дихроматического собрата?». Возможно, лишь только при способности воспринимать мир нормальным трихроматическим глазом человеку кажутся (при закрытом нормальном глазе) синими и желтыми те самые в норме красные и зеленые стимулы, что неспособен различить его дихроматический глаз. Возможно, цветовые ощущения, возникающие в сознании обычных двуглазых дихроматов, радикально отличны от любых ощущений, возникающих у нормальных трихроматов, пусть и одноглазых. Как видим, сие более осторожная точка зрения, ведущая ко второму варианту ответа: «Вполне вероятно, что никак...». Здесь мы вплотную подходим к проблеме т.н. *зрительной осведомленности*, о которой будем говорить в гл. 13.

### 3.2.1.6 Хроматические постобразы

В понимании цветового зрения важную роль играют феномены, относящиеся к категории т.н. *цветовых постобразов* — неких послеэффектов зрительного восприятия. Постобразы обнаруживают себя появлением в поле зрения высоконасыщенных цветовых пятен, остающихся в течение весьма продолжительного времени. Каждый из нас наблюдал ахроматические постобразы после короткой экспозиции интенсивным светом, например светом фотовспышки или солнца. Но при устойчивом и долгом наблюдении высокохроматичных стимулов мы можем наблюдать постобразы очевидно хроматические. Имеет смысл вновь взглянуть на рис. 2.1.8 в гл. 2, дабы убедиться в выраженности эффекта: в возникшем постобразе мы увидим красно-бело-синий американский флаг (хоть и не сильно насыщенный), поскольку цветовой тон ощущения от фактического стимула рождает *комплементарный цветовой тон* в постобразе. Компле-



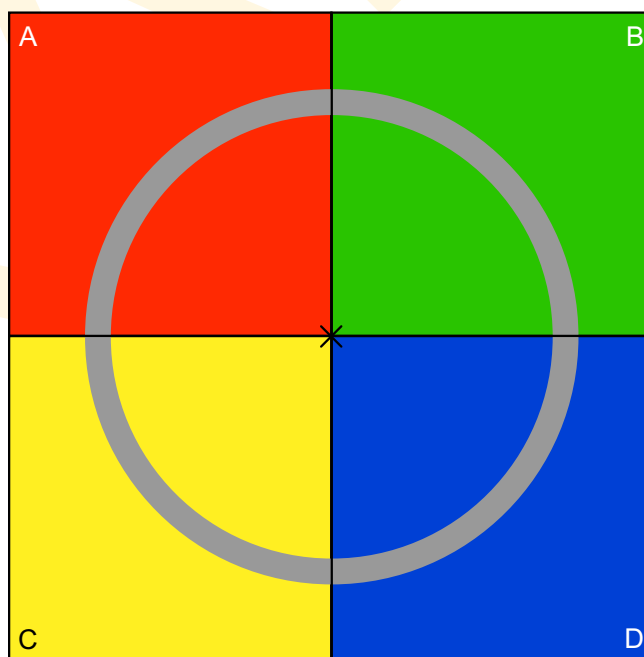
мент (дополнение) к любому цветовому тону расположено на диаметрально противоположной стороне цветового круга. К примеру, тон дополнительный к зеленому цветовому тону — красный, желтому — синий и т.п.

Существует простейший способ определить какой цветовой тон окажется комплементарным исследуемому: сосредоточить на первом взгляд, а затем спустя какое-то время перевести его на лист белой бумаги — зрительная система «автоматически» даст ответ.

### 3.2.1.7 Симульный хроматический контраст

Комплементарные цветовые тона играют важную роль в явлении т.н. *симульного хроматического контраста*, часто именуемого *наведенным цветовым оттенком*, поскольку комплементарность в данном случае — пространственно-зависимый феномен (но не темпорально). Классическая демонстрация эффекта симульного контраста — это предъявление наблюдателю двух стимулов с идентичными спектрами, но воспринимаемых по-разному под влиянием стимула, создаваемого их окружением. Иными словами, когда стимул предъявлен на интенсивно-окрашенном фоне, то фон этот может оказать определенное (и порой весьма заметное) влияние на характер цветового ощущения от исследуемого стимула. Отметим, что сегменты нейтрального круга на рис. 3.2.7, физически идентичны, но выглядят слегка окрашенными: на красном фоне сегмент воспринимается слегка зеленоватым, на зеленом фоне — красноватым, на желтом — синеватым и желтоватым — на синем фоне. Наведенный оттенок комплементарен цветовому тону фона, если тестовый стимул ахроматичен. Если исследуемый стимул хроматичный, то результирующий цветовой тон сдвинется в сторону комплементарного тона, а степень этого сдвига окажется зависимой от насыщенности фона и продолжительности наблюдения.

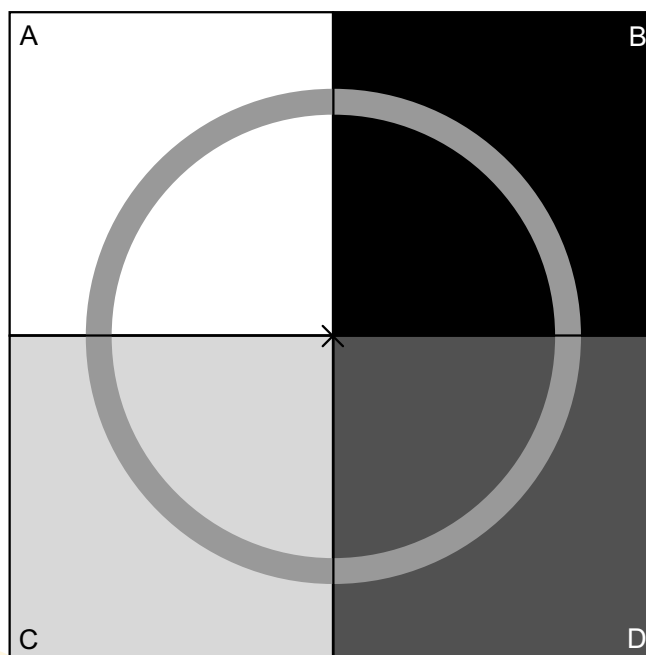
Эффект симульного контраста касается и ахроматического восприятия. На рис. 3.2.8 дан пример *симульного ахроматического контраста*: картина в целом аналогичная контрасту хроматическому, за исключением того, что фоны хроматически нейтральны. Отметим, что у сегментов кольца разная светлота, зависящая от фотометрической яркости фона: сегмент в темном квадрате воспринимается самым светлым, тот, что на белом фоне — самым темным, а оставшиеся два — занимают промежуточное положение. Если кольцо



**Рис. 3.2.7** Симульный хроматический контраст (наведенный оттенок). Серый круг отражает стимул, спектральное распределение энергии которого неизменно по всей площади круга. При этом перцепт круга приобретает определенный оттенок в зависимости от окружения: зеленоватый на красном фоне, красноватый на зеленом, желтоватый на синем и синеватый на желтом.

вырезать из страницы и рассмотреть на равномерно окрашенном фоне, то оно окажется совершенно однородным по светлоте.

**Маховы полосы.** Еще одним общеизвестным проявлением ахроматического контраста являются т.н. *маховы полосы*, названные в честь австрийского физика и психолога Эрнста Маха, открывшего это явление. Маховы полосы — это вариации светлоты (или субъективной яркости), возникающие на определенном вида краях. На рис. 3.2.9 видна сверх-светлая полоса на светлой стороне края и сверх-темная — на темной стороне. Ниже на графике представлен характер этой иллюзии: серой линией отмечен перепад фотометрических яркостей в рисунке, а черной — перепад светлоты. Ниже мы коснемся того, что явления контраста, в частности такие как маховы полосы — плод нейронной обработки определенного характера.



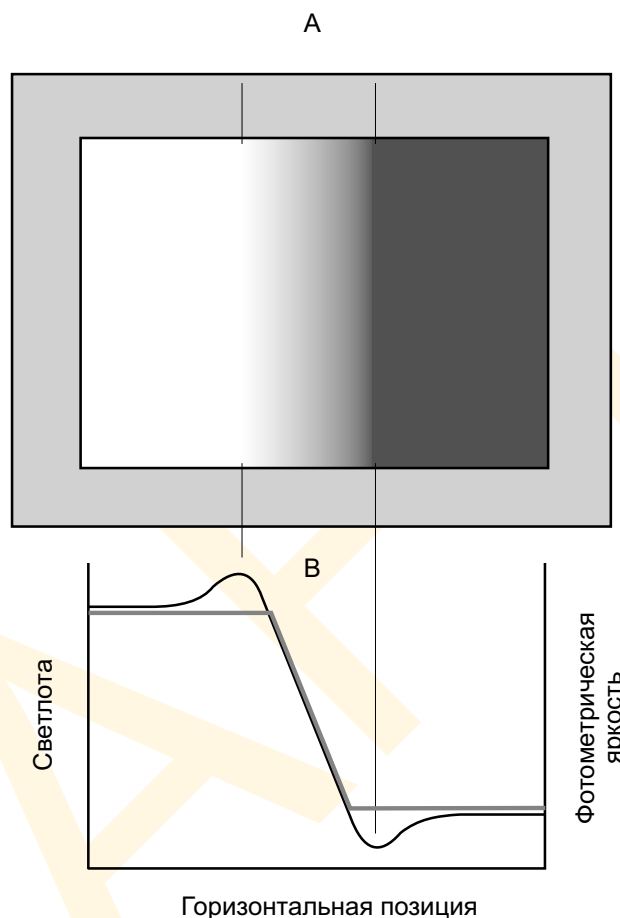
**Рис. 3.2.8** Симультанный ахроматический контраст. Серый круг отражает хроматически нейтральный стимул, фотометрическая яркость которого неизменна по всей площади круга. При этом перцепт круга варьирует по ощущению светлоты в зависимости от яркости фона.

### 3.2.1.8 Хроматическая адаптация

В разделе 1.1.3 мы обсудили феномен темновой адаптации — рост световой чувствительности зрительной системы после долгого пребывания в темноте. Противоположным эффектом является адаптация световая — снижение чувствительности зрительной системы после определенной экспозиции ярким светом. Поэтому вполне закономерно, что зрение способно в том числе к *адаптации хроматической*, которая суть снижение чувствительности к хроматичным стимулам, возникающее практически сразу же после их воздействия. Удостовериться в сказанном можно закрыв правый глаз и сосредоточив взгляд на черной точке внутри красно-зеленой фигуры рис. 3.2.10 (желательно при ярком свете); через полминуты закрыть левый глаз, посмотреть на рисунок правым и убедиться в том, насколько более насыщенной окажется фигура. Затем можно вновь адаптировать левый глаз по левой фигуре и посмотреть на рис. 3.2.10В — вначале левым (адаптированным) глазом, затем правым: левому глазу будет казаться, что красная полоска выглядит менее насыщенной внизу и более насыщенной сверху. Обратный эффект возникнет по зеленой полоске, которая сверху окажется менее насыщенной, а внизу — более. У желтой полоски возникнет резко выраженный хроматический сдвиг: она станет красноватой сверху и зеленоватой внизу. Аналогичные изменения произойдут и с синей полоской. Отметим, что по правому (неадаптированному) глазу описанных эффектов нет.

Хроматическая адаптация особо проявляет себя, если все поле зрения долгое время

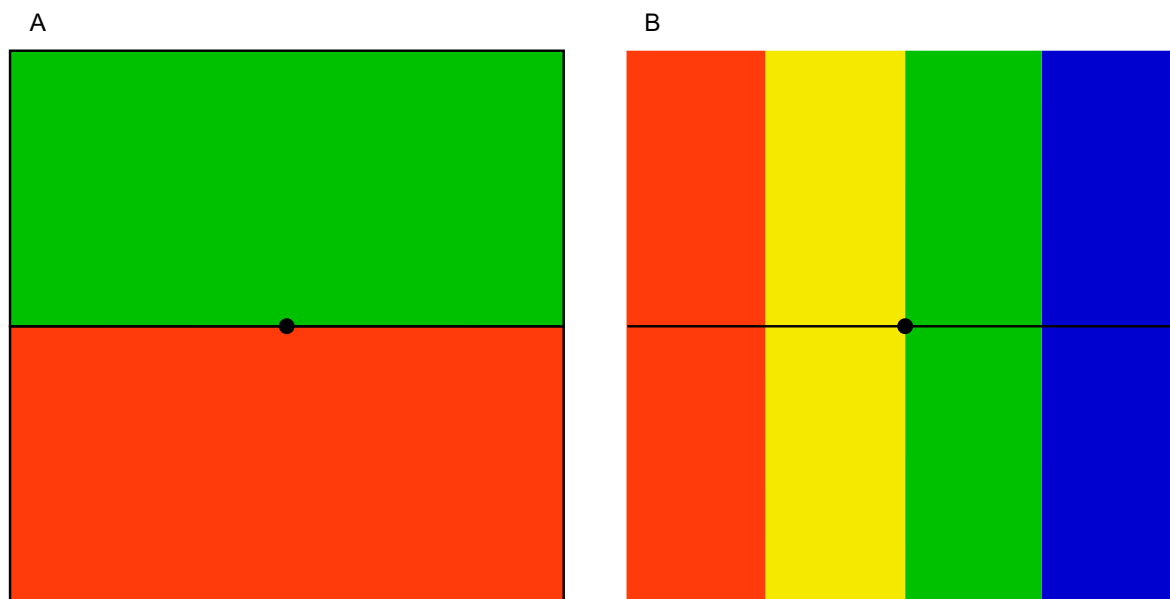
стимулируется пространственно равномерным хроматичным стимулом. Полностью равномерный стимул, заполняющий собой все поле зрения именуется *ганцфелдом*, что по-немецки значит «сплошное поле». Невзирая на свой исходный цвет любой ганцфельд, в конечном счете, начинает восприниматься как нейтральный. В сказанном легко убедиться, разрезав желтый пинг-понговый шарик на две половинки и приставив их к глазам: через пару минут мы будем видеть лишь сероватую пелену, но не равномерную желтую поверхность. Стоит лишь убрать эти импровизированные фильтры, как окружающий мир первое время будет восприниматься синеватым: сие суть т.н. *хроматическое последствие*, возникающее из-за того, что между хроматической адаптацией и последствиями существует очень тесная связь — хроматическая адаптация всегда вызывает хроматические последствия комплементарных цветовых тонов. Однако следует понимать, что хроматическое последствие — это не то же самое, что хроматическая адаптация (которая суть снижение чувствительности к данному типу стимуляции при долгой экспозиции): у последствий имеется тенденция проявлять себя «оппозитно» и всегда *после* стимуляции, в частности в ответ на нейтральный стимул, такой как свет от белой поверхности.



**Рис. 3.2.9** Маховы полосы. А — иллюзорные светлая и темная полосы на противоположных сторонах плавного края. В — сравнительный рафик динамики световой интенсивности (серая линия) с графиком, отражающим динамику светлоты (черная линия).

### 3.2.1.9 Феномен Вецольда — Брюкке

Большинство монохроматических стимулов по мере изменения интенсивности в восприятии своем дают легкий сдвиг цветового тона. Для иллюстрации сказанного рассмотрим следующий эксперимент: левая половина поля сравнения излучает монохромат с длиной волны  $\lambda_v$ , которая может меняться по воле экспериментатора; правая половина — монохромат с  $\lambda_e$ , который остается неизменным на протяжении всего опыта. По ходу эксперимента интенсивность свечения левого поля постепенно повышают. Если между правым и левым полями начинает появляться разница по цветовому тону, длину волны левого монохромата немного меняют до исчезновения этой разницы.



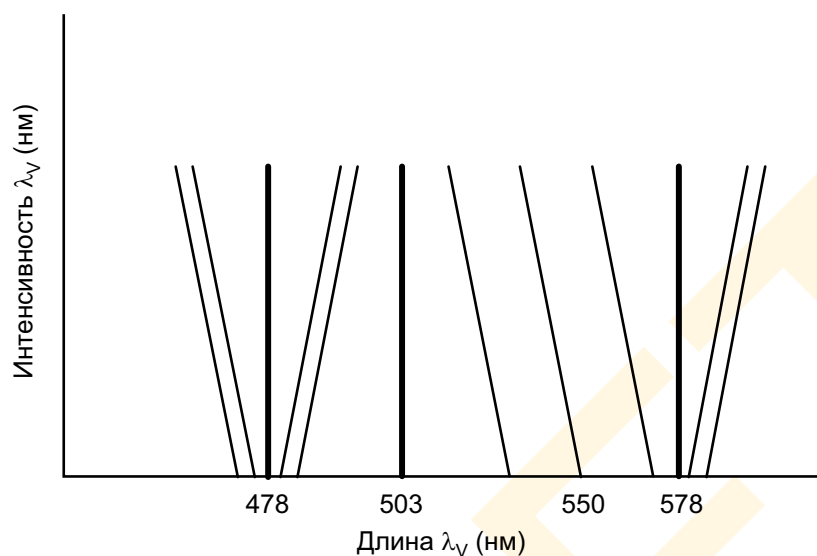
**Рис. 3.2.10** Хроматическая адаптация. Если на одну минуту сосредоточить взгляд на черной точке красно-зеленого квадрата (А), а потом перевести его на черную точку фигуры В, то нижняя красная полоска будет восприниматься менее насыщенной, чем верхняя из-за адаптации к длинноволновому стимулу, испускаемому нижней половиной фигуры А. Верхняя зеленая полоска также окажется менее насыщенной чем нижняя из-за пониженной чувствительности к средневолновому стимулу верхней половины фигуры А. То же коснется двух оставшихся полос.

Теперь обратимся к конкретному примеру, при котором левое поле испускает монохроматический стимул с длиной волны в 550 нм. По мере роста интенсивности этого стимула, цветовой тон его перцепта смещается в теплом направлении, т.е. желтеет. Для возврата к исходному цветовому тону необходимо понизить длину волны, на что указывает линия рис. 3.2.11, начинающаяся на отметке в 550 нм и смещающаяся в сторону коротких волн по мере роста интенсивности. Сия линия именуется *локусом постоянного цветового тона* (в данном случае желто-зеленого).

На рис. 3.2.11 показаны локусы цветовых тонов различных монохроматов. По мере изменения интенсивности большинства из них цветовой тон их перцепта меняется и, соответственно, требуется изменить длину волны монохромата, дабы удержать цветовой тон неизменным.

Стабильные локусы указывают на то, что у монохроматов трех длин волн цветовой тон перцептов по мере изменения интенсивности не меняется (Hurvich, 1981). Сии длины волн — 478, 503 и 578 нм — относятся к группе т.н. *инвариантных длин волн* (*инвариантных точек*). Соответствующие цветовые тона перцептов: по монохромату в 478 нм — синий, в 503 нм — зеленый и в 578 нм — желтый относятся к группе т.н. *однозначных цветовых тонов*, то есть ощущаемых идеально чистыми, и не «смешанными» из других.

Физиологический базис данного феномена, получившего название феномена Вефольда-Брюкке, дает оппонентная теория цветового зрения (см. раздел 3.2.2).



**Рис. 3.2.11** Приблизительные результаты эксперимента по измерению характера и степени выраженности эффекта Веинштейна — Брюкке. Мы видим, что три монохромата (478, 503 и 578 нм) инвариантны, то есть цветовой тон их перцептов не меняется по мере роста интенсивности этих стимулов.

### 3.2.2 Теории цветового зрения

Итак, мы описали несколько наиболее значимых феноменов цветового восприятия. Как уже отмечалось в гл. 2, факты вещь необходимая, но совершенно недостаточная для научного понимания физиологических механизмов, в частности — зрительных: нужна еще **теория**, четко соотносимая с этими фактами по отдельности и способная давать контролепригодные прогнозы новых.

Давайте теперь познакомимся с тем, как исследователи восприятия толковали психофизиологические факты, которые мы давеча обсудили.

#### 3.2.2.1 Трихроматическая теория

На заре науки о цветовом зрении звучала мысль о том, что на стимулы каждой из длин волн откликается свой специфический фоторецептор. Поскольку человек способен к восприятию тысяч самых разнообразных стимулов, во всех областях сетчатки должны располагаться тысячи специфических клеток. При этом кодированию прочих характеристик стимула в этой логике уже не было места.

Первую действительно научную теорию цветового зрения в 1777 г. предложил британский химик-красильщик Джордж Пальмер (известный также как Георг Гирсод'Гентилли), но почти одновременно с ним (в 1802 г.) те же идеи высказал британский философ сэр Томас Юнг:

«Совершенно невозможно вообразить, что в каждой чувствительной точке сетчатки находится бесконечное число частиц, способных колебаться в унисон со всеми возможными вибрациями света. Мы вынуждены думать, что число последних все же ограничено, например тремя принципиальными вибрациями... и что каждой из частиц придается большее или меньшее движение колебаниями, более или менее



отличными от унисона. Чувствительные нервные нити, вероятно, состоят из трех порций — каждая для одного из принципиальных цветов» (McAdam, 1970).<sup>1</sup>

Теория Пальмера — Юнга предполагала наличие трех типов фоторецепторов в глазу человека. Предполагалось также, что эти рецепторы порождают физиологически первичные цветовые ощущения красного, зеленого и синего. Все прочие цветовые ощущения толковались как комбинация трех первичных.

Позднее идею уточнили и углубили Джеймс Клерк Максвелл (1855) и Герман фон Гельмгольц (1867/1925), но сегодня она известна как *трихроматическая теория Юнга — Гельмгольца*.

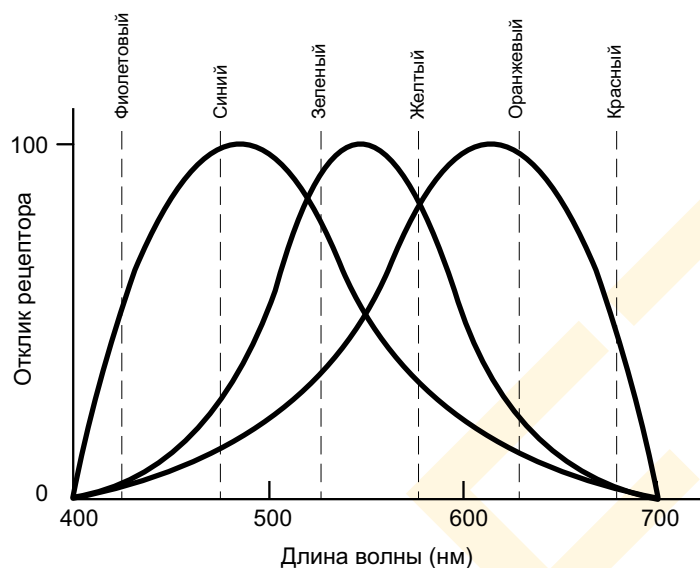
Несмотря на то, что слог Юнга выглядит нынче почти что древним, ученый четко сформулировал теорию, на которой базируется современный научный взгляд на цветовое зрение. Фактически Юнг понял то, что хроматическая информация регистрируется светочувствительными клетками весьма ограниченного числа видов, предположительно — трех, и что относительная активность этих клеток и кодирует хроматическую информацию.

Согласно Гельмгольцу отклики фоторецепторов трех разных типов суть функция от длины волны фотонов, воздействующих на эти рецепторы (рис. 3.2.12). Предполагалось, что рецепторы одного типа давали пик отклика на фотоны коротких волн (воспринимаются синими), другого типа — на фотоны средних волн (воспринимаются зелеными) и третьего типа — на фотоны длинных волн (воспринимаются красными). Рецепторы каждого типа в отдельности, сами по себе, если угодно «цветослепы», в том смысле, что не существует единого рецептора, позволяющего зрительной системе отличать свет разных длин волн. Но поскольку спектральные функции откликов рецепторов пересекаются, любой стимул любой длины волны неизбежно окажет воздействие на рецепторы всех трех типов, но, разумеется, в разной степени. Гельмгольц предположил, что такая стимуляция и есть суть цветового зрения (см. 3.2.13).

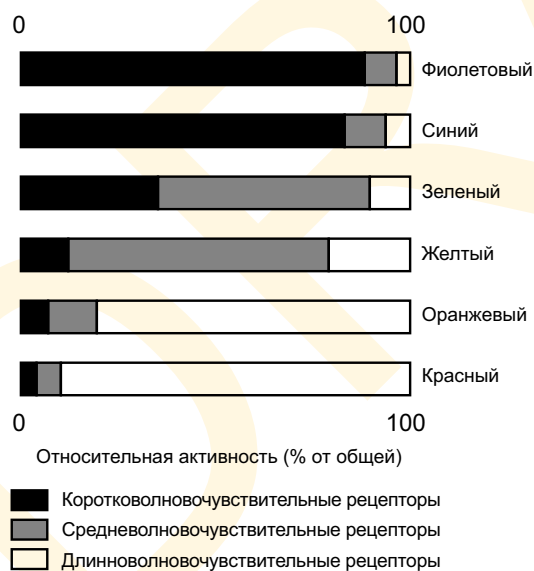
Трехрецепторная (трихроматическая, трехкомпонентная) теория способна объяснить множество феноменов цветового зрения, к примеру: трехмерность цветового пространства объясняется именно трехрецепторностью, и тот факт, что для визуального уравнивания любого стимула нужны три кардинальных, также объясняется работой рецепторов трех типов в связке с гипотезой о т.н. *пусковом паттерне* (см. раздел 2.2.3). Метамеры существуют потому, что стимулы самых разнообразных спектральных распределений способны сформировать на сетчатке один и тот же пусковой паттерн.

Трихроматическая теория также способна объяснить различные виды цветовой слепоты. Три варианта дихромазии — суть прямое следствие отсутствия в сетчатке колбочек одного из типов: у протанопов нет длинноволновочувствительных колбочек,

<sup>1</sup> Справедливости ради стоит отметить размышления Михаила Васильевича Ломоносова, озвученные им в докладе «Слово о происхождении света, новую теорию о цветах представляющее, в публичном собрании Императорской Академии Наук июля 1 дня 1756 года говоренное Михаилом Ломоносовым»: «Я приметил и через многие годы многими прежде догадками, а после доказательными опытами с довольною вероятностью утвердился, что природа эфирных частиц имеет совмещение с тремя родами действующих первоначальных частиц, чувствительные тела составляющих... От первого рода эфира происходит цвет красный, от второго — желтый, от третьего — голубой. Прочие цвета рождаются от смешения первых... Натура тем паче всего удивительна, что в простоте своей многохитростна, и от малого числа причин произносит неисчислимые образы свойств, перемен и явлений». Любопытно, что в своей работе Т. Юнг ссылается на доклад Ломоносова. — *Прим. Пер.*



**Рис. 3.2.12** Оригинальная трихроматическая теория цветового зрения, предполагающая, что цветовое ощущение возникает в результате световой активации гипотетических рецепторов трех различных типов, спектральные чувствительности которых хоть и перекрываются, но имеют максимумы в разных участках спектра. Согласно этой теории стимулы разных характеров приводят к разным картинам активации рецепторного аппарата и возникновению зрительных ощущений разных цветов.



**Рис. 3.2.13** Уровни активации трех типов рецепторов в трихромаической теории. Считается, что отличия в цветовых ощущениях — плод отличий в балансах активности трех систем рецепторов. Проиллюстрировано шестью цветовыми тонами.

у дейтеранопов — средневолночувствительных, а у тританопов — коротковолночувствительных. У монохроматов отсутствуют колбочки каких-либо двух типов.

Итак, трихроматическая теория оказалась очень удачной, потому что способна дать объяснение множеству фактов, базируясь всего лишь на нескольких и довольно простых гипотезах. В результате, в течение целого века она доминировала в науке о зрении.

### Монохроматическое уравнивание.

Чтобы сделать первый шаг в изучении трихроматической теории, давайте разберем то, почему особь — монохромат, т.е. у которой наличествует лишь один вид фотопигмента, не способна к цветовому восприятию окружающего мира. На рис. 3.2.15 показан спектр поглощения гипотетического фотопигмента сетчатки монохромата, демонстрирующий нам *вероятность* поглощения фотонов

как функцию от длины волны. К примеру, у фотона с длиной волны  $\lambda_b$  50% вероятности быть поглощенным, а у фотона  $\lambda_a$  — только 25%.

Количество родопсина в сетчатке многократно превышает количество всех колбочковых пигментов вместе взятых, благодаря чему исследователям удалось выделить его еще в 40-е годы прошлого века. Колбочковые же пигменты были сепарированы и исследованы лишь в самом конце XX столетия, благодаря успехам молекулярной генетики и генной инженерии (Merbs & Nathans, 1992).

Как мы уже показали в разделе 1.3, когда молекула родопсина поглощает фотон, запускается цепь неких специфических биохимических реакций, исходом которых является распад родопсина на две составляющих (опсин и ретиналь [витамин А]), выделение энергии и возникновение потенциала действия.

Британский психолог, профессор кембриджского университета Вильям Раштон, показал, что, при поглощении фотона молекулой фотопигмента, мера воздействия на эту молекулу постоянна и не зависит от длины волны фотона. Очень важно отметить, что поглотив фотон, молекула фотопигмента не кодирует длину волны этого фотона: вся информация о длине волны утрачивается. К примеру, несмотря на то, что фотон с длиной волны 400 нм несет больше энергии, чем фотон с длиной волны в 700 нм, при поглощении и того, и другого родопсином последовательность химических реакций и их характер будут одними и теми же. В своей знаменитой лекции «Зрительные пигменты и цветовые сигналы», прочитанной в Кембридже 17 апреля 1970 г., Раштон озвучил следующий принцип:

«Выходной сигнал рецептора зависит только от факта захвата фотона, но не от того, каков этот фотон».

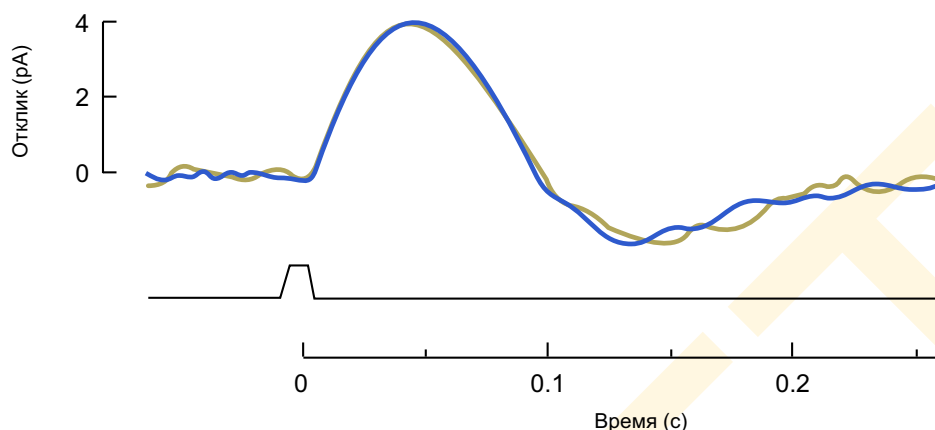
Сей тезис Раштон назвал *принципом унивариантности*, дабы подчеркнуть то, что характер отклика на световое воздействие у данного фотопигмента неизменен (рис. 3.2.14). Итак, мы говорим, что фотопигмент кодирует свет любой длины волны по единственной размерности — *вероятности поглощения*.

Отметим, что отклик единичного фотопигмента никак не может кодировать информацию об относительном спектральном составе света. Сие как раз и объясняет тот факт, почему мы не можем различить световые потоки разных спектральных распределений в скотопических условиях просмотра (Rushton, 1965; Naka & Rushton, 1966). Унивариантность *не* означает, однако, что фотопигмент равновелико откликается на свет разных спектральных составов: он в большей степени поглощает свет одних длин волн и в меньшей — других.

Итак:

⇒ Принцип унивариантности означает, что ежели фотон поглощен пигментом, то какова бы ни была длина волны этого фотона — мера его воздействия на клетку всегда одна и та же.

Вероятность поглощения можно измерить с помощью денситометра на просвет, схема которого показана на рис. 1.3.16 (гл. 1): тонкий слой фотопигмента наносят на идеально чистое предметное стекло; монохроматический свет организуют с помощью т.н. *монохроматора* — устройства в основе которого лежит призма (или дифракционная решетка), расщепляющая обыкновенный белый свет на спектральные составляющие; количество монохроматического света, прошедшего через пигмент и предметное стекло измеряется с помощью светочувствительного сенсора, расположенного в задней части аппарата. По выполнении замера пигмент удаляется и вновь производится



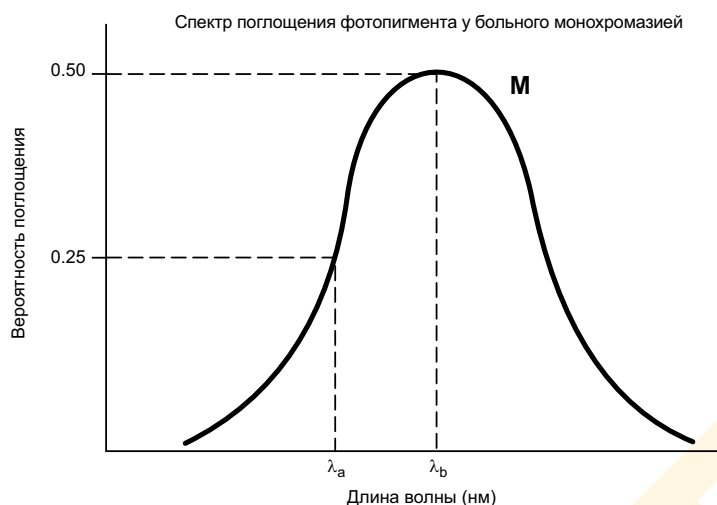
**Рис. 3.2.14** Принцип унитарности утверждает то, что поглощение фотона даст один и тот же нейронный отклик вне зависимости от длины волны этого фотона. Принцип прогнозирует то, что, манипулируя мощностью двух стимулов различных длин волн, можно уравнивать характеристики ответного фототока нейрона. На данной схеме продемонстрировано соответствие двух фототоков при откликах на тестовые монохроматы в 550 нм (желтоватая линия) и 659 нм (фиолетовая линия) и при соотношении их мощностей 9:1 (Wandell, 1995; Baylor et al., 1987).

измерение, но уже чистого светового потока. Разность между сигналом фотозлемента в одних и в других условиях измерения будет пропорциональна величине поглощения света пигментом. Чтобы выделить фракцию фотонов данной длины волны, поглощенных фотопигментом, мы должны разделить число фотонов (этой длины волны), прошедших через стекло и фотопигмент на число фотонов, прошедшее через чистое стекло.

Итак, способна ли особь, наделенная лишь одним видом фотопигмента, к различению стимулов по длинам их волн? Иными словами, ежели мы предъявим такому наблюдателю два стимула — один с длиной волны  $\lambda_a$ , а другой —  $\lambda_b$ , то сможет ли он отличить их? Ответ находим на рис. 3.2.15, где даны два образца, каждый из которых, условного говоря, испускает по 100 фотонов с длинами волн  $\lambda_a$  и  $\lambda_b$ . Поскольку *вероятности* поглощения по каждой из длин волн различны, то количество поглощенных фотонов тоже будет различным. По фотонам с длиной волны  $\lambda_a$  вероятность поглощения составляет 0.25 — следовательно поглотится 25 фотонов. Аналогично в отношении фотонов с длиной волны  $\lambda_b$  — коль скоро вероятность поглощения при такой длине волны 0.5, то поглотится 50 фотонов. Поскольку во втором случае фотонов поглощается вдвое больше, второй образец выглядит ярче. Таким образом особь — монохромат сможет различить эти два стимула по параметру их субъективной яркости.

Теперь давайте подумаем, что произойдет, ежели мы удвоим фотометрическую яркость образца, испускающего свет с  $\lambda_a$ , то есть дающего 200 фотонов в единицу времени, а интенсивность свечения второго образца ( $\lambda_b$ ) оставим прежней? Как видно по рис. 3.2.15 сие приведет к поглощению 50 фотонов по обоим длинам волн. Поскольку оба образца в этом случае окажут *одинаковое воздействие* на зрительную систему, они станут неотличимы друг от друга. Приведенный пример демонстрирует то, что особь, вооруженная лишь одним пигментом, не в состоянии различать стимулы по длинам доминирующих в них волн, поскольку всегда имеется возможность уравнивать результирующие перцепты, управляя интенсивностями этих стимулов.

Итак, хотя мы и знаем теперь, что особи — монохроматы не способны различать



$\lambda_a$		$\lambda_b$	
Кол-во испущенных фотонов	Кол-во поглощенных фотонов	Кол-во испущенных фотонов	Кол-во поглощенных фотонов
100	25	100	50
200	50	100	50

**Рис. 3.2.15** В восприятии больного, страдающего монохроматизацией, оба стимула окажутся неотличимы друг от друга, если их мощности подобрать так, что количество фотонов, поглощенных единственным фотопигментом (в данном примере палочковым пигментом родопсином), окажется одинаковым (нижний ряд таблицы).

лые объекты красными, а яркие зелеными. Однако подчеркнем, что заучивание происходит только по критерию субъективной яркости. Когда же больной монохроматизацией сталкивается с предметами естественного происхождения, скажем яблоками или бананами, задача несколько упрощается за счет наличия прочих признаков, таких как форма и фактура.

**Дихроматическое уравнивание.** Теперь посмотрим на графики рис. 3.2.16 А: мы видим, что у некоего индивида наличествуют два фотопигмента — М и L, с пиками поглощения на разных длинах волн и наложением графиков поглощения по большей части спектра. Ниже мы покажем, что такой индивид в состоянии различать стимулы по длинам их доминант (или непосредственно по длинам волн, если речь о стимулах

стимулы по их волновым доминантам, тем не менее, такие люди могут очень четко поименовать перцепты этих стимулов. Рассмотрим случай палочковой монохроматизации, т.е. тот редкий вид патологии, когда в сетчатке наличествуют лишь палочки.<sup>1</sup> Представим, что такому больному предъявлены три образца, каждый из которых в единицу времени испускает по 5000 фотонов и сказано вдобавок, что один образец голубой, второй — желтый, а третий — красный. Больной палочковый монохроматизацией стимул с доминантой на 505 нм воспримет ярким, с доминантой на 570 нм — менее ярким и на 600 нм — тусклым и, главное, в дальнейшем не допустит ошибки в выборе образцов.

Что при этом происходит? Коротко ответ звучит так: больной монохроматизацией именуется окраску предметов, ориентируясь на их субъективную яркость. В течение жизни он подмечает, что люди именуют туск-

<sup>1</sup> Напомним, что пик поглощения палочкового фотопигмента (родопсина) приходится на 507 нм (см. гл. 1).



монохроматических). Предположим, что наблюдателю предъявили два монохроматических стимула с длинами волн  $\lambda_a$  и  $\lambda_b$  и просят, управляя интенсивностью второго стимула ( $\lambda_b$ ), уравнивать по зрительному ощущению оба образца. Если наблюдатель сможет выполнить такое уравнивание, значит он не различает эти стимулы по длинам их волн (но только по интенсивности), а ежели не сможет, то это значит, что он обладает способностью к различению стимулов исключительно по доминирующей длине волны.

Из рис 3.2.16 А: следует также, что уравнивать образцы, манипулируя интенсивностью второго стимула ( $\lambda_b$ ) — невозможно, поскольку нет такого уровня интенсивности излучения, при котором число фотонов, поглощенных М- и L-пигментами по второму стимулу ( $\lambda_b$ ) уравнивается с числом фотонов, поглощенных теми же пигментами по первому стимулу ( $\lambda_a$ ). И как бы ни менялась интенсивность обоих стимулов, они никогда не окажут на зрительную систему одинакового воздействия: они будут неизбежно различны.

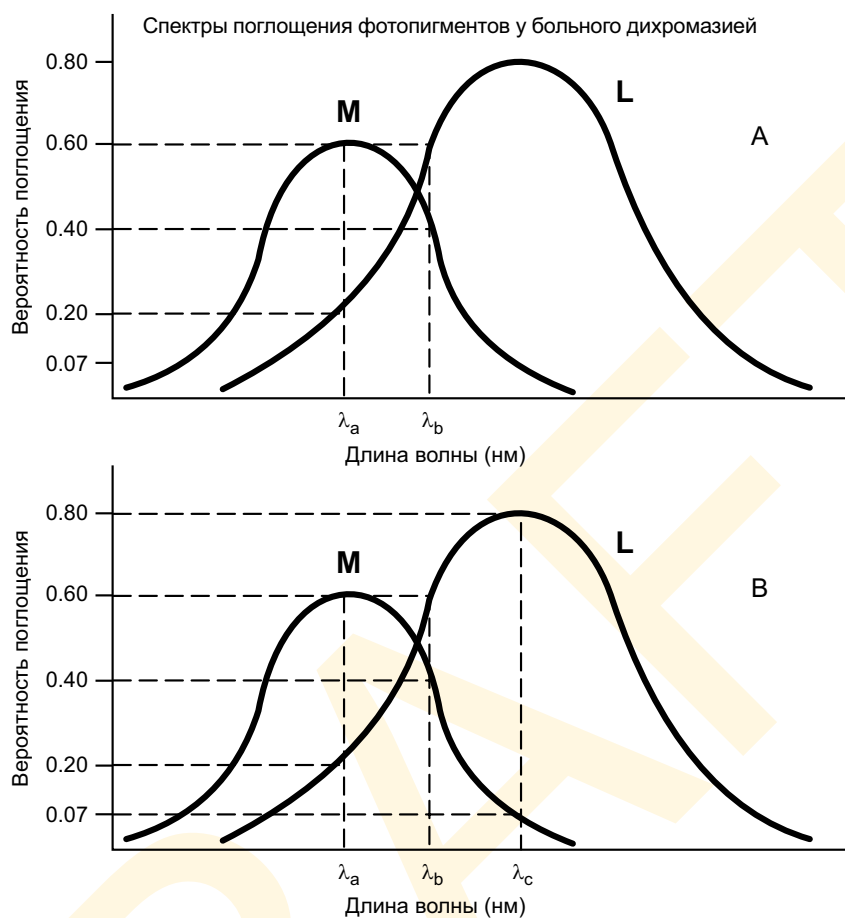
Теперь рассмотрим ситуацию, представленную рис. 3.2.16 В вновь имеются два монохроматических стимула — один с длиной волны  $\lambda_b$ , а другой — представляющий собой аддитивную смесь двух монохроматов с длинами волн  $\lambda_a$  и  $\lambda_c$ . Наблюдателю дана возможность отдельно менять интенсивности составляющих всех трех компонент, к примеру, можно изменить мощность стимула по  $\lambda_c$ , не затронув составляющие  $\lambda_a$  и  $\lambda_b$ . Итак, сможет ли наблюдатель, варьируя интенсивностями потоков по всем трем длинам волн, уравнивать образцы по зрительному ощущению? То есть, можно ли подобрать эти интенсивности так что, что М- и L-пигменты из первого и второго стимула заберут одинаковое число фотонов? Тщательный анализ рис. 3.2.16 В свидетельствует о том, что сие действительно возможно, то бишь о том, что наблюдатель, действительно может уравнивать воздействие этих стимулов на свое зрение (по каждому стимулу М-пигмент поглотит 200 фотонов, а L-пигмент — 300). С позиции наблюдателя это значит, что оба стимула, физически совершенно разные, будут выглядеть абсолютно одинаково.

Напомним, что описанный индивид — больной, т.н. *дихромат*. Люди, страдающие дихромазией, в какой-то мере способны различать стимулы по длинам волн, хотя и неполноценно. В их восприятии стимул  $\lambda_a$  будет всегда отличен от стимула  $\lambda_b$ , но последний можно уравнивать подходящей смесью  $\lambda_a$  и  $\lambda_c$ . Иными словами, дихромат способен выполнять визуальное уравнивание некоего монохроматического стимула с помощью *двух* других монохроматов.

**Трихроматическое уравнивание.** Доподлинно известно, что у большинства из нас зрение трихроматическое. Теоретический концепт, лежащий в основе трихромазии тот же, что и у дихромазии, но с тем лишь отличием, что в сетчатке присутствуют пигменты трех (а не двух) типов с перекрывающимися графиками спектров поглощения (рис. 3.2.17). Спектры эти были выведены на основе множества экспериментов.

Мы не будем проводить детальный анализ способности трихроматов к различению стимулов по длине волны, но воспользуемся принципами, выработанными нами при разговоре о дихромазии. Мы сразу увидим при этом, что способность трихроматов к селекции стимулов по длине волны намного выше, чем у дихроматов, но при этом по-прежнему неабсолютна.

Давайте определим, что есть трихромазия. Возьмем два стимула, состоящих из смеси двух монохроматов каждый (рис. 3.2.18): индивид с нормальным трихроматическим зрением — *нормальный трихромат* — манипулируя интенсивностями компонентов смеси, способен уравнивать оба стимула по зрительному ощущению. То есть, образцы будут восприниматься идентичными, поскольку дают равновеликое количест-



$\lambda_a$ и $\lambda_c$				$\lambda_b$			
Кол-во испущенных фотонов	Длина волны	Кол-во поглощенных фотонов		Кол-во испущенных фотонов	Длина волны	Кол-во поглощенных фотонов	
		М	Л			М	Л
300	$\lambda_a$	180	60	500	$\lambda_b$	200	300
300	$\lambda_c$	20	240			200	300
ИТОГО:		200	300	ИТОГО:		200	300

**Рис. 3.2.16** Манипулируя интенсивностями стимулов  $\lambda_a$ ,  $\lambda_c$  и  $\lambda_b$  можно добиться визуального равенства между смесью  $\lambda_a$  с  $\lambda_c$  и стимулом  $\lambda_b$  в восприятии больных дихромазией, поскольку и смесь, и стимул  $\lambda_b$  окажут на зрительную систему такого больного одинаковое воздействие. Каждая из половин поля сравнения даст 200 поглощений по хлорофиллу (М-пигмент) и 300 по эритрофиллу (L-пигмент).

во поглощенных фотонах каждым из трех фотопигментов, следовательно оказывают на зрительную систему одинаковое воздействие. Такие стимулы, напомним, именуются метамерами.

Трехволновое уравнивание трихромат выполнить не может: образцы, которые в восприятии дихромата совершенно одинаковы, в восприятии трихромата окажутся разными, что наглядно демонстрирует большую способность трихроматов к селекции стимулов по длинам их волн. Однако трихромат успешно выполнит уравнивание по четырехволновому набору и сие демонстрирует то, что и у него есть некие пределы в способности различать стимулы по длинам волн их доминант.

### 3.2.2.2 Оппонентная теория

Несмотря на очевидный успех трихроматической концепции имеют место факты и наблюдения необъяснимые ею. В частности, вся феноменология цветового восприятия — природа человеческих цветовых ощущений. Однако множество из феноменов цветового восприятия сумел объяснить немецкий физиолог Карл Эвальд Константин Геринг (см. рис. 3.2.19), с помощью своей *оппонентной теории цветового восприятия* или как ее еще называют — *теории оппонентной цветовой обработки* (Hering, 1878/1964).

#### Исдержки трихромазии.

Возможно наибольшим феноменологическим просчетом трихроматической концепции является то, что (хоть она и блестяще толкует все три вида дихроматической цветовой слепоты) она не дает объяснения тому, почему цветовые ощущения разделены на несовместимые пары: не существует красно-зеленых ощущений или же сине-желтых. При цветовой

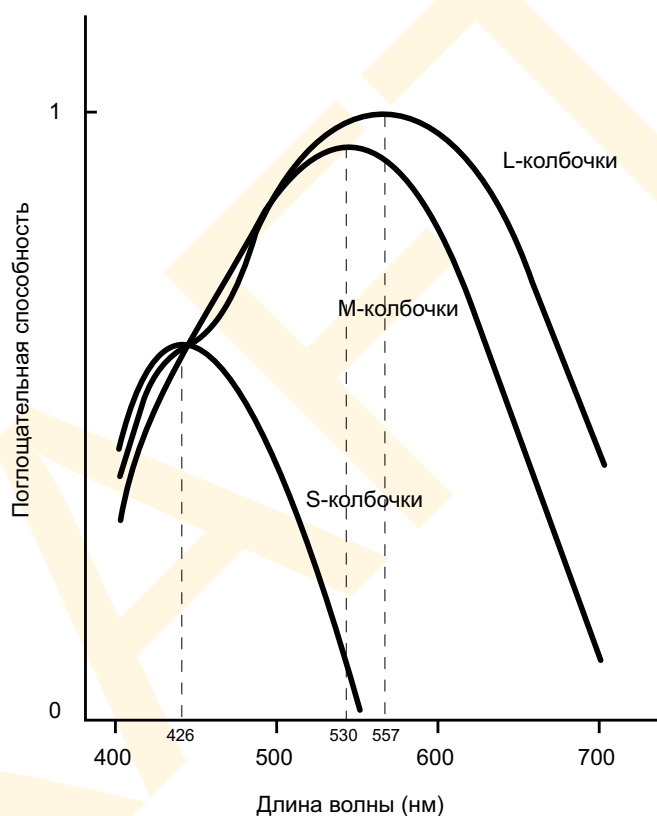


Рис. 3.2.17 Примерные спектры поглощения трех колбочковых пигментов человека, полученные Смитом и Покорни (1975).

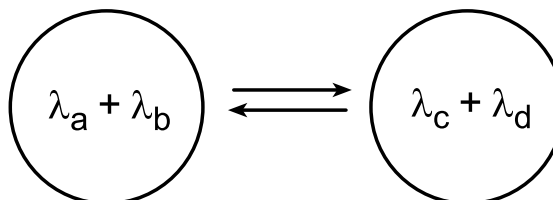


Рис. 3.2.18 Управляя интенсивностями четырех монохроматов, разбитых на две пары, нормальный трихромат способен уравнивать обе смеси по зрительному ощущению. Пара смесей создает два стимула совершенно разных физически, но оказывающих на зрительную систему одинаковое воздействие, то есть — метамерную пару.

слепоте ощущения никогда не выпадают единично (только красный, только синий или только зеленый) — как можно было бы вывести из трихроматической теории — но всегда попарно, например красный вместе с синим, зеленый вместе с желтым. Как видим — это важные моменты, которым трихроматическая концепция не в силах дать объяснения.

Еще одной загадкой, неразрешимой силами трихроматического концепта, является феноменология кардинальных стимулов и ощущений ими вызываемых. Юнг и Гельмгольц предположили существование лишь трех кардиналов: красного, зеленого и синего — сообразно пикам чувствительности предполагаемых рецепторов. Эти стимулы должны восприниматься «чистыми», то есть не должны выглядеть «смесью» других, как скажем оранжевый, воспринимаемый как «смесь» красного и желтого, или пурпурный — как смесь синего и красного.<sup>1</sup> В рамках трихроматической концепции желтый стимул не может являться кардинальным, потому что толкуется как смесь красного и зеленого кардиналов.

Недостаток трихроматического толкования желтого в том, что оно не вяжется с феноменологическими фактами: субъективное ощущение желтого никак не воспринимается как смесь красного и зеленого, подобно тому как, скажем, пурпурный воспринимается смесью красного и синего. Психологически желтый оценивается как кардинальный наравне с красным, зеленым и синим. От этого факта не отпереться, но трихроматическая теория и здесь не в состоянии дать внятного объяснения.

Сходные наблюдения касаются феноменологии восприятия световых смесей. Даже то, что трихроматическая теория в состоянии объяснить механизм уравнивания с помощью этих смесей, она не может истолковать невозможность объединения ряда ощущений. Как мы давеча говорили существуют стимулы, воспринимаемые одновременно красными и синими (пурпурные), одновременно красными и желтыми (оранжевые), одновременно синими и зелеными (голубые) и одновременно желтыми и зелеными (желто-зеленые). Сказанное касается любого парного сочетания описанных цветов за исключением двух: не существует цвета, который являлся бы субъективной «смесью» красного и зеленого, а также не существует цвета, «смешанного» из желтого и синего (сознательно абстрагируемся от опыта смешивания зеленой *краски* из голубой и

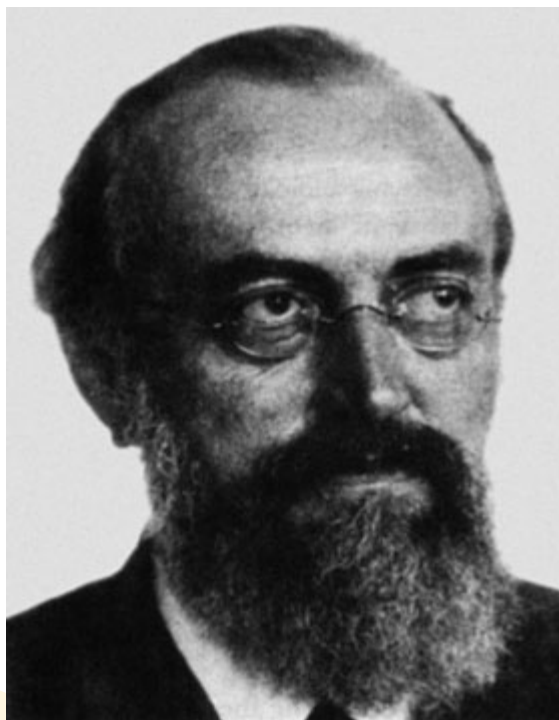


Рис. 3.2.19 Карл Эвальд Константин Геринг — основоположник оппонентной теории цветового зрения.

<sup>1</sup> Многие думают, что зеленый — это смесь голубого и желтого, но такая мысль приходит лишь из опыта смешения красок. Безусловно, существуют желто-зеленые и сине-зеленые оттенки, но для людей, не имеющих опыта работы с красками, существуют и чистые зеленые, воспринимаемые как не содержащие примеси желтизны или голубизны.

желтой). Подобно тому, как при цветовой слепоте цветовые ощущения выпадают парно, сие наблюдение указывает на диаметрально противоположность красного зеленому и синего — желтому.

Безусловное подтверждение присутствия оппонентных механизмов в цветовом восприятии появилось тогда, когда шведский физиолог Гуннар Свэтичин (1956), изучавший горизонтальные клетки в сетчатке рыб, не обнаружил, что клетки эти активируются при стимуляции сетчатки световыми потоками одних длин волн и подавляются потоками других. Спустя время оппонентные клетки обнаружились в сетчатке приматов, а также латеральном колленчатом теле их центральной нервной системы (DeValois et al, 1966; Wiesel & Hubel, 1966; Gouras, 1968).

Еще одно доказательство цветовой оппонентности находим в хроматических постообразах и наведенных оттенках (и при этом вновь сталкиваемся с несостоятельностью трихроматического толкования этих феноменов).

На основе этих (и ряда других) наблюдений Геринг убедительно показал, что существует четыре (а не три) первичных реакции, организованных в оппозиционные пары: красный против зеленого и синий против желтого. То есть, ученый сформулировал альтернативную теорию цветового зрения, положив в ее основу оппонентный принцип. Однако же Геринг принимал гельмгольцеву идею о первопричинности трех примитивных механизмов, но не соглашался с нею в частности. Мы говорим при этом, что геринговы механизмы *оппонентны* (см. рис. 3.2.20):<sup>1</sup> один откликается по красно-зеленой размерности, второй — по сине-желтой, а третий — по черно-белой (ахроматический отклик). Геринг полагал, что сие реализовано на рецепторном уровне, и что причина оппонентности в дуализме образования и распада ретиальной органики. К примеру, некие химические соединения, образуясь, вызывают красное ощущение, а распадаясь — зеленое.

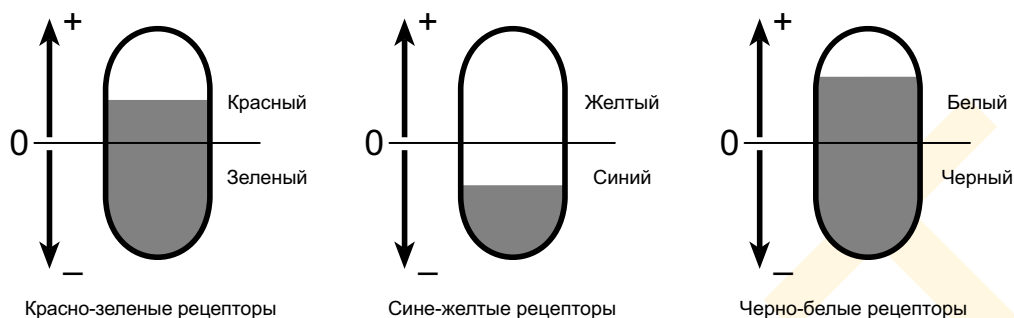
Оппонентная теория Геринга четко согласуется с феноменологией фактов, на которых она строится и объясняет физиологическую первичность красного, зеленого, синего и желтого тонов. Сходным образом оппонентная теория толкует выпадения по красно-зеленой и сине-желтой размерностям цветового пространства при цветовой слепоте, поскольку каждая пара в оппонентной логике — это часть единого механизма. Геринг полагал также, что комплементарные последствия возникают из-за того, что при повышенной активности в одном элементе системы (к примеру, в красном) истощается подлежащий ему биологический механизм и благодаря этому баланс смещается в сторону оппонентного элемента (зеленого). Появление наведенных оттенков при хроматическом симультанном контрасте можно истолковать аналогично, но в пространственном домене, а не темпоральном (временном).

Итак мы видим, что герингова теория весьма изящным и экономичным образом объясняет многие феномены, перед которыми бессильна теория трихроматическая.

**Теория двухэтапного цветового зрения.** Несколько десятилетий кряду между двумя научными группировками шли жаркие дебаты: сторонники Гельмгольца апеллировали к способности трихроматической теории прогнозировать объективные физиологические эффекты и яростно критиковали герингов подход к проблеме. Сторонники

<sup>1</sup> Несмотря на то, что красный/зеленый и синий/желтый это подлинно оппонентные размерности, в отношении черно-белой оси все не так просто: одним важным отличием является то, что красно-зеленых и сине-желтых оттенков не существует, тогда как серые — это субъективная «смесь» черного с белым. То есть, в отношении ахроматической размерности восприятия имеет место некоторое качественное отличие.





**Рис. 3.2.20** Герингова оппонентная теория цветового зрения. Геринг предположил, что рецепторы трех типов могут отвечать на стимуляцию по двум взаимнопротивоположным направлениям от нейтральной точки.

Геринга пеняли на слабые места трихромазии. Выходом из тупика явилось появление *теории двухэтапного цветового зрения*, часто именуемой также *теорией двойной обработки*, которая красивым и совершенно неожиданным посылом разрешила спор Гельмгольца — Геринга: **обе теории верны, но описывают разные этапы зрительного восприятия**. Основоположником этой идеи явился немецкий физиолог Иоганнес Адольф фон Криз (1905), но свою окончательную теоретическую форму концепция обрела только в 1957 г., когда американские физиологи Лео Гурвич и Доротея Джеймсон (Пенсильванский университет) предложили точную и, как говорят, *контролепригодную* формулировку, основанную на тезисе о двух последовательных этапах цветового восприятия.

Гурвич и Джеймсон подвергли геринговы оппонентные процессы количественному анализу и показали, что величину «синевы» в любом световом потоке можно измерить, смешав этот поток с оппозитным ему по восприятию желтым потоком, в результате чего результирующий стимул окажется ни синим, ни желтым. Начав с самых коротких волн видимого спектра, исследователи вывели то, сколько нужно желтого света для нейтрализации («обнуления») синевы в восприятии стимула исследуемого. Когда сию процедуру применили к потокам последовательно растущих длин волн, образовалась система измерения синевы по всем монохроматам спектра (см. рис. 3.2.21А). Оказалось также, что монохромат с длиной волны примерно 500 нм воспринимается однозначно зеленым и не требует желтой добавки. Свыше 500 нм свет желтеет, и ему требуется добавка синего света, чтобы «нейтрализовать» эту желтизну.

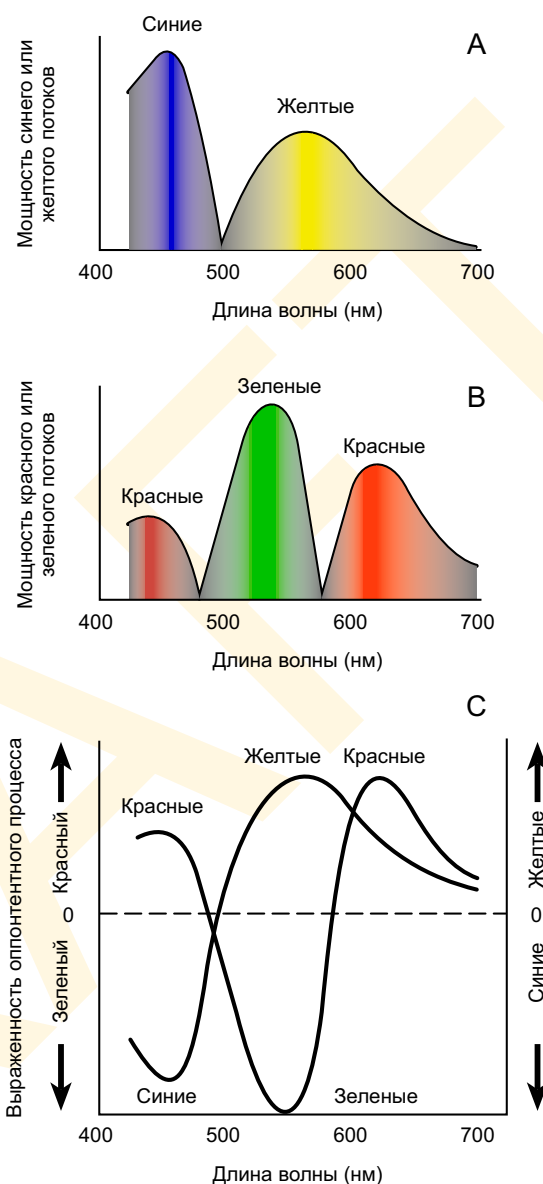
Исследователи проделали эту же манипуляцию с красными и зелеными стимулами, «нейтрализуя» их красноту потоками различных длин волн (рис. 3.2.21В). Отметим, что ощущения, вызванные монохроматами концов спектра, содержат в себе выраженный красный компонент, то есть позиции этих ощущений в цветовом пространстве стремятся друг к другу. В единый график данные Гурвича и Джеймсон сведены на рис. 3.2.21С, где оппонентность красно-зеленого и сине-желтого механизмов эксплицитно представлена за счет биполярности оси ординат и где каждой из пар принудительно задано положительное направление.

Приняв количественную формулировку оппонентной теории и ее взаимосвязь с теорией трихроматической, Гурвич и Джеймсон продемонстрировали то, что теории Гельмгольца и Геринга вовсе не противоречат друг другу (если, конечно, проигнорировать утверждение Геринга о том, что оппонентные механизмы реализуются рецепто-

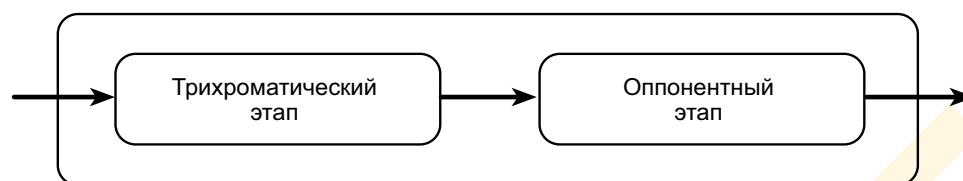
рами). Путем столь неожиданного синтеза исследователи показали, как, разделив цветовое восприятие на два этапа, можно объяснить и увязать между собой множество противоречивых фактов. По идее ученых гельмгольца трихроматизм — суть первый этап, выходные данные которого образуют вход второго этапа, являющегося собой тройку геринговых оппонентных механизмов (см. рис. 3.2.22). Такой подход объясняет не только оба набора феноменов, но также и собственные изыскания Джеймсон и Гурвич, в частности количественные измерения оппонентной структуры цветовых ощущений.

Сегодня доподлинно известно, что оба этапа (трихроматический и оппонентный) реализуются на уровне сетчатки, но важно показать, что рассуждения Гельмгольца, Геринга, Гурвича и Джеймсон о механизмах цветового зрения прозвучали *до того*, как было получено физиологическое подтверждение (что, кстати, не редкость в исследовании восприятия). Теории, основанные на сугубо поведенческом анализе часто (но не всегда) предшествуют знаниям о подлежащих физиологических механизмах. Одним критичным фактором в цепи событий, очевидно, является хронология экспериментальных технологий: поведенческий анализ выполняется с помощью психофизических и феноменологических методов, на разработку которых ушло больше века, но при этом соответствующим физиологическим методикам всего несколько десятилетий (и непрерывно появляются новые).

Наконец отметим, что существует еще один фактор, относящийся к марровским трем уровням информационной обработки (1982) и пальмеровско-кимчевской рекур-



**Рис. 3.2.21** Эксперимент Гурвич и Джеймсон по "обнулению" цветового тона. Линия В демонстрирует количества красного или зеленого "обнуляющих" потоков, необходимые для нейтрализации перцептов зеленого или красного соответственно. Линия А показывает требуемые количества синего или желтого потоков, необходимые для нейтрализации соответственно желтого и синего перцептов. С — обе кривые в оппонентном представлении (Hurvich & Jameson, 1957).



**Рис. 3.2.22** Двухэтапная теория цветового зрения, предложенная Гурвич и Джеймсон, в которой трихроматический этап (по Гельмгольцу) поставляет входные данные этапу оппонентному (по Герингу).

сивной декомпозиции (1986): дело в том, что всегда легче спускаться с более абстрактного функционального уровня к физическому исполнению, нежели идти в обратном направлении. С алгоритмическим принципом в руках (позволяющим декомпонировать систему до квази-нейронного уровня) физиологи четко понимают, что именно нужно искать на этапах восприятия и как интерпретировать полученные данные. Сами биологические данные можно собрать и без опоры на функциональную теорию, но при этом их научное значение так и останется непонятым.<sup>1</sup> Итак, первыми появились функциональные теории именно цветового зрения, и у физиологов в наличии давно были все необходимые инструменты и исследователи ясно представляли себе, что искать.

### 3.2.3 Физиологические механизмы

Анатомические и физиологические методы, позволяющие раскрыть суть процессов обработки информации, декодируемой зрением из длин волн стимулов разрабатывались в 50-х годах прошлого века. Как только они появились, так исследователи сразу начали поиск доказательств гельмгольцевой трихроматической теории на ретинальном уровне. При этом вопреки геринговым аргументам и невзирая на примирительный концепт Гурвича — Джеймсон, время было таким, что упорно искали доказательство работы оппонентных механизмов в высших отделах нервной системы, но не в сетчатке как таковой. Но одно из больших достоинств науки в том, что экспериментальный результат побуждает исследователей менять свои взгляды на действующие теории.

#### 3.2.3.1 Три колбочковых системы

В первой главе мы коротко рассмотрели то, что фоторецепторы, отвечающие за цветное зрение — это именно колбочки, но не палочки. То есть — как и предрекали Г. Пальмер, Юнг и Гельмгольц — в сетчатке нормальных трихроматов имеются три типа колбочек, каждая из которых содержит один из трех вариантов светочувствительного пигмента (цианолэйб, хлоролэйб и эритролэйб). С применением определенных методик удалось получить спектры поглощения этих пигментов, что позволило узнать долю поглощения света по каждой из длин волн видимого спектра (Brown & Wald, 1964; Marks, Dobelle & MacNichol, 1964). В результате получили три накладывающиеся друг на друга функции, показанные на рис. 3.2.17: пик поглощения *коротковолновочувствительных* (S-) колбочек приходится примерно на 426 нм,<sup>2</sup> *средневолновочув-*

<sup>1</sup> Мы столкнемся с изумительным примером такой ситуации в гл. 4, когда будем говорить о корковых клетках V1, изученных Хьюбелом и Визелем.

*ствительных (М-)* — примерно на 530 нм, а *длинноволновочувствительных (L-)* — на 557 нм (Schnapf, Kraft, & Baylor, 1987). Напомним, что количество колбочек каждого из типов и их пространственное распределение (заселенность) по сетчатке неравномерны: психофизические данные указывают на то, что S-колбочки составляют всего лишь 5-10% от общей популяции, и что L-колбочек примерно вдвое больше, чем колбочек М. То есть, соотношение L:M:S составляет примерно 10:5:1 (De Valois & De Valois, 1993). Более того, в самом центре фовеа S-колбочек нет вообще, а на угловом удалении от этого центра примерно на 1° их доля составляет порядка 6%.

Колбочки работают примерно так, как и представлял себе Гельмгольц, разве что отличаются спектры поглощения. Как и предполагал ученый, у лиц, страдающих цветовой слепотой в сетчатке отсутствуют один, либо даже два типа колбочек: у протанопов нет L-колбочек, у дейтеранопов — М-колбочек, у тританопов — колбочек S.<sup>1</sup> У аномальных трихроматов работают колбочки всех трех типов, но при этом содержание либо хлоролэйба, либо эритролэйба (либо обоих) отлично от такового у нормальных трихроматов. Иногда спектр поглощения хлоролэйба (содержится в М-колбочках) сдвигается в сторону эритролэйба (содержится в L-колбочках), а иногда наоборот (Neitz, Neitz, & Jacobs, 1993). В результате, у аномальных трихроматов пропорции кардиналов при визуальном уравнивании тестовых стимулов трихроматическими смесями оказываются отличными от таковых у трихроматов нормальных, а мера и характер этого отличия оказываются зависимыми от содержания и/или спектра того или иного пигмента.

**Экстраклеточная регистрация.** На сегодняшний день наиболее четкие данные о характере поглощения света колбочковыми фотопигментами получены путем измерений т.н. *колбочковых фототоков*, то есть, изменений силы электрического тока через мембрану колбочки при ее стимуляции светом. Отношение величины фототока к величине поглощения требует крайне тщательного анализа, поскольку зависимость этого тока от степени поглощения нелинейна. В данном разделе мы рассмотрим новые подходы к интерпретации результатов таких измерений и увяжем их с линейным характером поглощений колбочковыми пигментами.

Первыми, кто измерил колбочковые фототоки в сетчатке макаки были американцы Д. Бэйлор, Б. Нанн и Дж. Шнапф (Baylor, 1987). У макак три типа колбочек и поведение этих животных при решении большинства цветовых задач весьма сходно с поведением человека (DeValois et al., 1974).

Для измерения фототока Бэйлор, Нанн и Шнапф извлекали сетчатку из глаза обезьяны и разрезали ее на мелкие кусочки порядка 100 мкм в ширину, после чего для поддержания метаболизма помещали в физиологический раствор (благодаря чему электрический отклик рецепторов оставался весьма выраженным в течение нескольких часов). Собственно замер производился посредством метода, показанного на рис. 3.2.23: в стеклянную микропипетку с внутренним диаметром примерно 2-6 мкм (что всего лишь в десять раз больше длины волны видимого света) втягивался внешний сегмент фоторецептора из фрагмента ткани сетчатки, а стимуляция проводилась тонким лучом (белая полоска на фотографии).

<sup>2</sup> По некоторым данным на 440 нм.

<sup>1</sup> На самом деле у дихроматов общее число колбочек такое же, как и у нормальных трихроматов, о чем свидетельствуют клинические исследования протанопов и дейтеранопов (Cicerone & Nerger, 1989), но при этом либо хлоролэйб (в норме содержащийся в М-колбочках), либо эритролэйб (L-колбочки) заменен на один из оставшихся двух пигментов.

На рис. 3.2.24 показан результат стимуляции фоторецептора коротким световым импульсом: кривая представляет собой график мембранного фототока, возникающего после стимуляции световой вспышкой. Кривые на рис. 3.2.24 А — это отклики на монохроматический стимул с длиной волны 500 нм в определенном диапазоне интенсивностей.

На рис. 3.2.24 В дано то же, но по монохромату с длиной волны 659 нм.

До стимуляции имел место устойчивый внутренний ток положительно заряженных натриевых ионов внутрь фоторецептора через ионные каналы в клеточной мембране. Такой ток (именуемый *темновым током*), представляет исходный (базовый) уровень и поэтому принимается за ноль. Отметим, что величины на графике *двухфазны*, то есть в какой-то момент меняют свой знак.

Когда пигмент поглотил фотон, ток натриевых ионов внутрь клетки начинает падать. В темноте натриевый ток понижает отрицательную электрическую поляризацию внутренней поверхности клетки, но когда свет блокирует натриевый

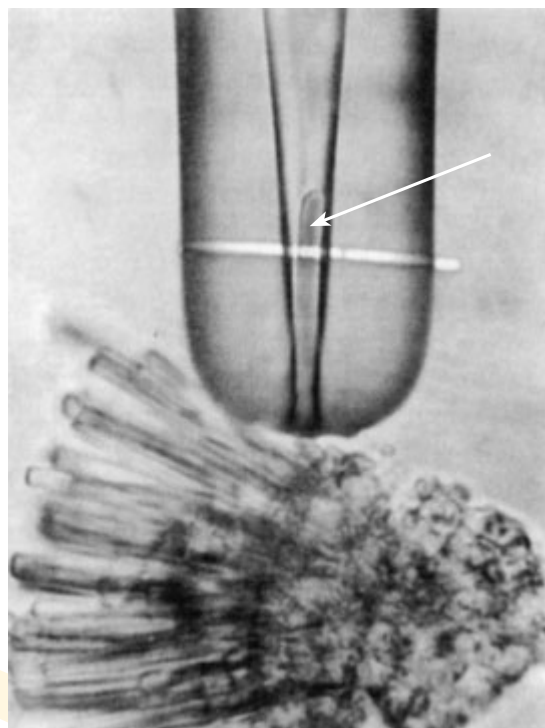


Рис. 3.2.23 Единичный рецептор (отмечен стрелкой) втянутый внутрь микропипетки. Белая полоска — монохроматический стимул. Внизу — фрагмент сетчатки.

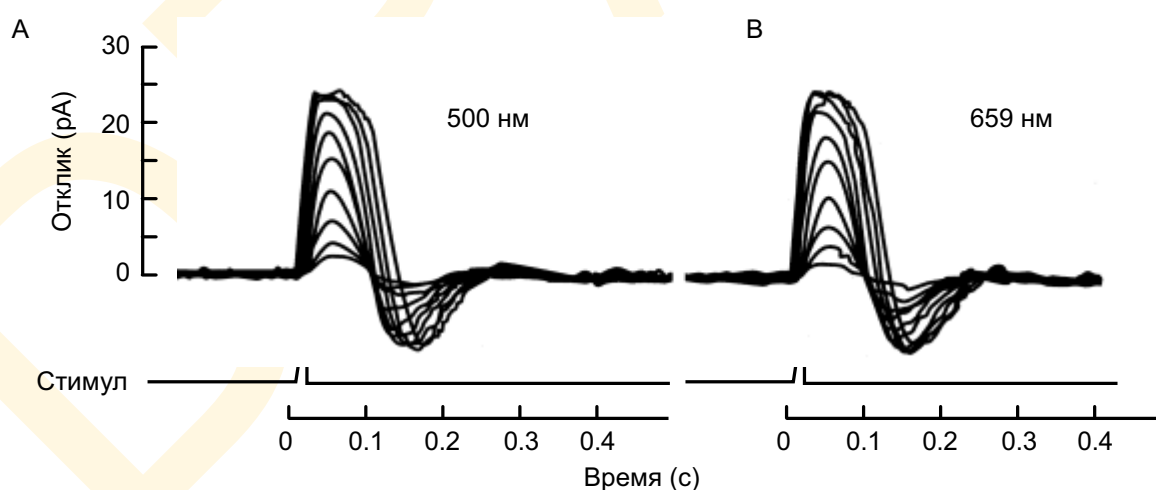


Рис. 3.2.24 Колбочковый фототок, возникающий в ответ на короткую стимуляцию, *бифазен*. Амплитуда этого тока растет по мере роста интенсивности стимуляции. Функция отклика на монохроматы разных длин волн одинакова: А — отклик на стимуляцию монохроматом в 500 нм; В — отклик на стимуляцию монохроматом в 659 нм.



ток, разность отрицательных зарядов между внутренней стороной мембраны и ее внешней стороной растёт. То есть, исходный отклик фоторецептора на световую стимуляцию — это *гиперполяризация*. После начальной блокады натриевого тока, ток этот активно восстанавливается, причем с избытком (т.н. *гиперкомпенсация*). В течение второй фазы рецепторного отклика общий фототок меняет направление: в начале отклика он идет в одну сторону, а затем в обратную, то есть, мы говорим, что результирующий отклик двухфазен.

Колбочковые спектры, полученные таким методом, оказались весьма схожи с результатами психофизических экспериментов по трихроматическому уравниванию.<sup>1</sup>

**Молекулярная генетика колбочковых фотопигментов.** Молекула колбочкового пигмента (независимо от его типа) состоит из двух органических соединений — хромофора и опсина. Хромофор (который един у всех колбочковых фотопигментов) — суть альдегидное производное ретинола (витамина А), именуемое *ретиналем*. Фотон, поглощенный хромофором, инициирует серию каскадных биохимических реакций, исходом которых является сигнал, идущий от колбочки к ретинальным нейронам следующего уровня. Опсин — зрительно инертная цепь аминокислот, залегающая между мембранами дисков внешнего сегмента и определяющая спектр поглощения молекулы фотопигмента. У каждого класса колбочек свой опсин.

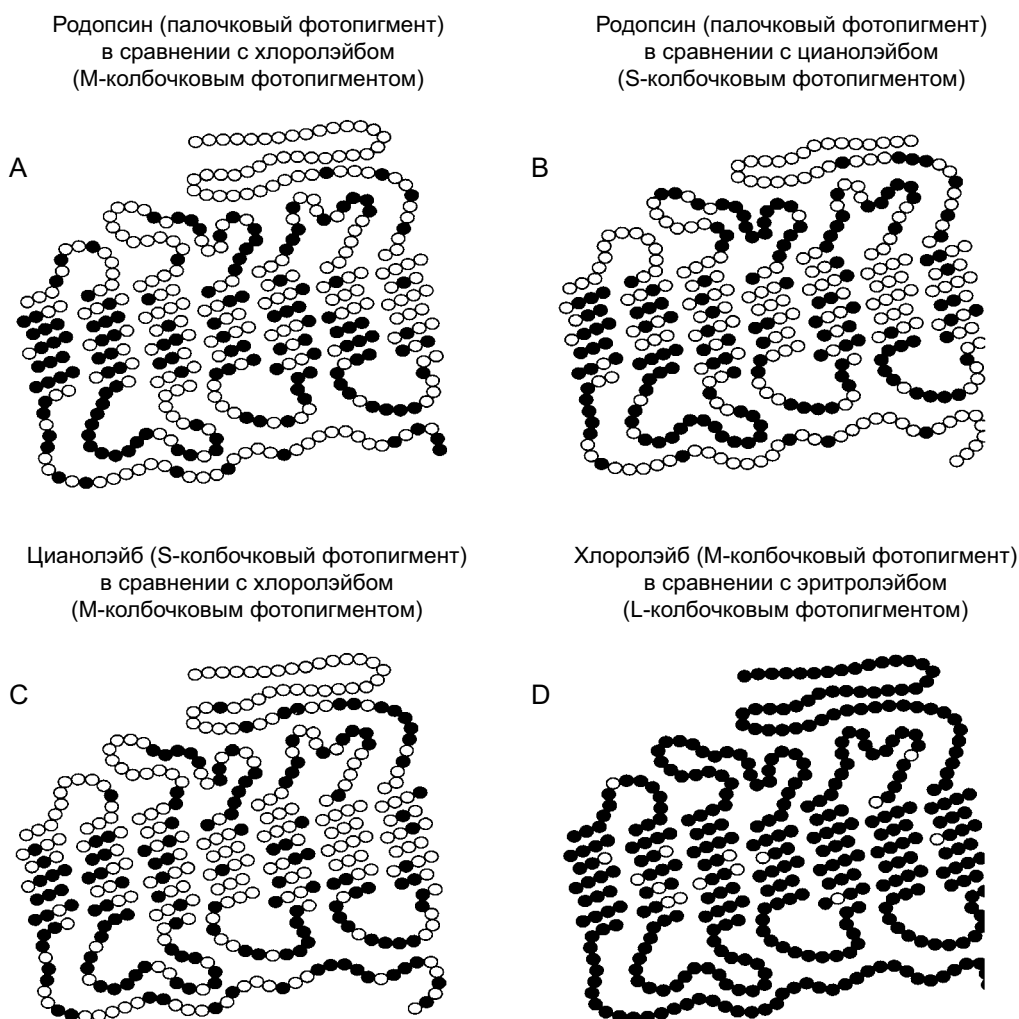
Гены, кодирующие молекулярную структуру колбочковых фотопигментов человека и выделены американским молекулярным биологом Джереми Натансом с помощью т.н. *метода рекомбинантных ДНК* (Nathans et al, 1986). Гены, кодирующие хлоролэйб и эритролэйб (М- и L-колбочковые пигменты соответственно), обнаружены, в X-хромосоме, что вполне закономерно, поскольку давно известно, что цветовые аномалии, при которых выпадает (или поврежден) либо М-, либо L-колбочковый пигменты, наследуются с полом. Ген, кодирующий цианолэйб (S-колбочковый пигмент), локализуется в 7-й хромосоме, а ген, кодирующий родопсин — в 3-й.

Гомология генов колбочковых фотопигментов гену родопсина указывает на то, что все четыре гена происходят от одного предшественника: М- и L-колбочковые опсины сильно похожи друг на друга, демонстрируя 98% гомологии (т.е. их ДНК-последовательности идентичны на 98%). Гомология S-колбочкового опсина М- и L-колбочковым опсином составляет 40%. Стало быть, ген, кодирующий S-опсин, выщепился из предшественника намного раньше (рис. 3.2.25).

В X-хромосоме обычно имеется одна копия гена L-колбочкового опсина и несколько копий М-колбочкового. При этом множественность копий М-колбочкового опсина никак не влияет на результат трихроматического уравнивания, а роль дополнительных генных копий остается пока невыясненной (Knau et al, 2002).

В продолжение своей воистину новаторской работы по молекулярной генетике цветового зрения Натанс определил спектры поглощения колбочковых фотопигментов человека путем т.н. *трансфекции*<sup>2</sup> специальной клеточной культуры генами фотопигментов (Merbs & Nathans, 1992). Трансфицированная клеточная культура выработала все три пигмента в количестве, достаточном для их заурядной макроспектрометрии. Пик поглощения S- и М-колбочковых пигментов пришелся на 426 нм и 530 нм соответственно. L-пигмент в человеческой популяции оказался представленным в двух вариантах: с пиком на 552 нм и на 557 нм. Обнаружение двух вариантов L-колбочкового фо-

<sup>1</sup> Детальное описание экспериментов по трихроматическому уравниванию см. Р.В.Г. Хант, Цветовоспроизведение, 2004, а также А. Френкель, А. Шадрин, Колориметрическая настройка монитора. Теория и практика, М., Август-Борг, 2005 г. — *Прим. пер.*



**Рис. 3.2.25** Сравнение аминокислотных цепочек (выведенных из нуклеотидных последовательностей): родопсина и хлоролэйба (A), родопсина и цианолэйба (B), цианолэйба и хлоролэйба (C), эритролэйба и хлоролэйба (D). Черные кружки — аминокислоты, общие обоим фотопигментам; белые кружки — аминокислоты, разнящие один пигмент от другого.

топигмента неудивительно, поскольку некоторое время назад было показано, что лица с нормальным цветовым зрением иногда разнятся по результатам трихроматического уравнивания (Neitz & Jacob, 1986).

<sup>2</sup> *Трансфекция* — процесс введения нуклеиновой кислоты в клетки человека и животных не-вирусным методом. Трансфекция обычно включает образование в плазматической мембране отверстий через которые внутрь клетки может проникать внеклеточный материал. Трансфицировать можно генетический материал, такой как ДНК или РНК, а также белки, например, антитела. При трансфекции часто используют сильное электрическое поле (электропорация) или электростатически заряженные липиды, способные к образованию *липосом* — структур, которые сливаются с плазматической мембраной, выбрасывая внутрь клетки заключенный в них материал. Известны и другие методы трансфекции. — *Прим. пер.*

### 3.2.3.2 Оппонентные клетки

Обнаружение трех типов колбочек, подтвердившее верность гельмгольцевой трихроматической теории, тем не менее не поставило желанной точки в дебатах вокруг механизмов цветового зрения. Американский психофизиолог Рассел де Валоис (1965) с сотрудниками (De Valois, Abramov, & Jacobs, 1966) измерил отклики латерального коленчатого тела макаки и обнаружил хромоселективные клетки, чьи отклики никак не стыковались с трихроматической теорией, но при этом худо-бедно согласовывались с оппонентной схемой Геринга: оказалось, что некоторые из клеток возбуждаются в ответ на стимуляцию сетчатки длинноволновым (красным) светом (рис. 3.2.26А) и тормозятся в ответ на ее стимуляцию светом средневолновым (зеленым) — т.н.  $R^+G^-$ -клетки. Другие отвечают инвертно: возбуждаются в ответ на зеленый свет и тормозятся в ответ на красный — т.н.  $G^+R^-$ -клетки. Плюс к тому, ученые обнаружили клетки, возбуждающиеся на стимуляцию сетчатки желтым светом и тормозящиеся на ее стимуляцию синим — т.н.  $Y^+B^-$ -клетки и наоборот — клетки, которые возбуждаются в ответ на синий свет, и тормозятся в ответ на желтый — т.н.  $B^+Y^-$ -клетки (см. рис. 3.2.26 В). Прочие же клетки нельзя назвать спектрально оппонентными: некоторые реагируют на стимуляцию сетчатки светом любых длин волн и тормозятся при отсутствии такой стимуляции ( $Wh^+Bl^-$ -клетки), и также, разумеется, наоборот: некоторые возбуждаются вне световой стимуляции сетчатки и подавляются при таковой ( $Bl^+Wh^-$ -клетки), что видно по рис. 3.2.26С. В совокупности описанные клетки коленчатого тела вполне обоснованно можно толковать как нейронную реализацию герингового оппонентного механизма. Спустя время, однако, поступили данные о том, что аналогичную картину отклика дают биполярные и ганглиозные клетки самой сетчатки.

### 3.2.3.3 Оппонентная рекомбинация колбочковых откликов

Итак, напрашивается вопрос: «Каким образом оппонентные отклики получаются из откликов трихроматических, колбочковых?». Несколько неожиданно прозвучит то, что точный ответ до сих пор неизвестен..., но предложен ряд гипотез, неплохо согласующихся с физиологическими данными. Одна из них описывает рекомбинацию колбочковых откликов, показанную на рис. 3.2.27 (De Valois, 1965; De Valois, Abramov, & Jacobs, 1968): оппонентный отклик класса  $R^+G^-$  образуется комбинацией возбуждающего сигнала от L-колбочек и тормозящего от колбочек М. То есть, по сути, берется разность между сигналом двух колбочковых систем ( $L - M$ ).

Сигнал оппонентных  $G^+R^-$ -клеток можно получить комбинацией возбуждающего сигнала от М-колбочек с тормозящим сигналом от L-колбочек ( $M - L$ ).

Сигнал  $B^+Y^-$ -клеток образуется из возбуждающего сигнала S-колбочек и суммы тормозящих сигналов колбочек М и L, что можно записать как:

$$S - (M + L) = S - M - L$$

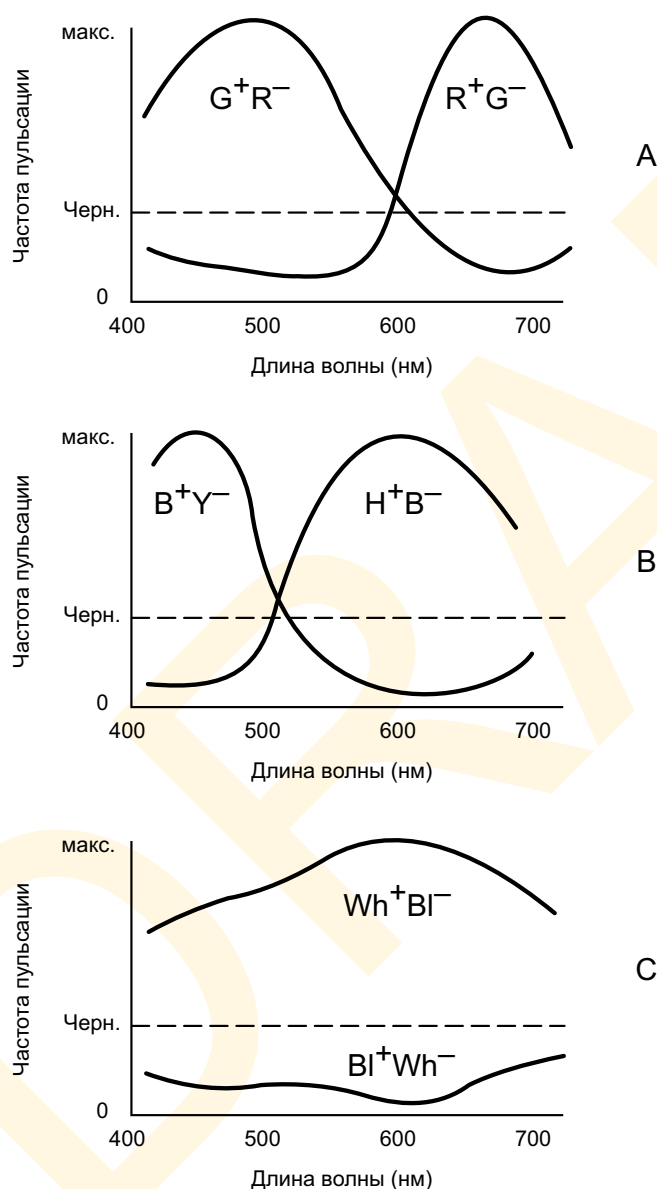
Сигнал клеток  $Y^+B^-$  получается из суммы возбуждающих сигналов от М- и L-колбочек и тормозящего сигнала от колбочек S, то есть:

$$(M + L) - S = M + L - S.$$

Сигнал клеток  $Wh^+Bl^-$  образуется суммированием возбуждающего сигнала от колбочек всех трех типов ( $S + M + L$ ).

Наконец, сигнал  $Bl^+Wh^-$ -клеток получается суммированием тормозящего сигнала опять же от колбочек всех трех типов:

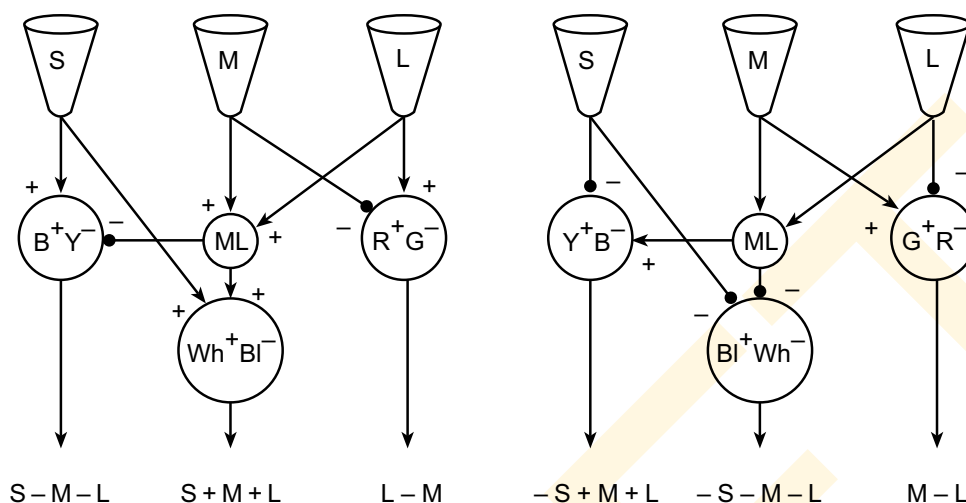
$$-S + (-M) + (-L) = -S - M - L.$$



**Рис. 3.2.26** Кривые нейронного отклика оппонентных клеток: относительные частоты пульсации клеток с рецептивными полями класса  $R^+G^-$ ,  $G^+R^-$ ,  $B^+Y^-$ ,  $Y^+B^-$ ,  $Wh^+Bl^-$  и  $Bl^+Wh^-$  латерального колленчатого тела мозга макаки как функция от длины волны стимула (De Valois, et al., 1966).

Итак, описанная двух-этапная модель цветового зрения — весьма изящный теоретический синтез, позволяющий учитывать множество фактов при некоторых допущениях. Однако же остается ряд неразрешенных вопросов, требующих совершенствования давеча описанного подхода. Большая часть из них касается коротковолновой стимуляции, оппонентный отклик на которую несколько не согласуется с перцепционными данными. К примеру, не совсем понятно, почему стимулы самых коротких длин волн (фиолетовые) воспринимаются содержащими красный компонент, тогда как красно-зеленый канал определен исключительно разницей между откликами М- и L-колбочек (можно сравнить данные рис. 3.2.21 С с частотами пульсации  $R^+G^-$ -нейронов на рис. 3.2.26А).

Ряд исследователей полагает, также что клетки L — М и S — (L + М) кодируют исключительно позицию доминанты стимула (проявляет себя цветовым тоном) и не играют никакой роли в кодировании его яркостной составляющей. Последнее, по их мнению, — миссия т.н. *ахроматических оппонентных клеток* (Lee et al, 1988;



**Рис. 3.2.27** Возможная организация нейронных цепей, реализующих переход от трихроматического этапа к опонетному. Оппонентные отклики образуются из сигналов S-, M- и L-колбочек за счет возбуждающих (стрелки) и тормозящих (точки) связей. Позиция ML означает суммарный эффект M- и L- колбочек.

Shapley, 1990). Колориметрическая же чистота стимула кодируется отношением активностей хроматических оппонентных и ахроматических оппонентных клеток (проявляет себя ощущением насыщенности).

Другие исследователи — приверженцы т.н. *двухфункциональной концепции* — считают, что и доминанта стимула, и его фотометрическая яркость кодируются одними и теми же клетками (Ingling & Martinez-Uriegas, 1983; Kelly, 1983; Lennie, 1984; Derrington et al, 1984; Shapley, 1990): их позиция состоит в том, что хроматически оппонентные нейроны несут двойную нагрузку, кодируя как информацию о длине волны стимула, так и о его фотометрической яркости. При этом, разумеется, двухфункциональный подход исключает то, что яркость стимула кодируется ахроматическими оппонентными клетками.

Короче говоря: сегодня доподлинно известно, что информация о доминирующей длине волны стимула кодируется хроматическими оппонентными нейронами, но по-прежнему не ясным остается механизм кодирования информации яркостной: либо сие реализуется теми же хроматическими оппонентными клетками, либо отдельными ахроматическими.

Наконец отметим, что Де Валоис и Де Валоис недавно предложили четырехэтапную модель цветового зрения, призванную объяснить некоторые неувязки в толковании процессов, идущих в зрительной коре.

### 3.2.3.4 Оппонентные сигналы в зрительных путях

В дополнение к сказанному выше имеется еще одно доказательство существования в зрительной системе оппонентных механизмов — факт существования оппонентных сигналов в зрительных путях. Вначале Де-Валоис установил некоторую связь между нейтральной точкой отдельных нейронов и нейтральными перцептами (DeValois et al., 1966), а спустя 18 лет британский нейрофизиолог Эндрю Деррингтон (с использовани-



ем некоего контрастного стимула, предъявленного наблюдателю на умеренно сером фоне) измерил отклики нейронов латерального колленчатого тела. Ученый установил то, что входной сигнал этим нейронам посылают колбочки двух разных типов и тем самым подтвердил наблюдения Де-Валойса.

Деррингтон разделил парвоклеточные нейроны (см. гл. 1) на две группы:

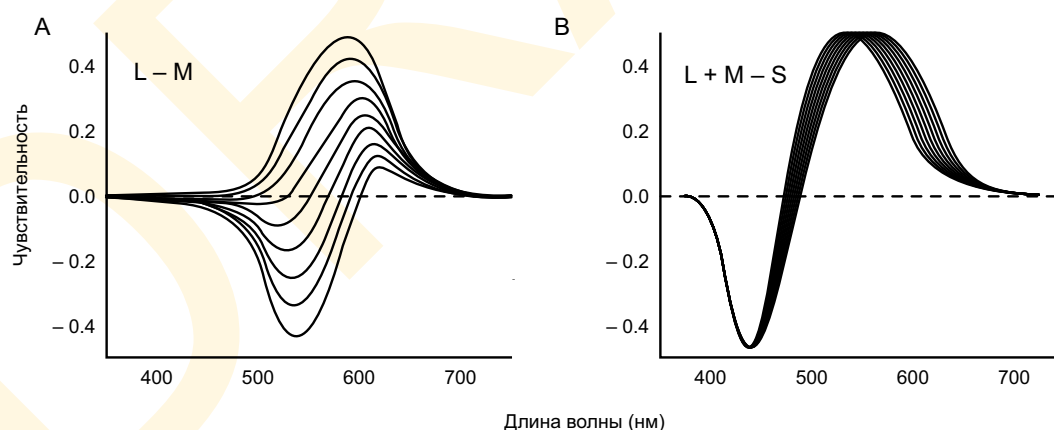
— нейроны *первой группы* получают оппозитный сигнал от L- и M-колбочек (см. рис. 3.2.28 А, где показана спектральная чувствительность таких нейронов — при этом хорошо видно что нейтральная точка варьирует внутри группы). Более того, график второй раз не пересекает нулевой точки в районе 470 нм, что было бы сходно с данными Гурвича — Джеймсон. Следовательно, нет строгих доказательств тому, что данная группа нейронов отвечает за красно-зеленую размерность цветового восприятия;

— нейроны *второй группы* получают сигнал от S-колбочек и оппозитный ему сигнал от комбинации колбочек L и M. У этих нейронов нейтральная точка находится в районе 500 нм (рис. 3.2.28 В).

Магноклетки в реализации оппонентного механизма почти не участвуют. Парвоклетки, а также некоторые из магноклеток управляемы положительной суммой сигналов от L- и M-колбочек и лишь некоторые из них реагируют на разницу между ними.

### 3.2.3.5 Репараметризация

Описанные нами в предыдущем разделе два первых этапа обработки цветовой информации составили простейший пример того как информационная система, начиная работать с одним представлением, преобразует его в другое представление (см. гл. 2): в нашем случае исходным представлением является представление трихроматическое, в котором каждая физическая комбинация длин волн стимулов спроецирована на трехмерное пространство, оси которого определены откликами S-, M- и L-колбочек (рис. 3.2.29А). Затем простые арифметические суммы и разности (рис. 3.2.28) дают оппонентное представление, т.е. преобразуют LMS-пространство в пространство с ося-



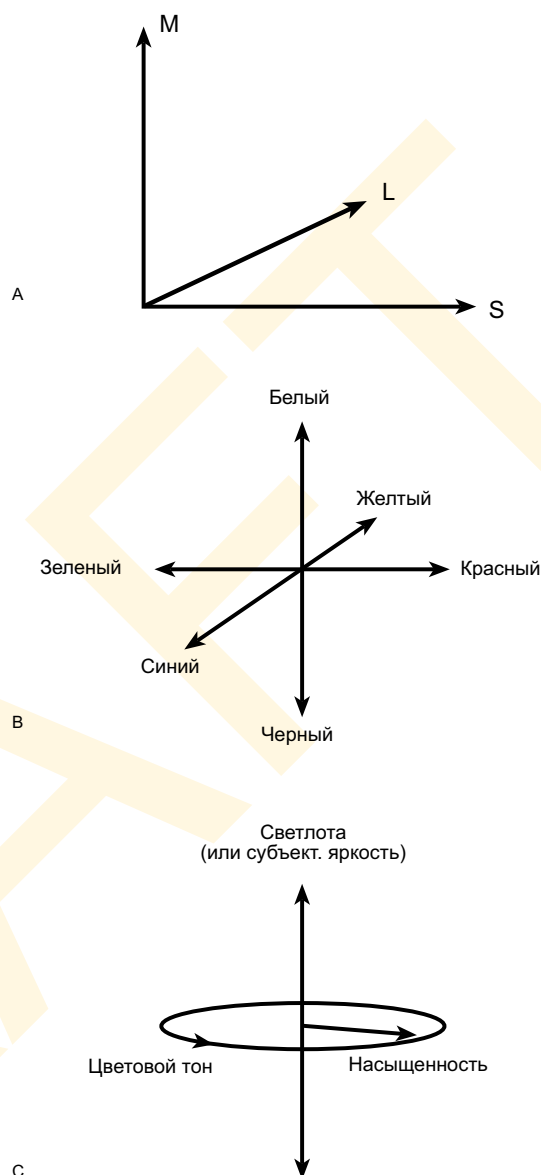
**Рис. 3.2.28** Оппонентные сигналы, измеренные в нейронах латерального колленчатого тела. Данные кривые спектральных откликов выведены из фактически измеренных откликов нейронов латерального колленчатого тела на разнообразные стимулы, предъявляемые на экране монитора. По отклику на хроматичные стимулы большую часть нейронов колленчатого тела в парвоцеллюлярных слоях можно разделить на две группы: первая группа (А) — получает оппонентный сигнал от L- и M-колбочек; вторая группа (В) — получает положительный сигнал от L- и M-колбочек и оппозитный сигнал от колбочек S (Derrington et al., 1984).

ми оппонентных откликов: черный-белый, красный-зеленый и синий-желтый (рис. 3.2.29В). Допустив, что существуют также поздние кортикальные представления, построенные на размерностях цветового тона, насыщенности и светлоты (т.е. как диктует феноменология цветовых ощущений), красно-зеленую и сине-желтую оси следует преобразовать в размерности цветового тона и насыщенности (3.2.29С).

Эффектом описанных преобразований является т.н. *репараметризация* трехмерного пространства колбочковых откликов в иное, эволюционно более успешное эксплицитное представление. Репараметризация — это процесс замены переменных, от которых непосредственно зависит поведение системы.

Несмотря на то, что термин «репараметризация» звучит наукообразно и пафосно, но вполне найдется бытовой пример, способный легко объяснить его. Горячая и холодная вода подается в наши дома по разным трубам и в простейшем исполнении управляется отдельно двумя кранами. Однако во многих современных кранах происходит преобразование этой двумерной системы в другую систему, более удобную в быту, например: смещение крана «вперед-назад» управляет общим объемом воды (холодная + горячая), а смещение «влево-вправо» регулирует температурный баланс (горячая — холодная). Преобразование колбочкового выходного сигнала (S, M и L) в выходной сигнал ганглиозных клеток сетчатки (R/G, B/Y и Bl/Wh) — суть аналогичная репараметризация. Сигнал размерности Bl/Wh представляет общий уровень отклика зрительной системы, тогда как сигнал размерностей R/G и B/Y — его «хроматический баланс»<sup>1</sup>.

Почему зрительная система оперирует



**Рис. 3.2.29** Репараметризация пространств. А — исходное трихроматическое пространство, основанное на размерностях, определенных колбочковыми откликами (LMS). В — пространство построенное на оппонентных осях (черно-белой, красно-зеленой и сине-желтой). С — перцепционное цилиндрическое пространство, построенное на размерностях светлоты (или субъективной яркости, цветового тона (угловая размерность) и насыщенности.

<sup>1</sup> Термин «хроматический баланс» автор берет в кавычки, потому что речь идет о единичном цветовом ощущении, но в отношении *изображений* сцен этот термин употребляется без кавычек. — Прим. пер.

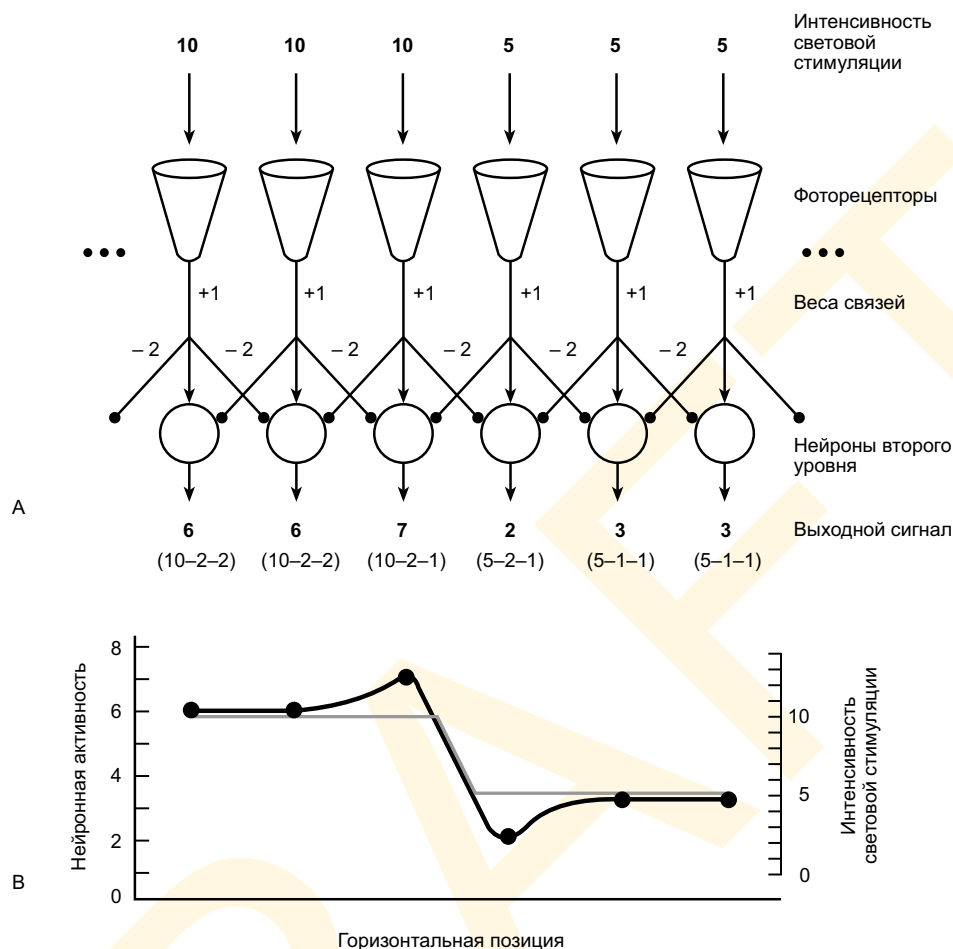
несколькими цветовыми размерностями? Дело в том, что по каждой из размерностей проходит информация, пригодная к решению разного класса задач. Биологический смысл трихроматического представления в том, что оно эффективно извлекает из рецепторного отклика минимально необходимую информацию о спектральном составе света. Но, к сожалению, информация эта не слишком ценна. Гораздо полезнее представление оппонентное, поскольку позволяет определить, какие отличия в уровнях освещенности (скажем, тень, упавшая на поверхность) и какие отличия в спектральном коэффициенте отражения (скажем у разноокрашенных поверхностей) привели к отличиям в стимуляции соседних областей сетчатки. Изменения в уровне освещенности — плод изменений в количестве падающего света (что ведет к изменению выходного сигнала Bl/Wh-системы), но при этом хроматический баланс сцены остается неизменным (то есть неизменен выходной сигнал R/G- и B/Y-систем). Однако изменения в спектральном коэффициенте отражения поверхности ведут к существенному изменению отклика как минимум одной из двух хроматических систем, но также могут дать (а могут и не дать) определенную реакцию Bl/Wh-системы.

А вот чем представление цветовой тон/насыщенность биологически эффективнее, чем представление оппонентное — сегодня не совсем понятно, хотя и очевидно, что первое необходимо для выявления физической причины хроматических изменений, то есть ответа на вопрос: «Что привело к хроматическим изменениям — изменение концентрации пигмента на поверхности или смена самого пигмента?».

### 3.2.3.6 Латеральное торможение

Физиологические механизмы, обсуждавшиеся до сих пор, способны объяснить многие из известных феноменов цветового восприятия, но чтобы истолковать явления симультанного ахроматического контраста и контраста хроматического необходимы дополнительные данные. Выполнить сие можно, начав говорить в том числе о пространственном взаимодействии между соседними областями сетчатки. Возможным механизмом обсуждаемых эффектов может явиться т.н. *латеральное торможение* — одно из *первазивных* (всепроникающих) явлений в физиологии нервной системы вообще и зрительной в частности. Латеральное торможение имеет место в первых слоях сетчатки и в целом объясняет возникновение маховых полос и симультанного хроматического контраста (Ratliff, 1965).

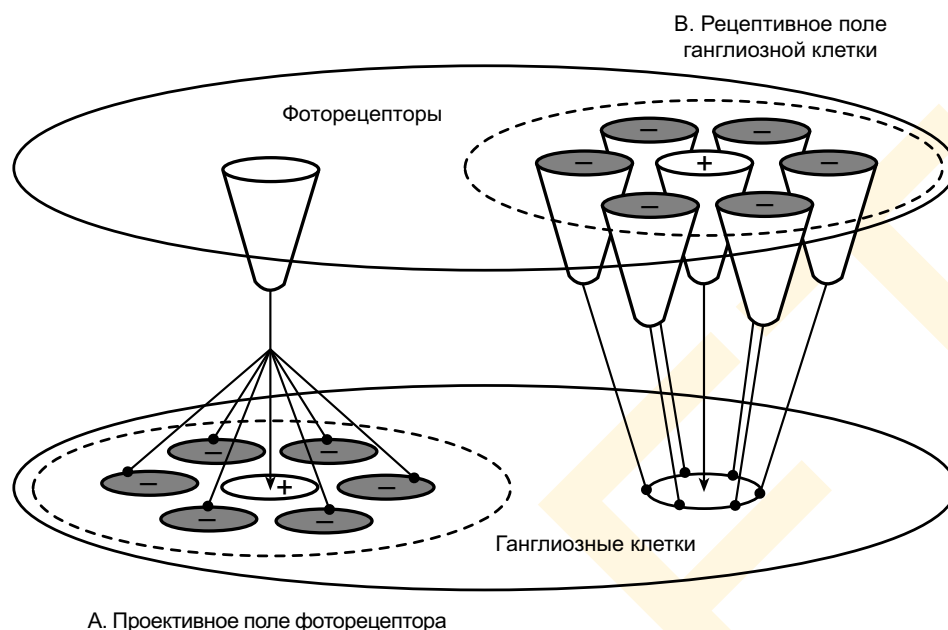
Явление латерального торможения, при котором одни нейроны тормозят соседние — функциональный атрибут нейронных сетей. На рис. 3.2.30 показана простейшая модель латерального торможения в одномерной двухслойной нейронной сети: каждый нейрон первого слоя возбуждает соответствующий нейрон (или группу нейронов) во втором слое, но подавляет при этом нейроны второго слоя, расположенные в сторону (латерально) от эффекторной (возбуждающей) связи. На схеме рис. 3.2.30 эффекторные связи между слоями (острия стрелок) отмечены положительными значениями, а тормозящие связи (черные точки) — отрицательными. Как и в большинстве моделей нейронных сетей выходной сигнал целевого нейрона вычисляется умножением активации каждого из входных нейронов на вес его связи с нейроном целевым и последующим суммированием этих произведений по всем клеткам. Когда мы выполним сие по всей *латерально-ингибиторной сети* (сети с латеральным торможением), показанной в нашем примере, то результирующая картина выходных сигналов во втором слое продемонстрирует нам маховые полосы. Проще говоря, сия нейронная сеть имитирует восприятие края с дополнительным перцептом в виде маховых полос.



**Рис. 3.2.30.** Латеральное торможение в нейронной сети, ведущее к эффекту симультанного контраста. А — каждый фоторецептор возбуждает пространственно согласованный с ним нейрон во втором слое клеток сетчатки (стрелка) и подавляет ближайших соседей по обеим сторонам (точка). В — графики интенсивности стимула (серая линия) и нейронной активности в данной сети (черная линия) как функции от горизонтальной позиции стимуляции. Отметим, что усиление отклика по обеим сторонам яркостной границы аналогично тому, что происходит при образовании маховых полос (см. рис. 3.2.9).

**Проективное поле.** Стоит сказать однако, что одномерная сеть — это далеко нереалистичная модель ретинальных процессов: латеральное торможение, однако, можно толковать в двухмерной логике, просто перевернув одномерную сетку вверх ногами (рис. 3.2.31А). Результатом явится т.н. *осесимметричная сеть*, в которой латеральное торможение, исходящее от данной рецепторной клетки, «расползается» по всем направлениям из эффекторного центра. Мы назовем сей паттерн связей *проективным полем* (Lehky & Sejnowski, 1988).

**Рецептивное поле.** Теперь, ежели мы отследим связи в обратном направлении, то есть от рецепторов к данной клетке, то выделим паттерн максимального отклика и назовем его *рецептивным полем*. Двухмерную сеть латерального торможения мы назовем при этом организованной по принципу *центр-окружение* (рис. 3.2.31В) — это значит, что клетки второго слоя выдадут максимум отклика на стимулирующий паттерн,



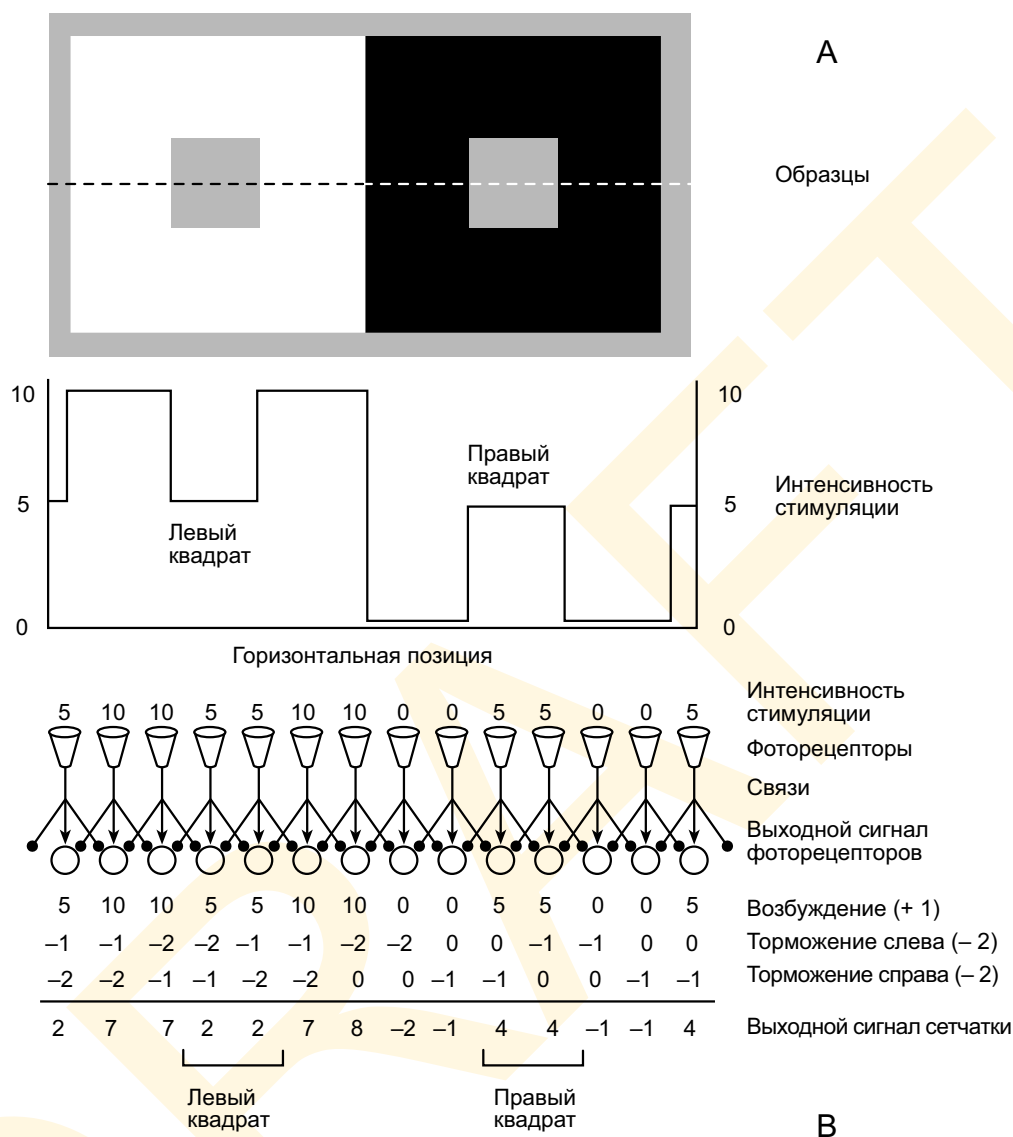
**Рис. 3.2.31** Проективное и рецептивное поля двумерной нейронной сети с латеральным торможением. А — единственный фоторецептор посылает как активирующие, так и тормозящие сигналы множеству ганглиозных клеток по схеме «центр/окружение». В — каждая ганглиозная клетка принимает как возбуждающие, так и тормозящие сигналы от множества фоторецепторов и также по схеме «центр/окружение». На данной иллюстрации показана простейшая нейронная сеть, организованная по принципу «активирующий центр/тормозящее окружение», но, разумеется, в сетчатке существуют проективные и рецептивные поля обратной организации — «тормозящий центр/активирующее окружение».

состоящий из яркой точки (активирующей эффекторный центр) с темным кольцом вокруг. Отметим, что ежели всё рецептивное поле простимулировано равномерным световым пятном, окружение *снижает* отклик второго слоя клеток. То есть, мы говорим, что ганглиозные клетки сетчатки млекопитающих организованы по принципу «центр-окружение».

В 4-й главе мы детально обсудим строение рецептивных полей ганглиозных клеток сетчатки.

**Латеральное торможение и контрастные эффекты.** Многие исследователи полагают, что латерально-ингибиторная архитектура позволяет истолковать как эффекты симультанного ахроматического контраста, так и маховы полосы. На рис. 3.2.32 показана схема возникновения ахроматического контраста за счет латерального торможения: отклик на тестовый образец, находящийся в светлом окружении, жестко заторможен, из-за высокой фотометрической яркости окружения. Поскольку отклик заторможен — образец воспринимается темным. Но такого эффекта мы не наблюдаем, если образец находится в темном окружении, поскольку фотометрическая яркость окружения так низка, что латерального торможения отклика на образец не возникает. Таким образом, мы говорим, что сеть с латеральным торможением прогнозирует эффект ахроматического симультанного контраста, т.е., что один и тот же образец воспринимается темнее в ярком окружении и светлее — в окружении темном. Сие четко





**Рис. 3.2.32** Латеральное торможение и светлотный контраст. А — образец (вверху) и динамика его фотометрической яркости на уровне пунктира (ниже). Отметим, что фотометрическая яркость обоих серых квадратов одинакова, тогда как светлота разная. В — латерально-ингибиторная нейронная сеть, прогнозирующая то, что левый тестовый квадрат будет выглядеть темнее правого.

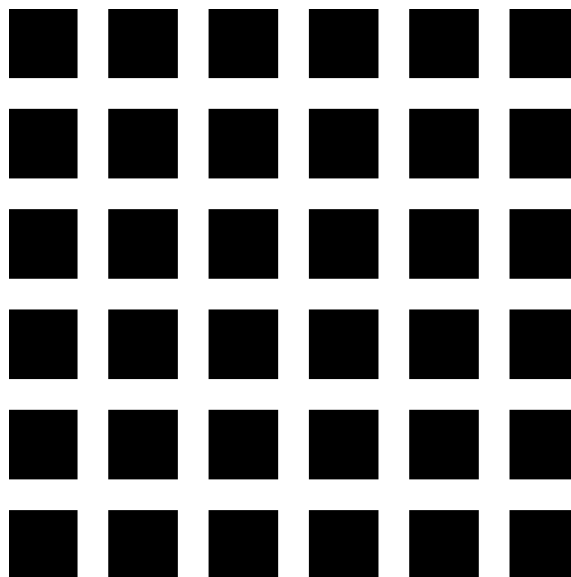
согласуется с феноменом ахроматического контраста, продемонстрированным рис. 3.2.8.

Отметим, что слабым местом изложенной концепции является то, что она не дает четкого представления о том, действительно ли латеральное торможение определяет возникновение симультанного контраста во всем диапазоне размеров стимула: латеральное торможение в сетчатке обычно распространяется на весьма малые углы зрения, тогда как симультанный контраст возникает порой в зонах стимуляции в  $10^\circ$  и более. Возможно в этом случае латеральное торможение имеет место на более высоких уровнях зрительной системы, возможно, также, что контрастный эффект возникает

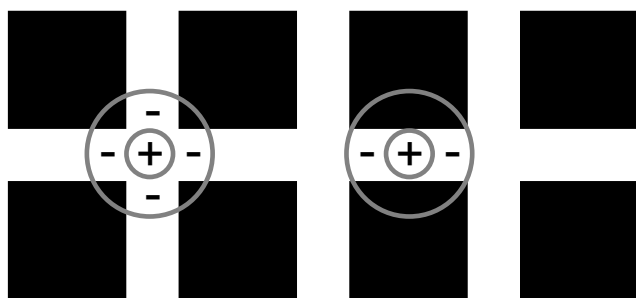
только возле краев, но некие дополнительные механизмы переносят его в центр и на периферию интересующей области.

Существует еще ряд контрастных феноменов, которые также можно объяснить фактом латерального торможения. Одно из таких явлений — *сетка Германа* (см. рис. 3.2.33), у которой в перекрестьях белых линий появляются темные пятна. Мы можем легко убедиться в сугубой иллюзорности этих пятен, если сосредоточимся на одном из перекрестий — темная точка пропадет: на всех перекрестьях пятна останутся видны, но только не на том, на которое мы смотрим.

Появление темных точек вполне объяснимо в логике латерально-ингибиторных сетей, потому что мера торможения на световых перекрестьях очевидно больше, чем в других светлых областях: каждое из перекрестий окружено световой областью с *четырёх* сторон, тогда как прочие только с *двух* (см. рис. 3.2.34). Тот факт, что темная точка исчезает, когда мы сосредотачиваем взгляд строго на перекрестье, можно объяснить тем, что в центральной ячейке эффекты латерального торможения меньше, чем на периферии. «Меньше» может означать, либо, что ингибиторные веса близки к нулю, либо что латеральное распространение почти не выражено, либо и то, и другое. Таким образом, мера распространения латерального торможения играет очень важную роль в объяснении контрастных эффектов, но по-прежнему остается предметом научных споров.



**Рис. 3.2.33** Иллюзорная сетка Германа: мнимые темные пятна появляются в перекрестьях белых полос, но кроме того, на котором зафиксирован взгляд.



**Рис. 3.2.34** Истолкование иллюзии германовской сетки в логике латерально-ингибиторных нейронных сетей. Клетки с рецептивными полями, организованными по схеме «возбуждающий центр/тормозящее окружение» на перекрестьях тормозятся больше, чем такие же клетки на белых полосах.

**Сложные контрастные эффекты.** Однако же существуют контрастные эффекты, объяснить которые латеральным торможением — да еще и на уровне сетчатки — невозможно. На рис. 3.2.35 дан сказанному убедительный пример (Todorovic, 1997): верхний круг выглядит ощутимо темнее нижнего, хотя имеет точно такую же фотометрическую яркость. На первый взгляд перед

нами еще один пример симультанного ахроматического контраста, но давайте обратим внимание на то, что большая часть верхнего круга (который выглядит темнее) закрыта *черными* квадратами, а того, что выглядит светлее — *белыми*. Сие диаметрально противоположно тому, что диктует логика латерального торможения, и наблюдаемый эффект свидетельствует о том, что в данном случае основу ахроматического контраста составляет гораздо более хитрый механизм, основанный, вероятно, на комплексных гештальт-факторах (Venary, 1924). Как видим, контрастный эффект в данном примере обратный, потому что квадраты воспринимаются лежащими на белом (А) и черном (В) фонах и прикрытыми, при этом, сверху черными (А) и белыми (В) квадратами. Контраст в данном случае возникает между кругом и копланарным (соплоскостным) ему фоном: верхний круг выглядит темнее, потому что воспринимается в контрасте с белым фоном, а нижний круг — светлее, потому что контрастирует с черным фоном.

**Динамические контрастные эффекты.** О еще более удивительном примере организационно зависимых контрастных эффектов недавно сообщили Агостини и Проффит (1993), которые предъявляли наблюдателям поле, состоящее из множества черных пятен и одного тестового серого, двигавшихся в определенном направлении, а также поле, состоящее из множества белых пятен и одного серого, двигавшихся в другом направлении. Оба тестовых пятна в покое имели один и тот же уровень серого, т.е. воспринимались равновелико темными. Но как только начиналось движение, так пятно, движущееся в компании белых пятен, начинало восприниматься темнее, того, что двигалось в группе черных. Поскольку все пятна располагались на общем фоне и поскольку эффект возникал только во время движения, стандартное объяснение феномена латеральным торможением, очевидно, несостоятельно. Вероятно, появление эффекта зависит от того, воспринимается ли пятно принадлежащим к группе белых или черным, благодаря их общему движению (в гл. 6 мы вернемся к разговору о перцепционных организации и классификации по общим свойствам).

### 3.2.3.7 Адаптация и последствия

**Хроматическая адаптация** весьма схожа с адаптацией световой: обе понижают чувствительность к свету после протяженной экспозиции константным стимулом за счет временного истощения ресурсов зрительной системы. Биологические процессы, лежащие в основе световой адаптации обратны тем, что действуют при адаптации темновой (см. раздел 1.3.2). Когда мы выходим из полумрака кинотеатра на залитую солнцем улицу, то вначале ощущаем слепяще-яркий свет, поскольку долгий период темновой адаптации перевел большинство молекул светочувствительного пигмента (как в колбочках, так и в палочках) в максимально сенситивное состояние, готовое к отбелке фотонами. Следовательно, в первый момент на ярком свету отклик рецепторов много сильнее обычного, поскольку находятся они в гиперчувствительном состоянии (и большинство из них этот отклик выдают). Продолжение экспозиции мощным солнечным светом вызывает затем световую адаптацию, в течение которой огромное количество молекул фотопигмента отбеливается. Сия массивная отбелка переводит рецепторы в менее чувствительное состояние и радикально снижает их световую реактивность. То есть, световая адаптация позволяет нам нормально воспринимать ярко освещенные сцены.

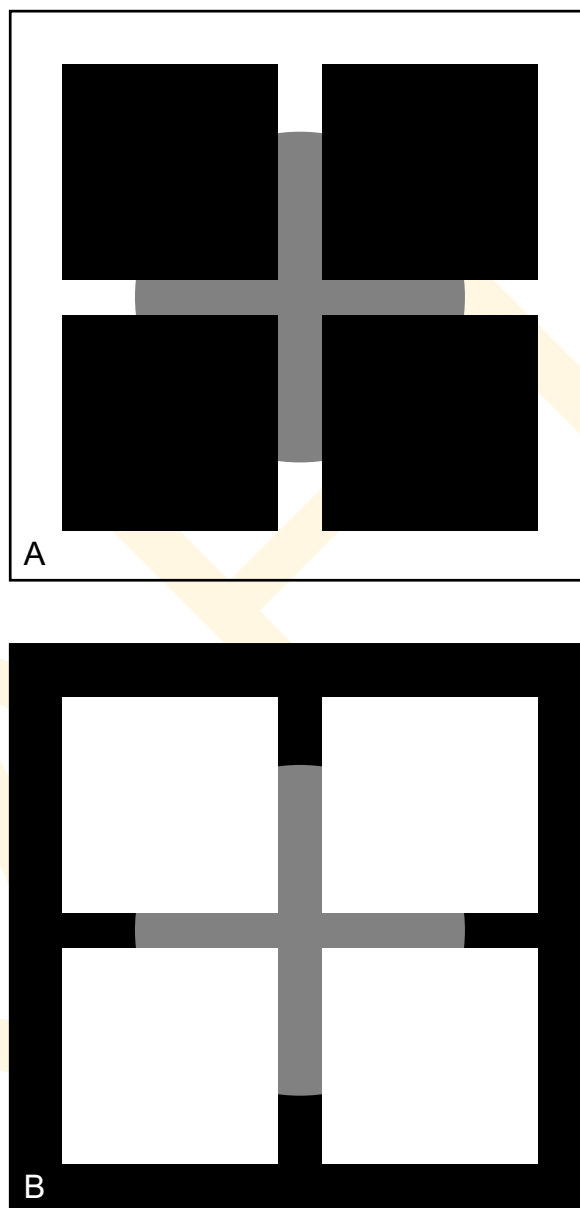
В случае хроматической адаптации колориметрически чистый стимул (воспринимается высоконасыщенным) понижает чувствительность колбочковых систем в разной степени, к примеру, в ситуации с желтым пинг-понговым шариком наибольшим

снижение оказывается в М- и L-колбочках и много меньшим в колбочках S. Так происходит из-за большего отбеливания молекул фотопигмента, в большей степени поглощающих фотоны, несомые адаптирующим стимулом. Дальнейший ход адаптации касается уже вышестоящих клеток зрительной системы (Mollon & Polden, 1979).

Результатом процесса (спустя время) явится то, что активность хроматической системы сдвинется в сторону от обычного восприятия: в варианте ганцфельда будет достигнуто полное равновесие частот пульсации всех нейронов сетчатки (что суть нормальный сигнал по нейтрально-серому стимулу). Именно поэтому желтый пинг-понговый шарик выглядит серым после полной хроматической адаптации.

**Хроматические последействия** — это эффекты, тесно связанные с хроматической адаптацией: длительная неподвижная экспозиция стимулом высокой колориметрической чистоты вызывает появление оппозитного (комплементарного) цветового тона в области сетчатки, отвечающей на нейтральный (скажем, белый) стимул. Например, адаптация по зелено-черно-желтому негативу (рис. 2.1.8) дает постобраз красно-бело-синего флага, если взглянуть на лист белой бумаги.

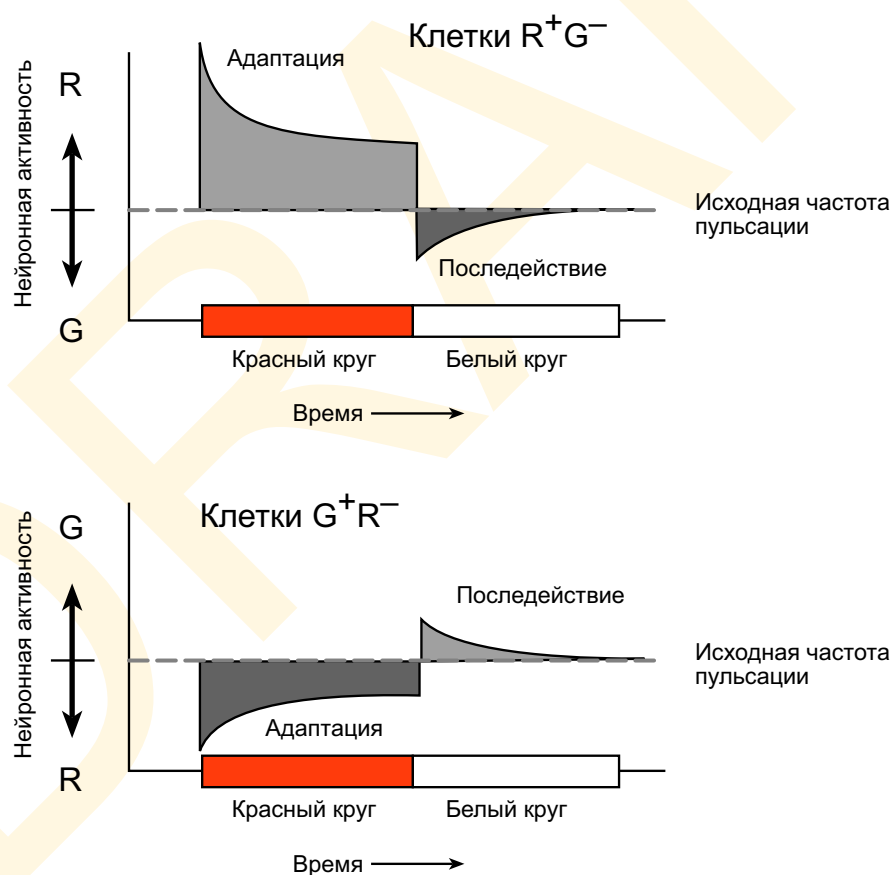
Хроматические постобразы демонстрируют базовое свойство всех последействий равно как и их отношение к адаптации. Протяженная экспозиция интенсивным стимулом рождает адаптацию к специфической форме стимуляции, селективно и временно понижая чувствительность зрительной системы к этому стимулу. В течение этого периода стимул, оппозитный адаптирующему, воспринимается почти нейтральным, то есть на такой стимул сетчатка дает такой же отклик, какой дала бы на нейтральный стимул в неадаптированном состоянии. А вот стимуляция белым (нейтральным), разумеется, приводит к появлению хроматического постобраза.



**Рис. 3.2.35** Иллюзия светлоты. Два круга одинаковой фотометрической яркости прикрыты четырьмя квадратами. Круг на белом фоне (А) выглядит темнее, чем такой же круг на черном фоне (В), хотя по логике симультанного контраста можно ожидать обратного эффекта.

В чем же состоит физиологический механизм хроматических постобrazов? Не удивительно, что постобразы возникают как следствие темпоральных (временных) изменений в отклике хроматически чувствительных нейронов в течение протяженной стимуляции. В качестве иллюстрации этого тезиса рассмотрим ситуацию, в которой сетчатка в течение определенного периода времени стимулируется круглым колориметрически чистым длинноволновым стимулом (в обычном кратковременном восприятии — высоконасыщенным красным): после периода повышенной пульсации у  $R^+G^-$ -оппонентных клеток данной области сетчатки временно снижается реактивность; вдобавок к этому соответствующие  $G^+R^-$ -клетки после длительного периода торможения становятся несколько *более* реактивными. Сказанное проиллюстрирует рис. 3.2.36.

В течение периода измененной чувствительности в определенных условиях будут возникать хроматические постобразы: если после длительной экспозиции красный адаптирующий стимул заменить нейтральным белым, сильно утомленные  $R^+G^-$ -клетки в ответ на белый свет будут пульсировать не столь активно как обычно, а клетки  $G^+R^-$  станут чуть более активными, поскольку освободились от торможения. Результат окажется таковым, что баланс между  $R^+G^-$  и  $G^+R^-$  сместится от нейтраль-



**Рис. 3.2.36** Нейронная теория хроматической адаптации и последействий. Оппонентные клетки снижают активность и тормозятся (адаптируются) после продолжительной экспозиции красным круглым стимулом. Когда красный стимул сменяется белым, отрицательное последействие в нейронной активности продуцирует возникновение хроматического постобраза с комплементарным цветовым тоном, угасающего по мере восстановления равновесия.



ного в тех клетках, что кодируют хроматическое воздействие стимула в адаптированной области сетчатки. В частности, активность системы  $G^+R^-$  окажется выше, чем у  $R^+G^-$ . Вне адаптации отклик обеих систем на нейтральный белый свет сбалансирован, но картина клеточной активности, возникающая после адаптации по интенсивно-красному стимулу, такова, какая вне адаптации возникает в ответ на интенсивно-зеленый стимул.

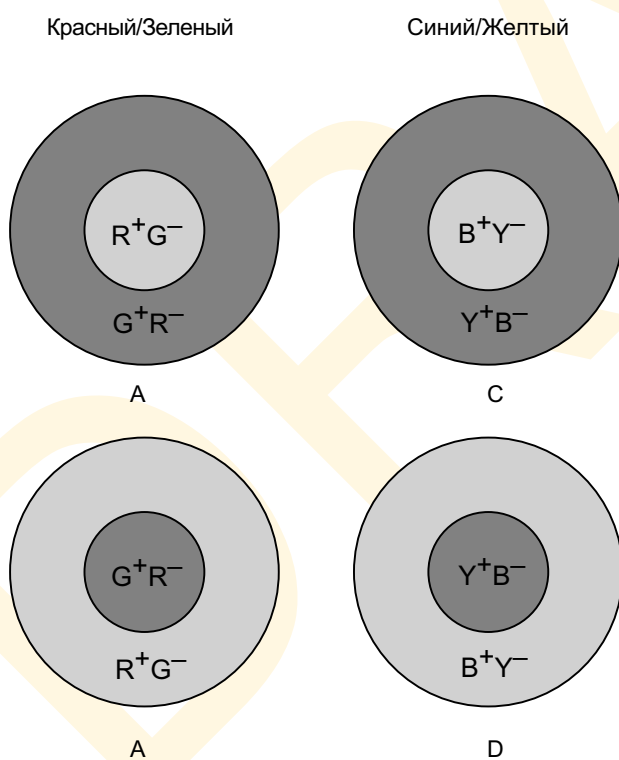
### 3.2.3.8 Дубль-оппонентные клетки

О хроматическом контрасте в целом можно говорить, как о результате работы т.н. *дубль-оппонентных клеток*, обнаруженных в зрительной коре (Michael, 1978). У дубль-оппонентных клеток пространственно оппонентная организация, аналогична той, что давеча обсуждалась нами в контексте ахроматического контраста, но также обладает и хроматически оппонентной структурой (рис. 3.2.37). Иными словами, центры и окружения в рецептивных полях этих клеток хроматически оппоненты: красно-зеленые оппонентные клетки возбуждаются в ответ на стимуляцию сетчатки длинноволновым (красным) светом в центре их рецептивных полей и тормозятся, когда центры стимулируются светом средневолновым (зеленым) —  $R^+G^-$ ; стимуляция окружений зеленым светом ведет, разумеется, к возбуждению, а красным — к торможению (и наоборот) —  $G^+R^-$ .

Сине-желтые дубль-оппонентные клетки будут возбуждаться при стимуляции центров своих рецептивных полей коротковолновым светом (синим) и тормозиться при его стимуляции средне-длинноволновым (желтым) —  $B^+Y^-$ . И наоборот — возбуждаться при стимуляции окружений желтым и тормозиться при их стимуляции синим —  $Y^+B^-$ .

Ежели мы присовокупим к разговору ахроматические клетки, в той или иной мере возбуждающиеся при стимуляции центров своих рецептивных полей любым светом и тормозящиеся вне такой стимуляции —  $Wh^+Bl^-$ , а также возбуждающиеся вне стимуляции окружений и тормозящиеся при их стимуляции любым светом —  $Bl^+Wh^-$ , то мы получим набор пространственно-оппонентных клеток по всем размерностям герингового пространства. Сие физиологический базис эффектов симультанного контраста.

В качестве иллюстрации того, как дубль-оппонентные клетки реализуют физиологию хроматического контраста, рассмотрим пример, в кото-



**Рис. 3.2.37** Рецептивные поля дубль-оппонентных клеток. Некоторые клетки в V1-области коры демонстрируют как пространственную, так и хроматическую оппонентность в своих рецептивных полях.

ром сегмент серого кольца, окруженный высоконасыщенным зеленым, может восприниматься розоватым (см. рис. 3.2.7). Вначале рассмотрим выходной сигнал клеток с архитектурой рецептивного поля «красный центр — зеленое окружение», стимулируемого по центру как раз серым кольцом, а по периферии интенсивно-зеленым фоном этого кольца: в данном случае на стимуляцию центра рецептивного поля серым нейроны ответят энергичнее, чем на стимуляцию серым всего рецептивного поля, поскольку на окружение будет действовать высокохроматичный зеленый. Сия сверхактивность клеток с  $R^+G^-$ -центрами и  $G^+R^-$ -окружениями в рецептивных полях (рис. 3.2.37А) ведет к тому, что центральная область стимула начинает восприниматься розовой.

Зеленоватость сегмента кольца на интенсивно красном фоне объясняется аналогично: за счет стимуляции окружения интенсивно-красным потоком активируются клетки с  $G^+R^-$ -центром в своих рецептивных полях и  $R^+G^-$ -окружениями (рис. 3.2.37В). В той же логике толкуются эффекты сине-желтого контраста. Прочие варианты хроматических контрастов можно объяснить, комбинируя приведенные выше рассуждения.

### 3.2.3.9 Высшие корковые механизмы

Выше мы намекнули на то, что где-то в глубинах мозга реализуется третий нейронный этап цветового восприятия, объединяющий выходной сигнал красно-зеленых и сине-желтых оппонентных систем в размерности, соответствующие цветовому тону и насыщенности (субъективная яркость или светлота уже представлены ахроматической системой). Вопрос сей сугубо умозрительный, поскольку никто еще не нашел корковых клеток, представляющих цвет по такой схеме. Однако тот факт, что цветовой тон, насыщенность и светлота доминируют в цветовом ощущении, убеждает нас в том, что соответствующие нейронные структуры рано или поздно будут найдены.

**Ахроматопсия.** Современное представление о высших корковых процессах получено в основном из нейрофизиологических наблюдений. Поражение определенных областей коры у людей приводит к различным видам недостаточности в обработке цветовой информации, и все эти виды отличны от обычной цветовой слепоты. Один из вариантов недостаточности цветового восприятия именуется *ахроматопсией*, проявляющей себя жалобами больного на то, что мир вокруг обесцвел или, как минимум, что его краски сильно потускнели (Meadows, 1974). Процесс затрагивает либо все поле зрения, либо часть его. Коррелированное нейронное поражение локализуется в области т.н. *престриарной коры*, т.е. кпереди от первичной зрительной коры, что возможно согласуется с областями цветовой обработки (Zeki, 1980).<sup>1</sup>

Ахроматопсия весьма отлична от цветовой слепоты, вызванной отсутствием одного или двух колбочковых пигментов. В одном весьма показательном случае пациент, страдавший этим недугом, успешно прошел т.н. *пороговый тест*, что свидетельствовало о нормальной работе колбочковой системы, но при этом совершенно не мог уравнивать и сортировать хроматичные стимулы, лежащие существенно выше порога (Mollon, Newcombe, Polden, & Ratcliff, 1980). Такие пациенты обычно могут охарактеризовать цвета объектов по памяти (к примеру, «небо синее», «банан желтый»), но не в силах изобразить что-то окрашенное, поскольку не могут выбрать нужную краску или карандаш. Проблема, видимо, состояла в том, что корковое поражение коснулось той

<sup>1</sup> Аналогичные области были найдены в престриарной коре обезьян.

части мозга, что работает с цветовой информацией исходно зарегистрированной колбочками сетчатки, в то время как когнитивный компонент цветowych характеристик предметов остался нетронутым.

**Ментальная цветовая недостаточность.** Второй формой высшей перцепционной дисфункции, вызванной поражением коры, является т.н. *ментальная цветовая недостаточность* (Kinsbourne & Warrington, 1964). Пациенты с этой патологией могут идентифицировать предметы по их окраске, но не знают, какими должны быть цвета объектов. Эти люди порой не могут нарисовать фигуру нужного цвета, и, вероятнее всего, затруднение связано просто с непониманием того, какой окраски должен быть тот или иной объект. Если верный цвет *назвать*, то больной легко выберет нужную краску. Описанная проблема часто коррелирует с расстройствами речи, что указывает на топографическую близость речевого и цветового центров в коре.

**Цветовая аномия** — третий вид коркового поражения, связанный с нарушениями цветового восприятия и проявляющей себя неспособностью к речевому именованию цветовых ощущений (Geschwind & Fusillo, 1966). Цветовые аномики не только не могут верно назвать тот или иной цвет, но они обычно неспособны и к обратному действию, то есть к корректной интерпретации названия цвета. Однако они вполне способны нарисовать предмет и верно раскрасить его, опираясь на ранее полученное знание о его цвете. При этом не в силах следовать словесным инструкциям по использованию тех или иных красок. Цветовые аномики испытывают полноценные цветовые ощущения, на что указывает их способность различать предметы по окраске, а также выполнять трихроматическое уравнивание. Память о цветовых характеристиках предметов не нарушена, но патология состоит в неспособности поименовать переживаемое цветовое ощущение. Феномены именования цветов нормальными наблюдателями мы рассмотрим в конце текущей главы, когда коснемся вопроса цветовой категоризации.

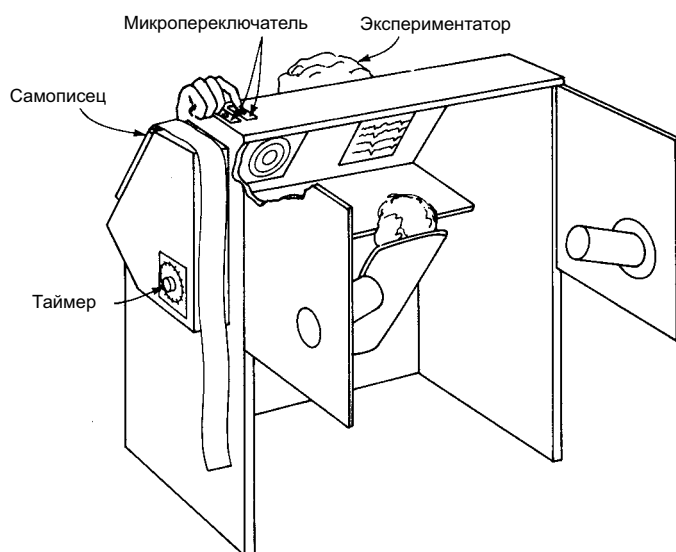
\* \* \*

Итак, мы видим, что описанные выше варианты корковой патологии весьма отличны от классических цветовых аномалий и не только потому, что первичная цветовая обработка остается незатронутой, но также и потому, что это виды *приобретенной* цветовой недостаточности, возникающие обычно после травм головного мозга. Классическая же цветовая слепота — всегда врожденная, и передается по наследству (Nathans, 1987, 1989).

### 3.2.4 Онтогенез цветового зрения

Коль скоро цветовое зрение генетически предопределено, то вполне естественны следующие вопросы: «Как рано младенцы начинают воспринимать мир цветным? Каков мир глазами новорожденного? Сразу ли он полон красок или по началу все-таки тускло-серый, а цветным становится только спустя дни, недели или даже месяцы после рождения?».

Хотя у нас нет окончательного ответа, мы точно знаем, что уже двухмесячный малыш располагает нормальным цветовым зрением. Откуда это известно, ведь такой маленький ребенок не может рассказать о своих ощущениях? Действительно, он неспособен выразить мысль словами, но в специально поставленном эксперименте может сообщить нам о своих ощущениях через *поведение*. В течение двух последних десятилетий были разработаны весьма изощренные методы, позволяющие узнать о том, как видят и слышат младенцы самых ранних возрастов. Один из таких методов, назван-



**Рис. 3.2.38** Экспериментальная установка слежения за зрительным поведением младенцев. Направление взгляда малыша отслеживает скрытый от него экспериментатор (Fantz, 1961).

ный *парадигмой зрительных предпочтений* (Fantz, 1958, 1965) оказался весьма эффективным и показательным.

Итак, главный вопрос в том, может ли младенец отличить стимулы разных цветностей, но одной фотометрической яркости? Малышам предъявляли два стимула разной цветности — один слева в поле зрения, другой справа (рис. 3.2.38). Преимущественное направление взгляда ребенка в течение тестового периода отслеживал скрытый от малыша экспериментатор, который не знал, какой стимул, на какой стороне расположен. В половине случаев стимулы меняли местами, дабы убедиться, что малыш именно следит за понравившимся ему образцом, а не просто уставился в одну точку. Есть основания полагать, что у ре-

бенка нет мотива следить за образцом, кроме как благодаря тому, что он способен видеть хроматические отличия между стимулами.

Было выявлено, что у двухмесячного ребенка работают все три колбочковых системы, более того, выяснилось даже, что и одномесячный ребенок обладает цветовым зрением, но пока неясно является ли он при этом полноценным трихроматом, поскольку имеются сведения, что в этом возрасте дети не могут отличить желто-зеленые и пурпурные стимулы от белого (Teller, Peeples, & Sekel, 1978). То есть, у новорожденных и младенцев первого месяца жизни, вероятнее всего, имеется некоторая нехватка цианолэйба. В любом случае ясно, что дети воспринимают мир цветным даже в самом раннем возрасте и как минимум качественно также, как и взрослые.

### 3.3 ФАСЕТНЫЙ ЭТАП ЦВЕТОВОГО ВОСПРИЯТИЯ

Весь предшествующий разговор (т.е. разговор об оптическом этапе цветового восприятия) был лишь прелюдией к разговору основному. Действительно, если описанные ранее принципы напрямую приложить к оптическому ретинальному изображению сложной сцены, прогноз результатов ее восприятия окажется радикально отличным от того, как человек видит ее на самом деле. В качестве примера рассмотрим восприятие серых элементов рис. 3.3.1. Любой, взглянув на снимок, скажет, что стена собора просто белая. Но ежели же мы озаботимся подсчетом числа градаций серого, воспринимаемых при внимательном разглядывании этой фотографии, т.е. собственно *ощущений*, возникших под действием света, отраженного от каждого ее элемента, то



обнаружим, что стена хоть и воспринимается белокрашенной, но с множеством серых теней.

Сей пример наглядно демонстрирует то, что восприятие т.н. *предметных стимулов* (т.е. стимулов, полученных отражением света от поверхностей объектов) много сложнее восприятия примитивных 2D-образцов, обсуждавшегося в разделе 3.2: наклон, текстура и освещение поверхностей в реальных сценах синергично влияют на результат нашего восприятия предметных стимулов.

Вспомним, что говоря о «цвете» поверхности, мы имеем в виду психологическое свойство, сопоставленное этой поверхности. Физически поверхности «бесцветны» до тех пор, пока свет, отразившись от них, не попадет в глаз наблюдателя. Физический параметр, в первую очередь влияющий на результат цветового восприятия поверхности — это *спектральная отражательная способность* этой поверхности (*спектральный коэффициент отражения, спектр отражения*): доля отраженного света (в процентах от упавшего) по каждой из длин волн. Спектральную отражательную способность можно отобразить графически по той же схеме, что и спектр излучения, но ось ординат при этом представит процентную долю фотонов «отскочивших» от поверхности (а не испущенных источником света). Непрозрачные поверхности отражают не весь упавший свет, но часть поглощают, равно как поверхности прозрачные поглощают часть идущего через них света, а часть попускают. В отношении прозрачных поверхностей мы говорим о *спектральной пропускательной способности* (*спектре пропускания*).

Важно понимать, что:

⇒ Спектр отражения поверхности — суть ее инвариантное (неизменное) свойство в том смысле, что ни по одной из длин волн она не меняется при смене освещения и тем более условий просмотра — поверхность всегда отражает свет в одних и тех же пропорциях по каждой из длин волн.

Именно поэтому спектры отражения поверхностей — это биологически весьма полезные свойства последних, позволяющие особи различать те объекты, в отношении которых идентификация по поверхности является для этой особи диагностически значимой, позволяющей отличить, к примеру, спелый плод от неспелого, дятла от синицы, шатенку от брюнетки и т.д.

Восприятие зрительной системой именно отражательной способности поверхно-



**Рис. 3.3.1** Пример константности светлоты. Стены собора остаются в сознании наблюдателя белыми, несмотря на хитроумную игру теней разной глубины, образованных архитектурным рельефом. Сей факт невозможно истолковать в простой логике оптического этапа цветового восприятия — потребуются рассуждения в логике фасетной обработки, т.е. учет ориентации поверхностей и их освещенности (Rock, 1984).



стей приоритетно, поскольку свет, воздействующий на сетчатку (и характеризующийся *спектральным распределением энергии излучения — спектром излучения*), определяется двумя факторами: спектром отражения поверхности и спектральным распределением энергии падающего на эту поверхность света. В этом можно легко убедиться в этом, приглядевшись к тому, что происходит, когда та или иная поверхность (спектр отражения которой, как мы уже сказали выше, всегда постоянен) освещается сильно отличными друг от друга источниками (см. рис. 3.3.2).

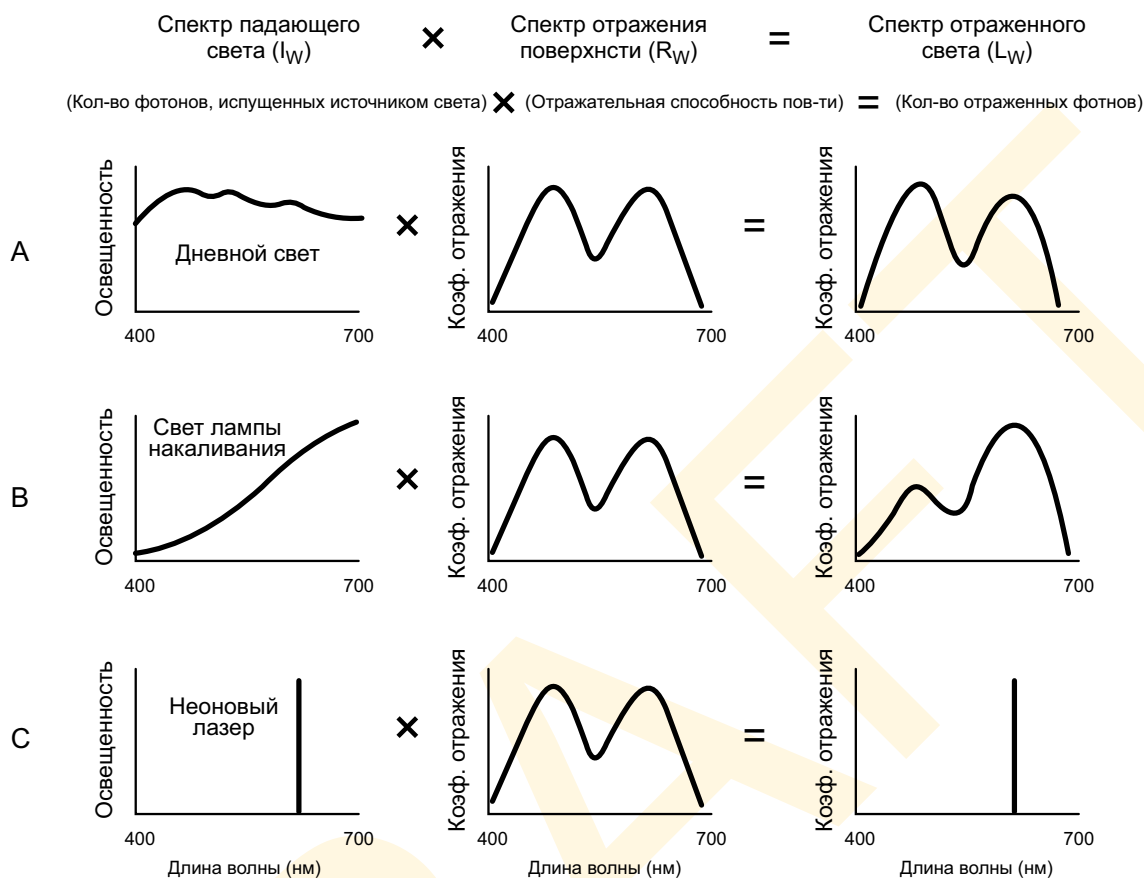
Чтобы вычислить спектральное распределение энергии света, воздействующего на сетчатку в данном ее участке ( $L_w$ ) число фотонов в падающем свете по каждой из длин волн ( $I_w$ ) следует умножить на спектр отражения поверхности, т.е. на процентную долю света, отраженного этой поверхностью по каждой из длин волн ( $R_w$ ):  $L_w = I_w \times R_w$ , где  $w$  представляет длину волны в диапазоне от 400 до 700 нм.

Таким образом, ежели свет, падающий на поверхность — это «белый» солнечный свет (т.е. свет с более или менее гладким спектром), то спектральное распределение энергии света отраженного будет почти копией спектрального коэффициента отражения поверхности. Цветовое ощущение от поверхности, возникающее при таком освещении условно можно назвать «истинным цветом» этой поверхности, т.е. четко согласованным с ее спектром отражения. При каком-нибудь ином, отличном освещении, скажем при свете ламп накаливания, спектральное распределение света, отраженного от той же самой поверхности, количественно окажется иным, хотя качественно останется примерно тем же (см. рис. 3.3.2В).

Теперь рассмотрим то, что происходит, когда та же самая поверхность освещается качественно иным светом, скажем лазерным монохроматом: умножение спектра отражения поверхности на спектральное распределение энергии монохроматического излучения дает совершенно другой спектр отраженного света (рис. 3.3.2С). Проекционное изображение на сетчатке в этом случае несет в себе энергию одной лишь длины волны, поскольку свет, упавший на поверхность, состоит из фотонов одной частоты, а отразившая их обычная поверхность (т.е. нелюминесцирующая), не в состоянии добавлять световую энергию, но только убавлять ее, поглощая часть фотонов. Приведенный пример исключителен и не встречается в повседневности (потому что монохроматические источники, разумеется, не используются для освещения помещений и улиц), но общий принцип рассуждений применим ко всем ситуациям.

Итак, когда мы глядим на какую-то поверхность, то спектр света, воздействующего на наш орган зрения, зависит не только от спектрального коэффициента отражения поверхности, но и от спектрального состава падающего на нее света. Теперь становится понятным, почему, скажем, в полиграфии и фотографии, а также в экспериментах по исследованию цветового зрения, так тщательно контролируется спектральный состав освещения.

Однако же путь, по которому зрительная система идентифицирует спектральную отражательную способность поверхности — это отнюдь не спектрометрия, построенная на физическом измерении спектров падающего света и света отраженного — инвариантность спектра отражения поверхности воспринимается независимо от состава освещения. Сие еще один пример необходимости в *инверсии*, обсуждавшейся в 1-й главе: перед зрительной системой стоит задача «откатить» результат взаимодействия между поверхностью и светом, восстановив тем самым свойства собственно поверхности. И вновь мы сталкиваемся с тем, что задача оказывается недоопределенной, поскольку глаз получает информацию только о проксимальной переменной (спектральный состав стимула), задаваемой совместно двумя дистальными переменными: спектром от-



**Рис. 3.3.2** Задача обеспечения цветовой константности: одна и та же поверхность, с одним и тем же спектром отражения (центральная колонка), в разных условиях освещения (левая колонка) может выдавать свет совершенно разных спектральных составов (правая колонка); цветовая константность необходима зрительной системе для восстановления инвариантных отражательных свойств поверхностей, невзирая на естественную переменчивость освещения.

ражения поверхности и спектральным составом падающего на эту поверхность света. Очевидно, что логически обе этих переменных восстановить невозможно, но мы знаем при этом, что зрительная система человека успешно справляется с этой задачей в широком диапазоне условий. Вновь мы приходим к мысли, что зрение человека пользуется неким дополнительным источником информации, вероятнее всего каким-то хитрым эвристическим.

### 3.3.1 Цветовая константность

Итак:

⇒ Способность воспринимать спектры отражения поверхностей вне зависимости от состава освещения и прочих условий просмотра, именуется *цветовой константностью*.

Разговор о цветовой константности мы поделим на две части: вначале обсудим простейший вариант — *константность светлоты*, а затем *хроматическую констант-*

ность. Первая и вторая суть слагаемые общей цветовой константности. Цветовая константность, в свою очередь — одно из слагаемых *константности зрительного восприятия* — способности воспринимать физические свойства объектов внешнего мира, неизменные в самых разнообразных условиях просмотра (в отличие от свойств оптических изображений на сетчатке, которые весьма переменчивы). Примеры константности восприятия мы обсудим в 7 главе, когда поведем разговор о пространственных свойствах объектов, таких как размер, ориентация, позиция и форма.

Сразу отметим, что цветовая константность не абсолютна: по мере смены условий освещения возникают легкие, но биологически очень важные изменения в цветовых перцептах поверхностей. К примеру, при зрительном восприятии естественных ландшафтов изменения эти позволяют нам интуитивно определять время дня (Zaidi, 1997).

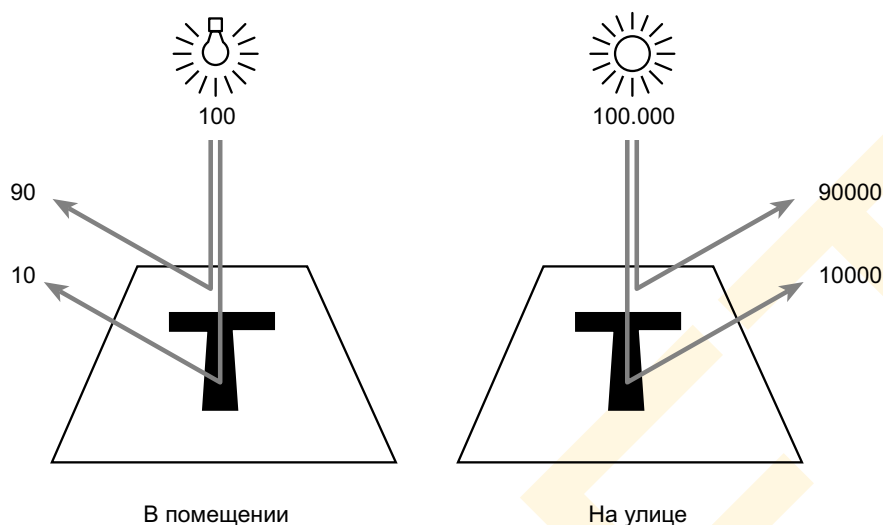
### 3.3.1.1 Константность светлоты

⇒ *Константность светлоты* — это факт неизменности зрительной оценки отражательной способности поверхности при смене освещенности и/или прочих условий просмотра.

Продемонстрировать константность светлоты можно, взглянув на страницу этой книги в помещении, при искусственном свете, а затем выйдя с нею же в руках на солнце. В обоих случаях бумага останется в нашем сознании белой, а текст — черным. И хотя в целом сие неудивительно, все-таки рассмотрим то, какие количества света отразятся от белой бумаги и черной краски в первом и втором случаях (см. рис. 3.3.3). Давайте договоримся, что в среднем при комнатном освещении на лист бумаги в единицу времени падает 100 неких условных единиц света и что бумага отражает 90% этого света в глаз наблюдателя, тогда как черная краска только 10%. То есть, 90 единиц света поступает от пробельных элементов листа — от бумаги, а 10 единиц от элементов печатных — от краски. В обычный солнечный день света как минимум в 100 раз больше (скажем 10000 абстрактных единиц) и, следовательно, 9000 единиц отразятся от бумаги и 1000 единиц от краски. Отметим, что количество света, поступившее от черных элементов листа при высоком уровне освещенности, почти, вдесятеро больше, чем количество света, поступающее в глаз от пробельных элементов страницы при тусклом комнатном освещении. Но даже при тусклом освещении мы все равно воспринимаем лист бумаги много более светлым, чем краску при освещении солнечном. В чем причина?

**Теории зрительной адаптации.** Одним из толкований механизма константности светлоты (а также и хроматической константности) является то, что зрительная система *адаптируется* к изменениям в освещенности сцены, понижая свою чувствительность на ярком свете и повышая ее при недостатке освещения (Hurvich, 1981, von Kries, 1905). Главная идея состоит в том, что ежели адаптация к различным видам освещения влечет за собой изменение уровня нейронной активности в некоем нейтральном диапазоне, светлота остается константной автоматически: уровень освещенности попросту факторизуется адаптацией, полагая общий (т.е. неспектральный) коэффициент отражения главным детерминантом нейронной активности. Поскольку адаптация — во многом миссия сетчатки, то имеются все основания предположить, что механизм достижения постоянства светлоты поверхностей — это механизм начального уровня.

Однако, несмотря на то, что изменение освещенности сцены запускает механизмы адаптации, последние не могут служить объяснением всех проявлений константности

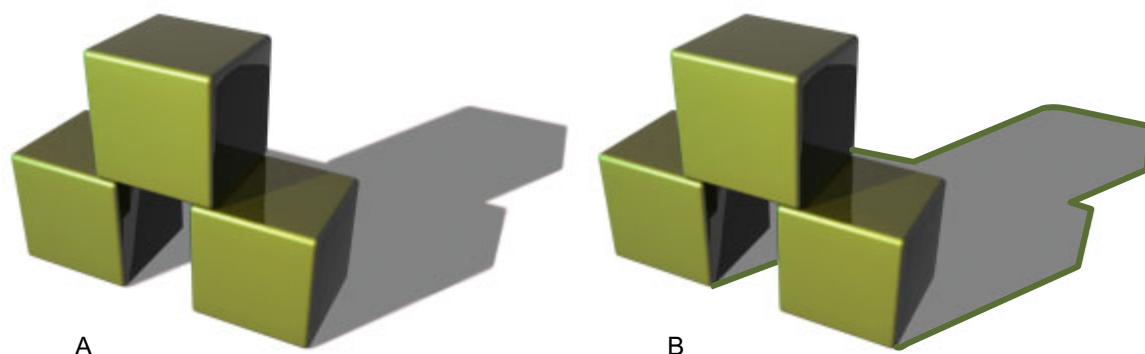


**Рис. 3.3.3** Константность светлоты поверхности при разных уровнях ее освещенности. Белая бумага выглядит белой и черная краска — черной вне зависимости от того, где рассматривается: в помещении при низкой освещенности или на открытом воздухе при ярком солнце.

светлоты: дело в том, что процесс адаптации требует определенного времени (не меньше минуты), тогда как осознание неизменности светлоты поверхности наступает моментально. Если погасить настольную лампу, ограничившись слабым светом, скажем, торшера в дальнем конце комнаты, страница с данным текстом все равно останется белой даже в полумраке и не станет серой даже на период наступления темновой адаптации, длящейся несколько минут (см. раздел 1.3.2, рис. 1.3.24): мы моментально ощущим, что свет резко ослаб и теперь его падает существенно меньше и на саму страницу, и на предметы ее окружающие. Если бы адаптация была моментальной, то мы бы не заметили того, что произошло. Сие потому, что адаптация решает проблему константности путем *устранения* информации об изменениях в освещении. Судя по всему, мы воспринимаем отражательную способность сцены отдельно от освещенности ее поверхностей, что иллюстрирует наш пример с настольной лампой.

Еще одним слабым местом толкования постоянства светлоты поверхностей через адаптацию является то, что некоторые виды зрительной константности проявляют себя в рамках одной сцены при постоянном освещении. Например, когда объект отбрасывает тень на поверхность (см. рис. 3.3.4А), затененный участок не воспринимается нами как темный (как если бы он был специально покрашен), но начинает выглядеть так, что на него падает меньше света именно из-за образования тени. Такое проявление константности светлоты невозможно объяснить адаптацией, поскольку классические световая и темновая адаптации — это глобальные феномены, касающиеся всего поля зрения. Любопытно, что ежели тень от предмета обвести черным контуром (см. рис. 3.3.4В), то затененная поверхность начнет восприниматься темнее (Hering, 1905). При этом нам неизвестен адаптационный механизм, объясняющий радикальные отличия между восприятием обыкновенной тени и тени, обведенной контуром.

Итак, хотя адаптационные механизмы действительно могут вызвать долгосрочные изменения в зрительном отклике на изменение уровня освещенности сцены, они не могут дать общего объяснения феномену константности светлоты. Следовательно, мы должны искать ответ в каком-то ином направлении.



**Рис. 3.3.4** Восприятие теней. А — обычная тень воспринимается как отличие в уровнях освещенности поверхности, но не как окрашенный участок. В — искусственная обводка радикально меняет восприятие: поверхность выглядит темнее и специально окрашенной.

**Бессознательные инференции и реляционная концепция.** Две классических теории константности светлоты предложили в свое время те самые ученые, что яростно спорили о механизмах цветового зрения: Герман фон Гельмгольц и Эвальд Геринг. Гельмгольц пытался объяснить постоянство светлоты, применяя свой общий принцип бессознательной инференции (см. раздел 2.1.4): ученый полагал, что зрительная система бессознательно «знает» (из предыдущего опыта), каково соотношение между коэффициентом отражения поверхности ( $R$ ), освещенностью ( $I$ ) и фотометрической яркостью результирующего светового потока ( $L$ ), воздействующего на сетчатку. Далее ученый рассуждал так, что зрительная система «отыгрывает» назад эту ситуацию, учитывая каким-то образом уровень освещенности сцены: если известна фотометрическая яркость света, падающего на поверхность и фотометрическая яркость света, воздействующего на сетчатку, то коэффициент отражения поверхности легко вычислить, разделив второе на первое ( $R = I/L$ ). Однако слабое место идеи Гельмгольца в том, что зрительная система не располагает информацией об уровне освещенности: ежели бы оно было так, то, вероятно, решить такое уравнение не представляло бы труда, по крайней мере в принципе — в нормальных просмотровых условиях зрение человека весьма нечувствительно к абсолютному уровню освещенности, хотя и может порой уловить легкие изменения (Schubert & Gilchrist, 1992).

Геринг предложил принципиально иное объяснение, полагая, что информация об уровне освещенности сцены извлекается непосредственно из проксимального стимула и без всяких посредников вроде инференций. Геринг полагал, что хотя *абсолютные* величины света, поступающие в глаз от разных участков данной поверхности меняются радикально при смене освещенности, *отношение* между световыми потоками остается неизменным (или почти неизменным). Следовательно, Геринг предположил, что восприятие светлоты зависит от *относительной фотометрической яркости (контраста)* между соседними поверхностями.

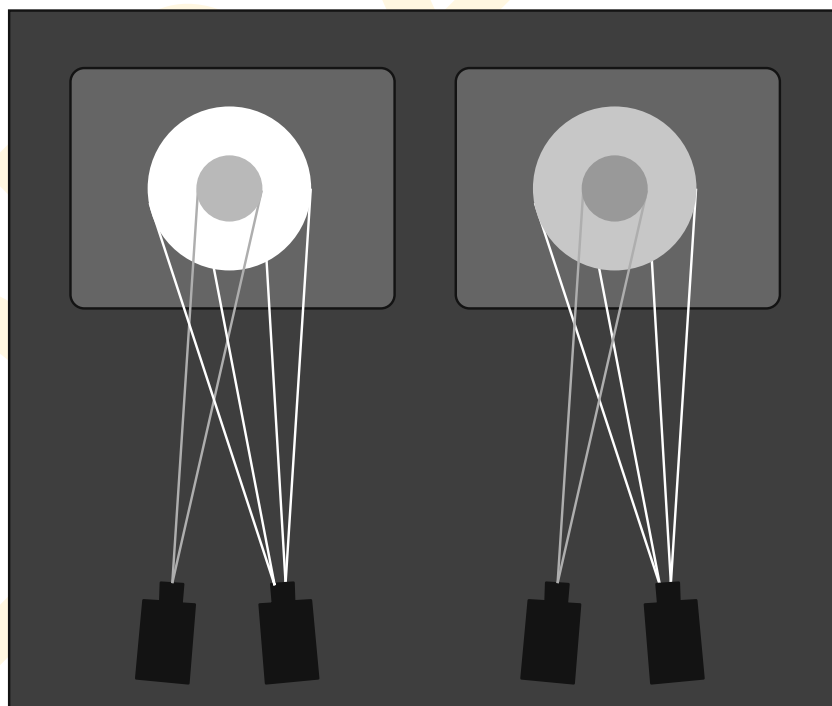
Много позже весьма специфическую версию геринговой идеи предложил американский психолог Ганс Валлах (1948), также утверждавший, что светлота определяется яркостными соотношениями. Чтобы проиллюстрировать, как сие может обеспечить ее постоянство, вернемся к нашему давешнему примеру с запечатанным листом



бумаги (см. рис. 3.3.3), и заметим, что отношение, количества света, отраженного от ее пробельных элементов к таковому от элементов печатных составляет 9:1 и остается неизменным, как в условиях тусклого освещения (90:10), так и на ярком солнечном свете (9000:1000). Следовательно, постоянством соотношений фотометрических яркостей можно попытаться объяснить константность светлоты поверхностей, не предполагая при этом, что зрительной системе нужны дополнительные источники информации о фактической освещенности сцены. Отметим также фундаментальное отличие между бессознательной инференцией и реляционным механизмом: реляционный подход для толкования константности светлоты факторизует освещенность (то есть оперирует величинами, что остаются неизменными при смене освещенности), тогда как бессознательная инференция принимает во внимание освещенность для активной коррекции яркостной информации.

Валлах протестировал реляционный концепт простым, но очень эффективным экспериментом (рис. 3.3.5). Он осветил потоками двух разных интенсивностей соседние участки равномерно-белого экрана. Центральный круглый участок освещался одним прожектором, а окружение — вторым (оба прожектора были скрыты от наблюдателя). Путем раздельного управления фотометрической яркостью обоих источников Валлах манипулировал как абсолютными, так и относительными величинами освещенности экрана по центру и в окружении. При этом комплектов было два: в первом комплекте освещенности центра и окружения были фиксированы, а во втором освещенность центра управлялась наблюдателем. Перед испытуемым ставилась задача: управляя освеще-

Сtimул и окружение  
фиксированной мощности      Сtimул переменной мощности,  
окружение фиксированной мощности



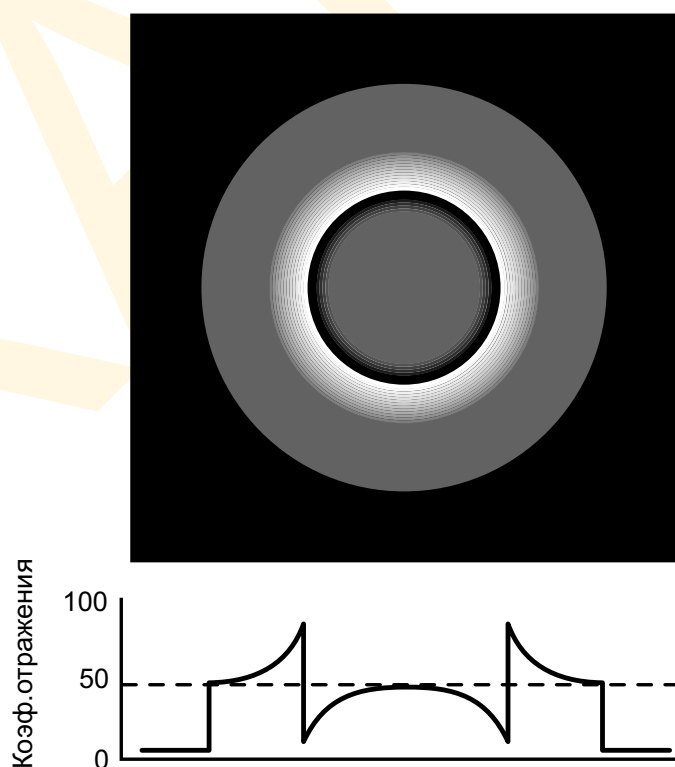
**Рис. 3.3.5** Валлахов эксперимент по истолкованию механизма константности светлоты. Наблюдатели, управляя освещенностью центра во втором комплекте, уравнивали его светлоту со светлотой центрального круга первого комплекта.

ценностью центра во втором комплекте, уравнивать его светлоту со светлотой центрального круга первого комплекта.

Результаты опыта четко «сели» в валлахову реляционную концепцию: уровень освещенности центрального круга во втором комплекте наблюдатели установили таким, что отношение фотометрических яркостей центрального круга и его окружения оказалось равным этому отношению в первом комплекте. Отметим при этом, что, собственно константность светлоты в данном эксперименте напрочь отсутствовала, поскольку отражательная способность поверхности — это отражательная способность равномерно белого экрана. Отличия по светлотам были плодом исключительно отличий в освещенностях (и, следовательно, в фотометрических яркостях отраженных потоков). Однако воспринимались они наблюдателем, как отличия именно в отражательных способностях поверхностей и по причинам, о которых речь пойдет позже.

**Краевые эффекты.** Ежели соотношение фотометрических яркостей световых потоков, воздействующих на орган зрения, определяет восприятие светлоты (как предполагал Валлах), то реляционная концепция должна давать и алгоритм вычисления этих соотношений. Однако прежде следует определиться в том, как производить вычисления — глобально, по всему полю или только по яркостным краям (т.е. резким перепадам световой интенсивности)?

Несколько неожиданно прозвучит то, что наука о зрении интересуется лишь локальными перепадами фотометрических яркостей на краях. На рис. 3.3.6 дан простой, но очень интересный пример, демонстрирующий т.н. иллюзию Крэйка — о'Брайана — Корнсвита: яркостный профиль изображения (показанный внизу) указывает на то, что у центральной области и ее окружения фактически одинаковые фотометрические яркости за исключением внутренней кромки, где локальное изменение яркости образует легко читаемый край; при этом центральный диск выглядит заметно темнее своего окружения, даже глубоко внутри, где яркость та же, что и у окружения. Если яркостный край прикрыть кольцом, вырезанным из непрозрачной бумаги, то визуальная разница между центральным кругом и его окружением моментально исчезнет. Сие свидетельствует о том, что информация от краев определяет воспри-



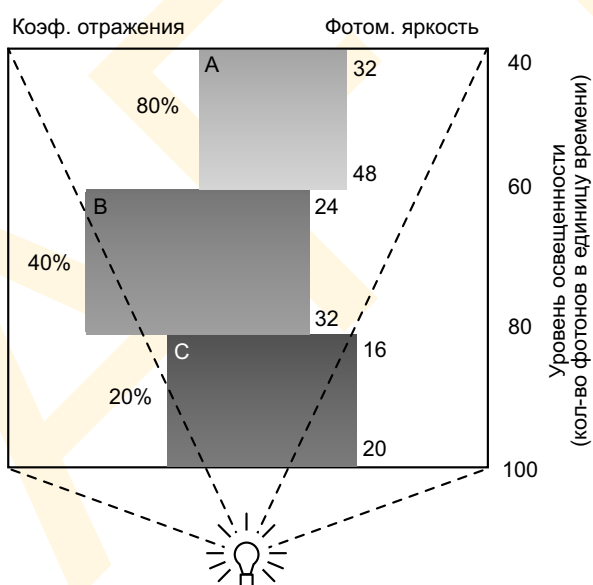
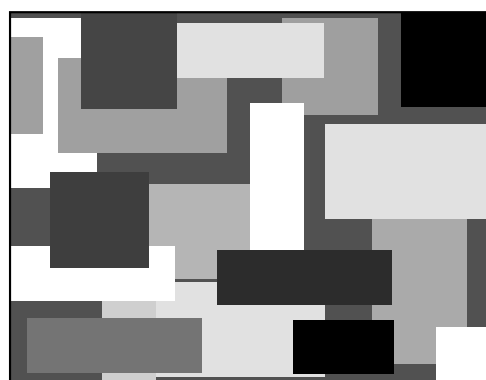
**Рис. 3.3.6** Иллюзия Крэйка — о'Брайана — Корнсвита. Центр внутреннего круга воспринимается темнее периферии круга внешнего, несмотря на то, что фотометрическая яркость у обоих одинаковая. Отличия по яркости имеют место только возле границы между кругами, что хорошо видно по графику внизу.

ятие относительных светлот целых областей, поскольку диссеминируется (распространяется) зрением на все эти области.

**Ретинекс-теория.** Ежели светлота поверхности задается отношением локальных яркостей на краях, значит зрительная система каким-то образом собирает эти отношения по всему изображению. Как сие происходит? Своеобразный ответ предложили американские исследователи Эдвин Герберт Лэнд и Джон Мак-Канн (1971) в т.н. *ретинекс-теории цветовой константности*. Подчеркнем, что теория разрабатывалась для того, чтобы истолковать общую цветовую константность, но, коль скоро константность светлоты — это слабое общее цветовой константности, мы вправе использовать данные Лэнда — Мак-Канна в таком аспекте.

Лэнд и Мак-Канн пытались понять, каким образом зрительная система человека выделяет светлоты из сложных двухмерных массивов стимуляции, названных исследователями *мондрианами* (см. рис. 3.3.7А).<sup>1</sup> В серии весьма вольных экспериментов ученые показали, что элементы мондрианов сохраняют светлоту в широком диапазоне освещений.

В условиях, показанных на рис. 3.3.7В искусственный источник света располагался внизу изображения и на некоем упрощенном мондриане создавал градиент освещенности от низа к верху (внизу ярче — вверх темнее). Поскольку фотометрическая яркость в каждой точке — это математическое произве-



$$\frac{48}{24} \times \frac{32}{16} = \frac{1536}{384} = \frac{4}{1}$$

Вычисление яркостных соотношений на краях

**Рис. 3.3.7** Мондриан в градациях серого. *Вверху:* мондриан наподобие того, что использовали Лэнд и Мак-Канн (1971). *Внизу:* упрощенный мондриан с градиентом освещенности, падающей по мере удаления от источника света. Несмотря на плавное изменение освещенности, соотношение между яркостями на сторонах краев остается таким же, как соотношение коэффициентов отражения поверхностей. Расчет, приведенный внизу, демонстрирует то, как зрительная система может «вычислять» соотношение коэффициентов отражения поверхностей, оценивая яркости на сторонах краев.

<sup>1</sup> Потому что они напоминали картины популярного в те годы голландского абстракциониста Пита Мондриана.

дение освещенности в этой точке и коэффициента отражения в ней, задачей зрительной системы является восстановление отражательной способности каждого из участков, исходя из сугубо яркостной информации. Сие безусловно недоопределенная задача инверсии, поскольку единственное, что есть в наличии — это фотометрическая яркость стимула в данной точке, определяемая освещенностью и коэффициентом отражения поверхности.

Ретинексная теория предполагает, что светлота — плод т.н. *глобальной интеграции* соотношений яркостей на краях. Рис. 3.3.7В проиллюстрирует то, как такая интеграция выполняется в отношении набора из трех участков, чьи коэффициенты отражения взяты с упрощенного мондриана. Отношение коэффициента отражения участка А к таковому у С равно 4:1 (80% / 20%). Те же значения можно получить путем последовательного умножения отношений яркостей на краях между этими участками. То есть отношение яркостей у А и С можно найти, умножив отношение яркостей у А и В (48/24) на отношение яркостей у В и С (32/16). Как уже было сказано, такое вычисление даст соотношение 4:1. Сие демонстрирует то, что яркостная информация на краях достаточна для восстановления относительных коэффициентов отражения по всем участкам. Однако абсолютные величины восстановить невозможно, поскольку освещенность мондриана факторизована. Отметим также, что плавное понижение освещенности от низа к верху мондриана (от 100 до 40 фотонов в единицу времени) не может помешать анализу, потому что отношения фотометрических яркостей берутся только на краях, которые столь узки, что градиентом яркостей можно пренебречь.

Лэнд и Мак-Канн утверждали, что зрительная система человека представляет собой плотную сеть т.н. *ретинексов*, работающую по описанной выше схеме последовательного умножения или какому-то ее эквиваленту. 2D-версия такой схемы была разработана Хорном (1974) и действовала одновременно по всему полю зрения, благодаря чему относительные коэффициенты отражения по всем участкам мондриана можно было восстановить, не завися ни от общих изменений в освещенности, ни от ее градиентов.

**Точка отсчета.** Восприятие светлоты не ограничивается одним лишь определением отношений яркостей. Когда мы рассматриваем поверхности объектов сцены, мы воспринимаем каждую как имеющую определенный уровень светлоты (от черного до белого). Однако яркостные соотношения актуальны только в отношении соседних участков. Плюс к тому, ежели образец отражает в пять раз больше света, чем его фон, мы не сможем понять, то ли сам образец белый, а фон серый, то ли образец этот серый, а фон черный. Одних лишь соотношений недостаточно для ответа на этот вопрос. Сие проблема выбора *точки отсчета* в восприятии светлоты, то есть того, каким образом происходит мэппинг яркостных отношений на ахроматическую ось цветового пространства.

Всегда считалось, что в большинстве ситуаций зрительная система работает как *эвристически фиксированная*: предполагает, что область наибольшей фотометрической яркости — это белое, после чего шкалирует все остальные области поля зрения относительно этого белого (Land & McCann, 1971; Wallach, 1948). Мощное экспериментальное доказательство сего тезиса получено совсем недавно, когда наблюдателям предъявляли мондриан, чьи коэффициенты отражения варьировали в диапазоне от темно-серых до черных. Когда в поле зрения не было ничего, кроме самого мондриана, наиболее яркий участок воспринимался как белый, а наиболее темный — как средне-серый (Cataliotti & Gilchrist, 1995).

Правило наибольшей яркости, выработанное в описанном (и многих других) экспериментах, очевидно, работоспособно, но следует помнить при этом, что участок наи-

большей яркости должен восприниматься хоть и *белым*, но не *самосветящимся*. Классический пример последнего — Луна, поверхность которой по отражательной способности своей — средне-серая и, разумеется, ничего не излучающая, но только отражающая солнечный свет. Однако на ночном небосклоне она воспринимается как самосветящийся объект и недаром именуется «ночным светилом».

Факторы, влияющие на то, что предмет воспринимается самосветящимся весьма сложны и зависят как от пространственной конфигурации объекта, так и от его фотометрической яркости (Li & Gilchrist, 1996). Область малого пространственного размера на темном фоне (как, скажем, Луна) — это идеальные условия для восприятия отражающего объекта самосветящимся.

**Световые и вещественные края.** Давеча мы остановились на том, что, судя по всему, соотношения яркостей на краях определяют восприятие светлоты соседних поверхностей, что эти отношения можно интегрировать по всей сцене и что область с наибольшей фотометрической яркостью обычно воспринимается зрением как белая. И как же теперь мы будем объяснять константность светлоты? К сожалению, проблема эта пока так и остается нерешенной.

Дело в том, что фотометрические яркости по сторонам краев в ретинальном изображении происходят от двух разных видов краев в сцене:

⇒ *Вещественных краев*, которые суть изменения фотометрической яркости, вызванные изменением отражательной способности (коэффициента отражения) двух поверхностей, проекции которых оказываются край в край на сетчатке.

Вещественные края образуются тогда, когда поверхности состоят из разных материалов или по-разному покрашены.

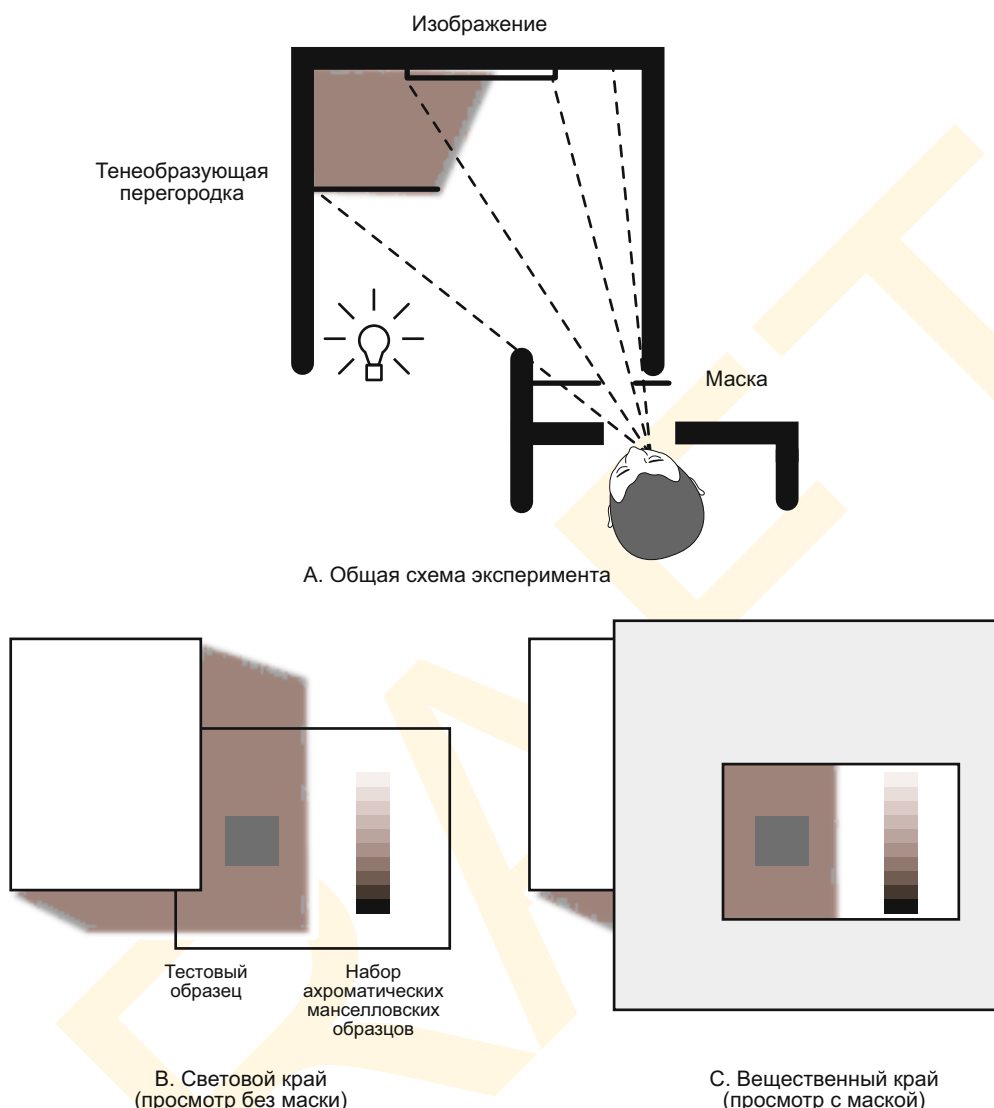
⇒ *Световых краев* — которые суть изменения фотометрической яркости, вызванные разными количествами света, падающего на одну и ту же поверхность с равномерного коэффициента отражения.

Световые края могут появиться, когда на поверхность падает тень, рефлекс от глянцевого поверхностей, направленные световые пятна (скажем, на театральной сцене), либо же при смене ориентации поверхности.

Ретинекс-идея хорошо работает в отношении множества вещественных краев, но начинает давать грубые сбои, когда сталкивается с краями световыми. Причины сего неочевидны: ретинекс-концепт предполагает, что освещение всех краев меняется плавно и потому «не озабочен» расчетом краеобразующих перепадов яркостей. То есть, ретинекс-теория допускает константность светлоты только если яркостные края, представленные в ретинальном изображении — это производное вещественных краев сцены.

То, что генез краев критично важен в формировании ощущения светлоты поверхности весьма изящно продемонстрировал американский физиолог Алан Гилхрист (Рутгерский университет): в одном из экспериментов (1988) ученый поместил наблюдателей в ситуацию, в которой им предъявляли сложный стимул, состоящий из хитроумного сплетения светов и теней (рис. 3.3.8А). Исследователь создал две ситуации: фотометрическая яркость центральной области теста в первом и во втором случаях была одной и той же, но разной у окружений. На рис. 3.3.8В показан край, воспринимаемый как световой; на рис. 3.3.8С — как вещественный. В обоих случаях в левой (темной) половине располагался тестовый квадрат чья светлота измерялась уравниванием





**Рис. 3.3.8** Эксперимент А. Гилхриста по выявлению особенностей восприятия световых и вещественных краев. А — Через специальную маску, либо без нее, наблюдателю предъявлялось наполовину затененное изображение. В — без маски наблюдатели воспринимали границу между тестовым образцом и уравнивающей матрицей как световой край, то есть образованный тенью. С — с маской край воспринимался как вещественный, т.е. как плод разности в коэффициенте отражения поверхностей (Gilchrist, 1988).

с наиболее схожим элементом из клина сравнения справа (светлая область). В обоих случаях наблюдатели выбирали поле клина наиболее схожее с тестовым по критерию светлоты.

В первом случае (рис. 3.3.3В) затемнение фона тестового образца вызывалось наведенной тенью, и, следовательно, край между светлой и темной половинами поля воспринимался световым. Во втором случае (рис. 3.3.8С) маскирующий экран устранял всю информацию из затененной части. Не видя подтверждения тени, наблюдатели воспринимали центральный край как стык поверхностей с разными коэффициентами отражения, т.е. вещественным краем. Таким образом, главное отличие между двумя

ситуациями состояло в том, что центральный край в первом случае воспринимался как край световой (В), а во втором — как вещественный (С). Однако нам важно понять, влияет ли данное отличие на наблюдательскую оценку светлоты поверхности.

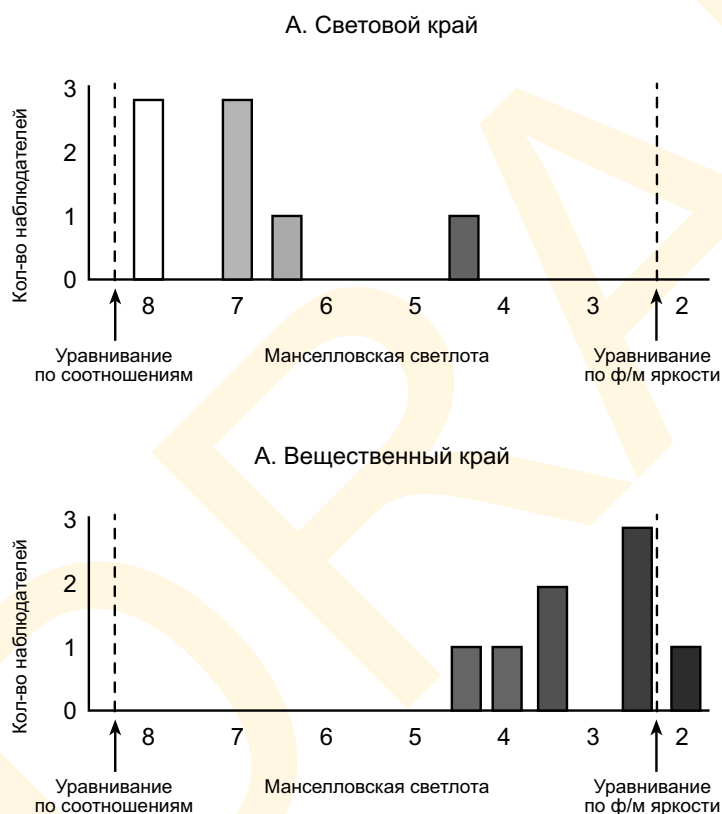
Гилхрист обнаружил огромную разницу между светлотами в первых и вторых условиях (что видно по схеме рис. 3.3.9).

В условиях вещественного края (рис. 3.3.9В) наблюдатели выбрали серый образец, фотометрическая яркость которого была лишь чуть-чуть выше яркости тестового образца. То есть, они очень плотно подошли к константности светлоты, допустив, что центральный край — плод отличия исключительно в коэффициенте отражения при постоянной освещенности обеих половин поля зрения. Испытуемые в этой ситуации, судя по всему, интегрировали яркостную информацию по всему полю зрения, что, собственно, и прогнозирует ретинекс-теория Лэнда — Мак-Канна.

Однако в ситуации светового края (рис. 3.3.9А) наблюдатели выбрали образцы, фотометрическая яркость которых радикально отличалась от таковой у образца тестового. Их выбор основывался на константности светлоты, но при допущении, что центральный край — это исклю-

чительно плод разности в освещенностях. Согласование основывалось на равенстве соотношений фотометрических яркостей тестового образца и его непосредственного окружения без интеграции данных по этому краю.

Разницу между результатами по двум экспериментальным ситуациям невозможно истолковать лишь в логике локальных яркостей, поскольку фотометрические яркости соответствующих областей в центре полей зрения идентичны. Наиболее подходящая интерпретация звучит так, что в ситуации световых краев наблюдатели восприняли центральный край как разницу в освещенностях поверхности по обеим сторонам этого края, и, следовательно, их зрительная система подавила процесс определения относительных светлот (выполняемый путем интегрирования яркостных отношений по этому краю). Вместо этого она допустила, что отражательная способность по обеим сторонам по-



**Рис. 3.3.9** Результаты гилхристового эксперимента. На гистограмме показано число наблюдателей, которые выбрали ту или иную светлоту серого из набора манселловских образцов, которые по их мнению, визуально уравнивали тестовый образец в обоих вариантах просмотра, т.е. при световом и вещественном краях. А — при световом крае большинство испытуемых выбрали светло-серые образцы. В — при имитации вещественного края большинство испытуемых выбрали темные образцы (Gilchrist, 1988).

верхности центрального края одна и та же и что все отличия — результат падения тени. В ситуации вещественного края, однако, яркостные отношения по центральному краю «вычислялись» в ретинексной логике, о чем свидетельствует почти идеальное равенство по яркости.

Итак, мы приходим к тому, что интерпретация светлот в сложном стимуле, содержащем как световые, так и вещественные края, вероятнее всего требует двух отдельных представлений: *карты* областей, отличных количеством падающего на них света, и *карты* коэффициентов отражения по разным областям. Относительные субъективные яркости областей первой карты интегрируются по световым краям сугубо отдельно от краев вещественных. Аналогично относительные светлоты отражающих поверхностей во второй карте интегрируются по вещественным краям отдельно от краев световых.

**Отличие световых краев от вещественных.** Несмотря на сказанное выше, в цепочке рассуждений по-прежнему отсутствует одно звено — ответ на вопрос: «Каким образом зрительная система определяет то, какой край в сцене образован отличиями в освещенности поверхности, а какой отличиями в отражательной способности поверхностей?». Существует несколько эвристических посылов, позволяющих реализовать сие критично важное отличие: *нерезкость*, *планарность*, *краевой контраст* и *хроматика*.

*Нерезкость.* Один из факторов — это резкость или нерезкость края. Световые края, образованные тенями, либо световыми пятнами, обычно слегка размыты. В отсутствие обратной информации зрительная система предполагает, что острый край между соплоскостными областями — это край вещественный. К примеру, если неподвижное световое пятно с острым краем положить на равномерно окрашенную поверхность, зрительная система автоматически интерпретирует его как область с повышенным коэффициентом отражения на фоне с пониженным коэффициентом. Собственно поэтому валлаховы образцы, организованные как центр-окружение (рис. 3.3.5), воспринимаются как отличные по отражательной способности, но не по освещенности, даже несмотря на то, что воспроизведены за счет отличий в интенсивностях света прожекторов.

*Планарность.* Если информация об удалении объектов от наблюдателя указывает на то, что две области не копланарны — край между ними воспримется скорее как световой, нежели как вещественный (даже если окажется острым). Объяснением служит то, что у плоскостей разной глубины расположения и/или ориентации обычно разная освещенность.

*Краевой контраст* — третий фактор, представляющий собой величину отношения фотометрических яркостей на крае: световые края обычно дают намного больший перепад яркостей, нежели края вещественные. Поверхность высокой белизны как правило отражает не более 90% падающих на нее фотонов, а глубоко черная не менее 10%. Следовательно, соотношение коэффициентов отражения обычно не превышает 10:1. При этом перепад освещенностей поверхностей при их затенении может составлять 1000:1 и более. Следовательно, если соотношение яркостей на крае порядка 10:1 и выше, то эвристика срабатывает в пользу отличий по освещенности.

*Хроматика* также помогает зрительной системе различать световые и вещественные края (чего мы вскорости коснемся, когда будем говорить о хроматической константности). Главным фактором, вероятнее всего, является то, что отличия в освещенности всегда оставляют неизменными цветовой тон и отчасти насыщенность, тогда как отличия в коэффициентах отражения обычно влекут за собой и хроматические от-

личия (т.е. отличия по цветовому тону и/или насыщенности). В целом, если цветовой тон или насыщенность варьируют по сторонам края, то, вероятнее всего — это вещественный край; если варьирует только субъективная яркость, то, вероятно, это край световой.

Итак, прежде чем закончить разговор о константности светлоты, вернемся к противостоянию Гельмгольца — Геринга. Вспомним, что Гельмгольц полагал механизмом константности светлоты бессознательную инференцию, основанную на информации об освещенности, тогда как Геринг считал, что константность светлоты базируется на «прямом» реляционном восприятии краевого контраста. На первый взгляд герингов контрастный подход очевидно предпочтительнее (как минимум в валлаховой формулировке и подкрепленный ретинексной теорией Лэнда — Мак-Канна). Однако же эксперименты Гилхриста подняли очень серьезный разговор о том, что края разных типов требуют разных подходов к интерпретации светлоты поверхностей.

Общая же картина такова, что включает в себя и гельмгольцев компонент (в том смысле, что световые эффекты следует учитывать при расчете уровня светлоты поверхностей). Сие не является собственно бессознательной инференцией, основанной на учете абсолютных величин освещенности, как полагал Гельмгольц, но сие необходимо процессу классификации (или интерпретации) краев (как возникших либо в результате изменений в освещенности) либо же в отражательной способности поверхности и последующему выбору алгоритма инференции на базе этой интерпретации. Следовательно, и Геринг, и Гельмгольц оба правы, но каждый по своему, и адекватная теория светлоты требует синтеза обоих подходов. Сей теоретический синтез пока еще не выработан, но очевидно, где-то рядом.

### 3.3.1.2 Хроматическая константность

Давеча мы рассмотрели факты и теории, касаемые ахроматического компонента цветового восприятия, т.е. черно-белой размерности цветовых ощущений. Теперь нам предстоит перенести полученные знания в хроматический домен. К сожалению, добавка еще двух размерностей существенно усложнит и запутает и без того сложный и путанный разговор. Однако это вовсе не значит, что для зрительной системы обеспечение хроматической константности в восприятии объектов — это особо сложная задача, бо хроматическая информация поставляет дополнительные ориентиры, облегчающие достижение общей константности восприятия.

Итак:

⇒ *Хроматическая константность* — это плод восприятия неизменности спектральной отражательной способности поверхности вне зависимости от изменений в условиях освещения этой поверхности и условиях ее просмотра.

Чаще всего мы не задумываемся о хроматической константности как собственно о проблеме, потому что предметы вокруг нас сохраняют свой цветовой тон и насыщенность в очень широком диапазоне освещений, к примеру: в облачную погоду, ясную, утром, днем, вечером, при свете ламп накаливания, люминесцентных лампах, галогенных и пр. — розы остаются красными, фиалки синими, подсолнухи желтыми. Но спектры излучения у разных источников ощутимо разные, в результате чего в глаз наблюдателя от одной и той же поверхности летит свет разных спектральных распределений (рис. 3.3.2).

Важнейшим механизмом обеспечения хроматической константности предметов

является хроматическая адаптация. Если рецепторы сетчатки (и всю последующую нейронную цепь) утомить постоянным воздействием стимула, в котором преобладает, скажем, средневолновый компонент (например, через зеленые солнцезащитные очки), то частота пульсации М-колбочек упадет, весьма активно устранив тем самым избыточное средневолновое влияние на зрительную систему. Сие изменение в восприятии посредством хроматической адаптации — шаг в направлении обеспечения хроматической константности.

Однако адаптация — это не единственный механизм хроматической константности, более того, даже не главный (причины сего те же, что были рассмотрены нами в разговоре о константности светлоты).

**Сужение рамок.** Хроматическая константность, реализуемая зрительной системой человека, ежели не совершенна, то по крайней мере весьма добротна. Хроматические перцепты (результаты цветового восприятия) по каждому участку красочного мондриана большей частью определены спектрами отражения этих участков, причем даже несмотря на то, что свет, воздействующий на сетчатку (спектр излучения), фактически является произведением спектра отражения и спектра падающего света. То есть, зрительная система каким-то хитрым образом способна отделить спектральную информацию о падающем свете от спектральной информации о поверхности. Теоретически — это очередной вопрос инверсии без однозначного решения. Но работа зрительной системы доказывает то, что проблема по большей части решается, поскольку человек вполне успешно пользуется хроматической константностью в самом широком диапазоне условий просмотра. Строя эвристические версии, в целом (но, однако, не всегда) истинные, зрительная система явно добавляет некую информацию к исходной информации от сенсорной стимуляции. Да, сегодня мы толком не знаем, как реализуется хроматическая константность в нашем восприятии окружающего мира, но все же мы потихоньку начинаем понимать, как сие происходит.

Сложность толкования механизма хроматической константности восприятия реальных сцен состоит в том, что по каждой поверхности в поле зрения имеют место шесть неизвестных параметров (три по свойствам падающего света, три по отражательной способности поверхностей), но только три известных показателя по участкам ретинального изображения (отклик трех колбочковых систем на воздействие стимула, который суть плод взаимодействия падающего света с поверхностью). Теоретически задача не имеет решения, поскольку неизвестных больше, чем известных. Однако в естественных условиях существует минимум три аспекта, благодаря которым задача становится не столь недоопределенной — это:

- относительная стабильность состава падающего света в данной сцене;
- однообразие спектров естественных источников света;
- относительное постоянство набора спектральных коэффициентов отражения поверхностей.

Рассмотрим каждый из них.

*Относительная стабильность состава падающего света* — первое и весьма «спасительное» обстоятельство, проявляющее себя тем, что хроматическое воздействие падающего света примерно одно и то же (или порой даже строго одно и то же) по всем областям поля зрения, причем даже тогда, когда интенсивность освещения поверхностей варьирует благодаря теням, расстоянию до источника, ориентации этих поверхностей и пр.

По тем участкам ретинального изображения, у которых уровень яркости неизменен, число параметров сокращается вдвое, к примеру: если в мондриане имеется 10 ко-



планарных участков с различной отражательной способностью, равноосвещенных единичным источником, то ретинальное изображение не станет варьировать по 60 параметрам (шесть по каждому участку), но только по 33 (три по коэффициентам отражения каждого из 10 участков плюс три по постоянному освещению). Отметим, что даже в столь ограниченных обстоятельствах, проблема не имеет однозначного решения, но при этом диапазон возможных решений радикально сокращается. Отметим однако, что приняв такую схему, зрительная система должна как-то отличать световые края от вещественных (см. ниже).

*Однообразие спектров естественных источников света* — второе полезное обстоятельство, состоящее в том, что спектры освещения естественных сцен весьма однообразны.

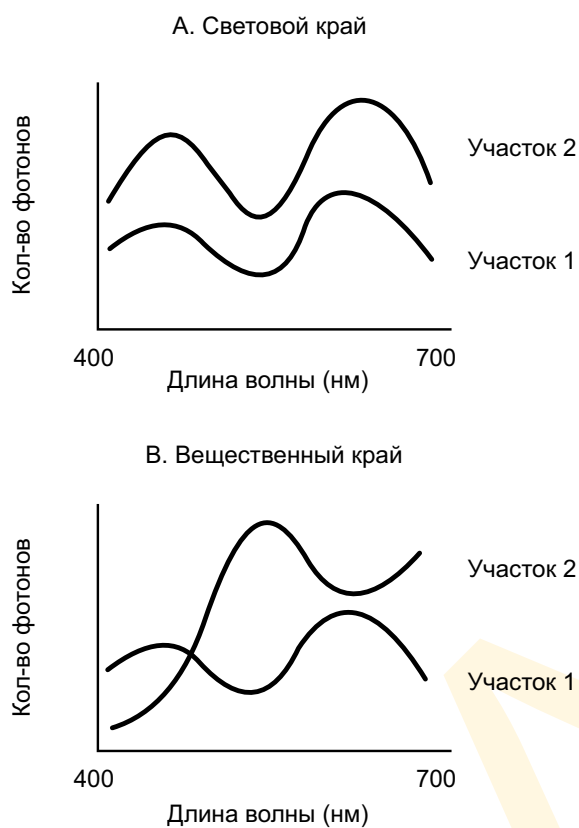
Разумеется, естественное освещение переменчиво (облачно/солнечно; утро/вечер), а искусственное (лампы накаливания и люминесцентные лампы) еще вариативней (см. рис. 3.3.2). Однако строго окрашенный свет встречается исключительно редко — в основном это сценический свет или экспериментальный (например, при исследовании зрительного восприятия). В таких условиях цветовое восприятие дает очень серьезные сбои. Скрамность набора естественных осветителей означает, что в достижении веридикальности восприятия зрительной системе не приходится бороться со всем арсеналом проблем цветовой константности (Judd, McAdam, & Wyszecki, 1964).

*Относительное постоянство набора спектральных коэффициентов отражения поверхностей* — третье обстоятельство, проявляющее себя тем, что диапазон спектральных коэффициентов отражения поверхностей, встречающихся в естественных условиях, невелик, а те, что имеются обычно гладкие (не зубчатые) и легко разбиваются на небольшое число групп (Cohen, 1964).

Американские нейрофизиологи Лоренс Мэлони и Брайан Ванделл (Стэндфордский университет) разработали алгоритм достижения приблизительной хроматической константности на базе всех трех перечисленных факторов (Maloney & Wandell, 1986). Алгоритм этот предполагал наличие единичного общего осветителя и моделировал различные источники и различные поверхности как линейную комбинацию лишь из двух (или из трех) базовых спектров освещения и базовых спектров отражения. Исследователям, в итоге, удалось продемонстрировать эффект, схожий с хроматической константностью в цветовом восприятии окружающего мира человеком (Brainard & Wandell, 1991, 1992).

**Световые и вещественные края.** Обеспечение хроматической константности восприятия объектов окружающего мира, в конечном счете, требует той же визуальной информации, что и обеспечение константности светлоты. В состав этой информации входят различия между световыми и вещественными краями. Поскольку основными данными, необходимыми для выявления этих различий, являются хроматические свойства ретинального изображения, рассмотрим последние.

Американцы Джон Рубин и Витман Ричардс (1982) выяснили, что вещественные края представлены на сетчатке, как правило, совсем иными спектрами излучения, нежели края световые. Важным моментом является то, что простое изменение в освещенности, какое возникает при затенении или при смене ориентации поверхности, создает на сторонах края спектры излучения качественно очень близкие друг к другу (рис. 3.3.10А). Однако изменения в спектральном коэффициенте отражения поверхностей, какие возникают при смене материала или пигмента, приводят к образованию качественно весьма отличных спектров излучения отраженного от этих поверхностей



**Рис. 3.3.10** Спектральные свойства световых и вещественных краев. А — спектральные распределения энергии по обеим сторонам светового края в целом сходны между собой. В — спектральные распределения энергии по сторонам вещественного края ощутимо разнятся. (Rubin & Richards, 1982).

возникает перекрест спектральных графиков (Rubin & Richards, 1982).

Оппозитный наклон, показанный на рис. 3.3.11В, возникает при наклоне графиков в разных направлениях, то есть тогда, когда один идет вверх, другой вниз (Rubin & Richards, 1989).

Тот факт, что сии два варианта условий независимы друг от друга, демонстрируется всеми четырьмя графиками рис. 3.3.11. Мы видим, что только последний вариант (рис. 3.3.11D) указывает на отсутствие как перекреста, так и оппозитного наклона, иллюстрируя тем самым случай, когда спектральные отличия между участками скорее всего образуют световой край.

Действительно ли зрительная система отличает световые края от вещественных по схеме Рубина — Ричардса или нет, до сих пор неизвестно. Дело в том, что цветовое зрение человека не может выполнить полноценной спектрометрии стимула, воздействующего на сетчатку по причинам, которые мы обсуждали в прошлой главе. Однако у нее имеются два самостоятельных хроматических канала — красно-зеленая и сине-желтая оппонентные системы, которые могут являться биологическим механизмом, работающим по схеме Рубина — Ричардса. В целом, вполне возможно, что усло-

вета (рис. 3.3.10В). Коль скоро зрительная система способна различать характер краевых спектров, это значит, что она может отличать световые края от вещественных в том числе за счет хроматической информации. Как уже отмечалось в разговоре о константности светлоты, сия информация оказывается необычайно полезной в зрительной реконструкции отражательной способности видимых поверхностей.

Рубин и Ричардс предложили два критерия, маркирующих изменения в спектральном коэффициенте отражения поверхностей: *спектральный перекрест* и *оппозитный наклон*. Критерии эти определены отношением между двумя спектрами световых потоков, образованных парой интересующих нас поверхностей сцены, но оцененных лишь по двум длинам волн. Идею проще всего растолковать по соответствующим графикам — рис. 3.3.11А и 3.3.11С.

Спектральный перекрест появляется тогда, когда спектральные измерения по двум участкам выявляют оппозитные отличия в энергии световых потоков по выбранной паре длин волн, в результате чего

вия отличия световых краев от вещественных «обсчитываются» дубль-опponentными клетками, кратко описанными в разделе 3.2.3. Отметим, что сие весьма грубые количественные показатели, поскольку представляют отличия лишь по красноте/зелени и синеве/желтизне между соседними участками сцены.

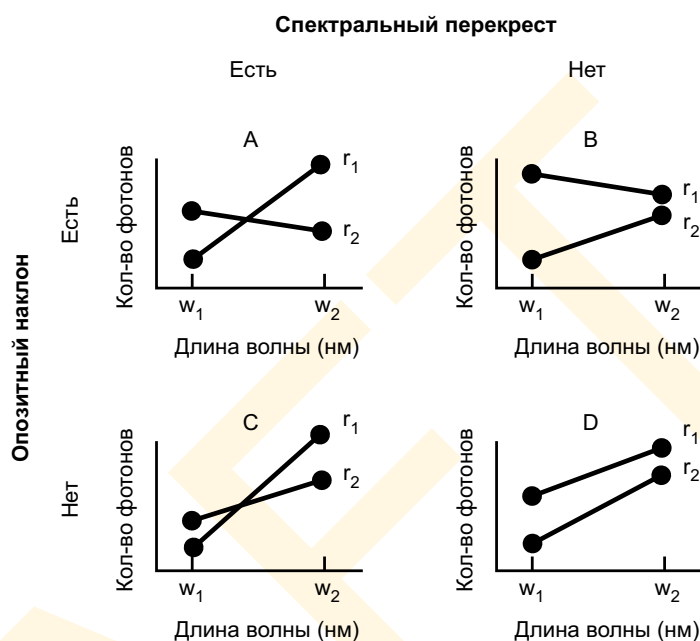
**Эволюционная подоплека (гипотеза Шепарда).** Американский оптометрист Карл Шепард (1992) продумал эволюционное толкование трехмерности цветового зрения и его опponentной структуры, основанное на цветовой (в частности, хроматической) константности. Ученый предположил, что цветовое зрение человека трехмерно потому что дает особи цветовую константность восприятия объектов внешнего мира в рамках естественной подвижности дневного света, по трем степеням свободы (Judd et al., 1964), то есть по трем перцепционным направлениям:

1. *Светлее-темнее.* Физически задается общим количеством света — от полуденного солнца до глубокой тени при луне.

2. *Сине-желтый баланс.* Физически задается балансом относительных количеств коротких («синих») волн и прочих с центром в «желтой» части спектра (определяется рассеянием коротковолнового света молекулами воздуха).

3. *Красно-зеленый баланс.* Физически задается балансом относительных количеств длинных («красных») волн и прочих с центром в «зеленой части» (определяется угловой высотой Солнца и количеством водяных паров в воздухе, поскольку пары эти селективно поглощают длинноволновый свет).

Коль скоро эволюционная роль цветового зрения в том, чтобы позволить особи стабильно воспринимать спектральную отражательную способность поверхностей при вариациях дневного света, то один из путей достижения сего — это эффективное устранение этих вариаций зрительной системой внутренними компенсаторными механизмами. Таким образом, черно-белая, красно-зеленая и сине-желтая размерности цветового зрения вполне могут оказаться эволюционно детерминированными, призванными компенсировать естественную динамику дневного света. При этом стоит отметить, что рассуждения Шепарда, хоть логичны и эффективны, но пока еще не получили надежного эмпирического подтверждения.



**Рис. 3.3.11** Спектральная эвристика Рубина — Ричардса, позволяющая различать световые и вещественные края. Соседние участки ( $r_1$  и  $r_2$ ) сэмплированы по двум длинам волн ( $w_1$  и  $w_2$ ) и рассортированы согласно тому пересекаются отрезки (A и C) или нет (B и D). К световому краю относится только D-вариант (Rubin & Richards, 1982).

### 3.3.1.3 Онтогенез цветовой константности

Константность светлоты в цветовом восприятии объектов окружающего мира и хроматическая константность не являются врожденными признаками. И хотя сей вопрос только-только начали изучать в экспериментах по селективному рассматриванию и адаптации, результаты свидетельствуют, что хроматическая константность и константность светлоты отсутствуют у новорожденных, но развиваются в первые месяцы жизни.

Наиболее продуктивными экспериментами по изучению онтогенеза цветовой константности явились опыты, поставленные на базе т.н. *адаптационной парадигмы* (Dannemiller, 1989): детей адаптировали к определенному стимулу и до тех пор, пока те не уставали смотреть на один и тот же предмет. После чего им либо вновь предъявляли тот же самый стимул, либо другой. При этом время, потраченное ребенком на тестовый стимул, замерялось. Стандартный результат был таков, что дети охотнее смотрели на новый стимул, нежели на тот, которым были адаптированы. Таким путем удавалось определить, может или нет ребенок увидеть разницу между двумя вариантами стимуляции.

Американский исследователь Джеймс Даннемиллер (1989) использовал эту методику для изучения цветовой константности у детей 2-4 месячного возраста: малышей вначале адаптировали стимулом, исходящим от поверхности двойной окраски (т.е. с двумя различными спектрами отражения), освещенной дневным светом (рис. 3.3.12А). После чего им предъявляли одно из четырех тестовых изображений (рис. 3.3.12В): ту же поверхность при том же освещении («исходный» стимул); ту же поверхность, но под другим освещением («поверхность-константный» стимул); другую поверхность, но под исходным освещением («осветитель-константный» стимул); наконец, другую поверхность при другом освещении («новый» стимул).

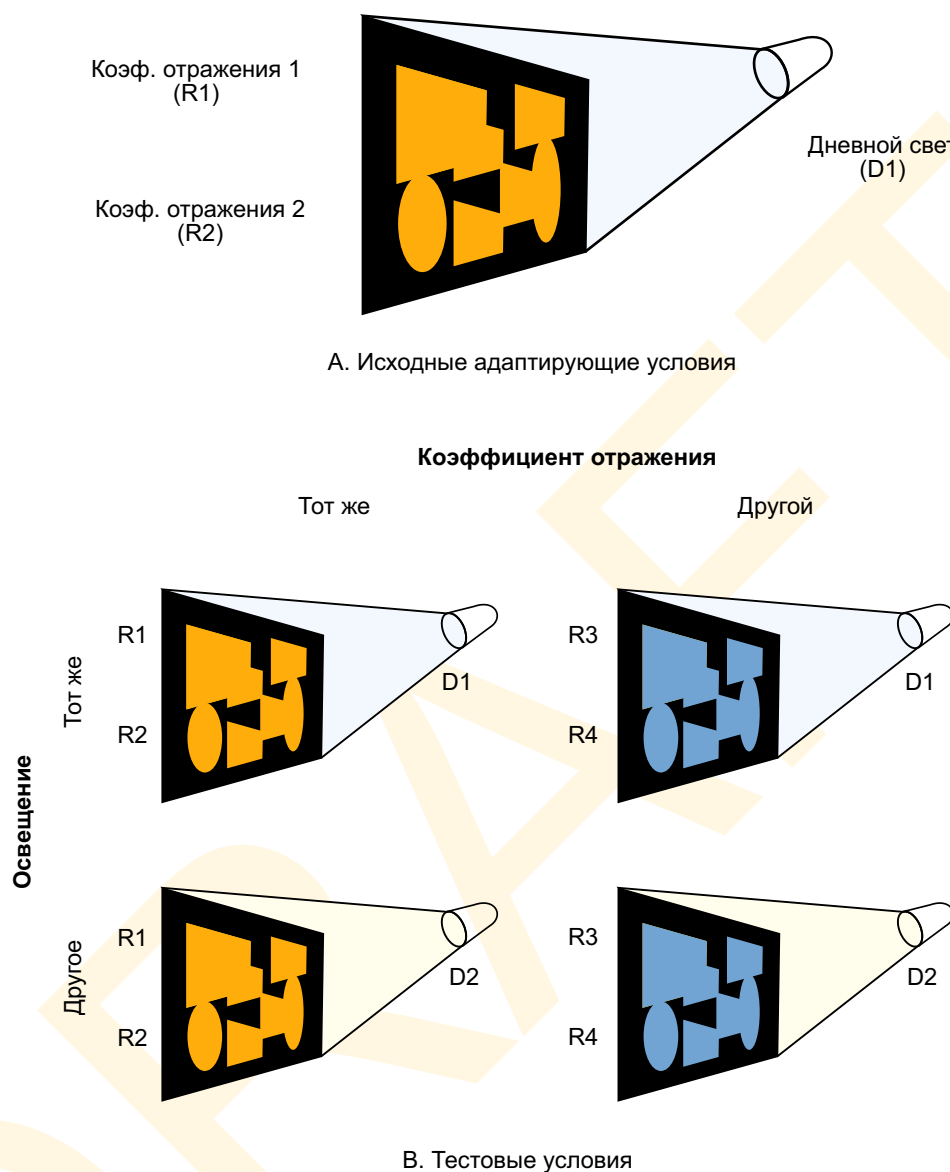
Двухмесячные малыши существенно дольше глядели на «поверхность-константный» тест, нежели на «исходный» — так долго они глядели только на совершенно «новый» стимул. То есть, картина визуальных откликов свидетельствует о том, что у детей столь раннего возраста отсутствует хроматическая константность.

Четырехмесячные дети дали совершенно иной результат. Они глядели на «поверхностно-константный» тест не так долго как на «исходный», но при этом долго смотрели на «осветитель-константный» и «новый» тесты. Сие четко указывает на наличие у четырехмесячных младенцев ожидаемой хроматической константности в двух разных условиях освещения.

Даннемиллер пришел к выводу, что у человека хроматической константности нет от рождения до 2-4 месячного возраста.

## 3.4 КАТЕГОРИЙНЫЙ ЭТАП ЦВЕТОВОГО ВОСПРИЯТИЯ

Когда хроматические свойства поверхности или светового потока восприняты и внутренне представлены, это еще не значит, что процесс цветовой перцепции завершен: практически всегда человеческое восприятие выходит за рамки чистых ощущений и классифицирует воспринятое, как принадлежащее к той или иной категории, после чего обычно уже категоризированному перцепту дает определенную словесную метку. К примеру, свежую кровь, спелый помидор и некоторые сорта яблок мы называем в целом «красными», даже несмотря на то, что их зрительные перцепты обладают своими специфическими цветовым тоном, насыщенностью и светлотой. Траву, листву

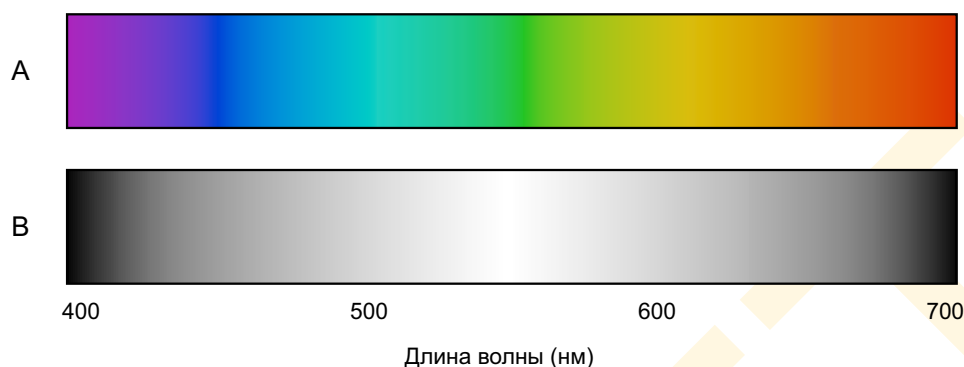


**Рис. 3.3.12** Эксперимент по изучению онтогенеза цветовой константности. А — ребенка адаптировали комплексным стимулом, полученным путем отражения дневного света (D1) от поверхности с двумя разными спектральными коэффициентами отражения (R1 и R2). В — после адаптации предъявляли одно из изображений при исходном (D1) и при новом освещении (D2).

и некоторые другие сорта яблок мы классифицируем как зеленые, невзирая на порой сильные отличия по цветовому тону и насыщенности их перцептов. И сие лишь два примера того, как человек разделяет непрерывное трехмерное цветовое пространство на дискретные цветовые категории.

Перцепционно значимые цветовые категории легко проиллюстрировать на примере радуги (рис. 3.4.1 А). Почему мы воспринимаем радугу состоящей из окрашенных полос? В физической логике радуга — это всего лишь плавный энергетический градиент от волн длиной в 400 нм до волн в 700 нм. Если бы ее перцепт оказался по цветовому тону таким же плавным и недифференцированным, то мы увидели бы «ахроматиче-





**Рис. 3.4.1** Сравнение градиентных перцептов, стимулированных одним и тем же физическим континуумом при полноценном цветовом зрении (А) и при полной цветовой слепоте (В). Полосчатое восприятие волнового градиента нормальным цветовым зрением — базис категоризации цветовых перцептов.

скую радугу» (рис. 3.4.1 В) Но очевидно, что сие не так. Отдельные полосы, отличные друг от друга по цветовому тону — плод деления перцепта радуги на качественно разные категории. И хотя категории эти плавно переходят одна в другую, мы четко воспринимаем хроматическую разницу между ними. Данный феномен закладывает экспериментальный базис словесного деления или как говорят *вербализации* цвета как минимум по нескольким категориям.

Вторым важным фактом из сферы цветовой категоризации является то, что в отношении цветовых категорий у нас на вооружении имеется несколько общих словесных символов: «красный», «синий», «желтый» и т.д. и, похоже, что сие далеко не случайно. Однако остается не совсем понятным, что же первично: перцепционное деление цветов на категории или все-таки вербализация этих цветов? Действительно, мы категоризируем цвета так, а не иначе, потому что у нас уже имеются их названия или же мы даем названия согласно категориям цветовых перцептов зрительной системы? В данном разделе мы рассмотрим ряд доказательств того, что большинство базовых цветовых категорий в отношении цветовых названий психологически первичны. Мы также обсудим внутреннее устройство этих категорий и то, как они соотносятся друг с другом. Но поскольку этап цветовой категоризации жестко связан с этапом цветовых именовании мы не допустим грубой ошибки, если начнем с последних.

### 3.4.1 Вербализация цветовых ощущений

Понимание того, как люди описывают цвета словесно, очень и очень важно. Такие слова в языке, как «красный», «зеленый», «черный», «белый», «розовый», «коричневый», «голубой», «красно-коричневый», «золотой», «серебряный», «лимонный», «оранжевый», «лиловый», «малиновый» — это малая толика из словаря цветовых названий. Однако вообразим дюжину разноцветных карандашей или пару сотен красок из коробки: есть ли в этом хаосе хоть намек на какой-то порядок?

Прорывом в нашем понимании цветовой категоризации явилось *межкультурное исследование цветовых названий*, то есть того, как представители разных языков и культур именуют цветовые ощущения. Американские когнитивные антропологи

Брент Берлин и Пауль Кэй (Беркли) в 1969 г. провели и опубликовали фундаментальное исследование по этой проблеме.

До выхода в свет работы Берлина и Кэя вопрос вербализации цветовых ощущений относили к т.н. *культурологическому релятивизму*, суть которого в том, что каждая культура (и язык) фиксируют свою собственную идиосинкразическую (уникальную) структуру всего разнообразия цветовых ощущений.

У лингвистов культурный релятивизм озвучивается в форме гипотезы Сапира — Хорфа, названной так в честь горячих приверженцев релятивизма Эдварда Сапира и Бенжамина Ли Хорфа. Одним наиболее ярким примером концепции Сапира — Хорфа является то рассуждение, что в языке эскимосов порядка дюжины названий оттенков льда и снега. Согласно Сапиру и Хорфу для обозначения тех или иных категорий в каждой культуре рождаются собственные слова, если в том возникает необходимость, к примеру: различать виды льда и снега эскимосам намного важнее, чем, скажем австралийским аборигенам, которые снега и льда никогда не видели.

Последователи культурологического релятивизма не только утверждали роль культуры в формировании структуры цветовых категорий, но напрочь отвергали при этом значимость физических, физиологических и психологических факторов в этом процессе. Одна из наиболее известных монографий гласит:

«В видимом спектре наличествует непрерывный градиент оттенков, которые американцы описывают такими словами как “красный”, “оранжевый”, “желтый”, “зеленый”, “синий”, “фиолетовый” — или как-то еще. Но ни в самом спектре, ни в его восприятии человеком нет ничего, что заставляло бы нас называть их именно так» (Gleason, 1961).

Такие утверждения были полностью опровергнуты исследованиями Берлина и Кэя, которые начали поход на релятивистскую точку зрения, сосредоточившись на малом наборе цветовых названий. Они предложили концепцию *базовых цветовых наименований* как лингвистических описаний цвета, отвечающих следующим требованиям:

1. Базовое цветовое наименование должно представлять собой т.н. *моноксему* — единичный лексический элемент (“красный”, «розовый», «коричневый» и пр.). Данный критерий исключает такие цветовые описания как «сине-зеленый», «светло-синий» и т.п.

2. Базовое цветовое наименование должно выступать в роли т.н. *первичного хроматического эталона*, т.е. в первую очередь относиться к цветовому перцепту объекта, нежели к самому объекту или материалу из которого тот выполнен. Данный критерий исключает такие термины как «золотой», «серебряный», «лимонный» и т.п.<sup>1</sup>

3. Базовое цветовое наименование должно быть сугубо *целевым*, т.е. применимым к широкому спектру объектов, что исключает использование таких слов как «блондинка» (цвет волос и только) или «гнедая» (масть лошади и только).

4. Базовое цветовое наименование должно регулярно употребляться в языке, что исключает такие малопонятные цветовые определения как «охра», «ультрамарин» и т.п.<sup>2</sup>

<sup>1</sup> В этой логике очевидно следует исключить и такое прилагательное как «оранжевый», но авторы сделали исключение, поскольку существуют культуры, где «оранжевый» является базовым цветовым наименованием вне его сопряжения с собственно плодом (апельсином) или аналогично окрашенным предметом.

В английском языке в рамках изложенных критериев набор базовых наименований сокращается до одиннадцати: «красный», «зеленый», «синий», «желтый», «черный», «белый», «серый», «оранжевый», «фиолетовый», «коричневый» и «розовый».<sup>1</sup> Существует несколько пограничных вариантов, таких как «беж» и «хаки», но мы все же будем следовать схеме Берлина — Кэя.

Когда исследователи непосредственно проанализировали 20 языков и еще 78 через литературные источники, выявились две исключительно любопытных закономерности. Оказалось, что в языках существуют:

1. *Шестнадцать базовых цветовых наименований.* Однако вначале Берлин и Кэй (1969) выделили только 12 базовых цветовых терминов по всем языкам: 11 в английском плюс один означающий светло-синий (аналогично тому как в английском светло-красный именуется розовым). Впоследствии список добавили еще четыре (которые в английском не попадают в разряд базовых): *теплый* (красный или синий), *холодный* (синий или зеленый), *светло-теплый* (белый, или красный, или желтый) и *темно-холодный* (черный, или синий, или зеленый) (Kay & McDaniel, 1978).

2. *Двадцать два подмножества по 16 базовым цветовым наименованиям.*

Более того, полученные данные не были случайными, но подчинялись строгой зависимости, показанной на рис. 3.4.2.

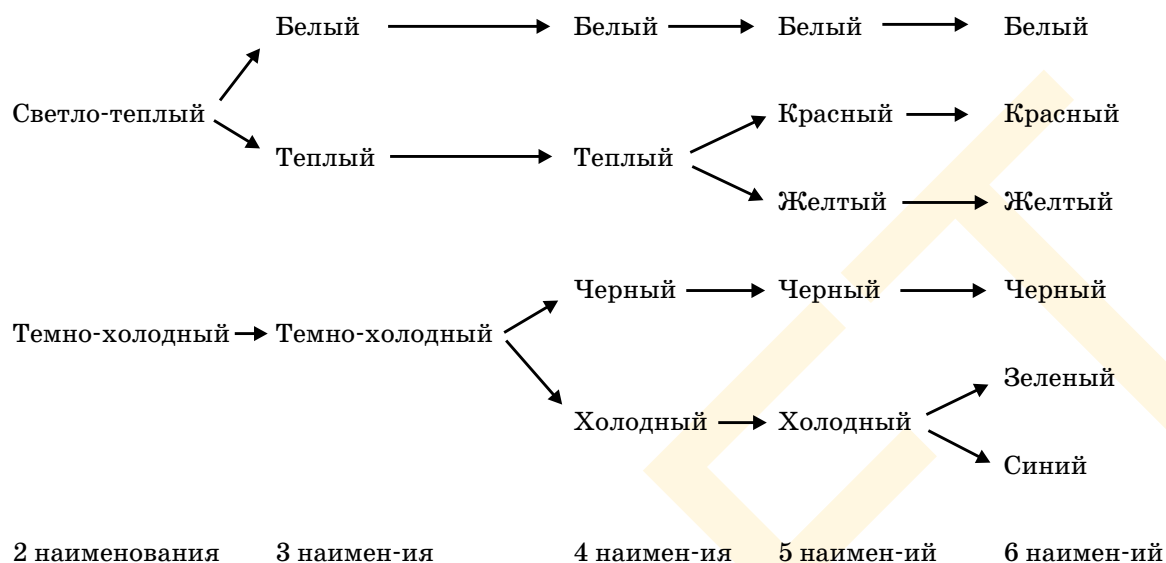
Правило 22 подмножеств можно интерпретировать как последовательность появления и накопления в языках цветовых наименований. Рис. 3.4.2 послужит примером того, что ежели в некоем примитивном языке имеются лишь два базовых цветовых наименования, то это будут непременно «светло-теплый» и «темно-холодный»; если их три, то это всегда будут «белый», «теплый» и «темно-холодный»; и так далее. Каждый этап развития добавляет к списку одно новое базовое наименование. В итоге, появилась еще одна гипотеза, гласящая, что данная схема представляет временную последовательность, согласно которой в развивающемся языке появляются новые слова.

Данные Берлина и Кэя несколько подмывают концепцию культурологического релятивизма, поскольку совершенно не согласуются с ней: действительно, чистый культурологический релятивизм никак не прогнозирует экспериментальных результатов, добытых этими исследователями. Факты, явно опровергающие релятивистскую точку зрения, одновременно утверждают альтернативную идею *лингвистической универсальности* — гипотезу о том, что язык определен либо инвариантными физическими характеристиками материального мира, либо инвариантными биологическими особенностями организма (или и того, и другого). В нашем контексте очевидна биологическая детерминанта, к примеру: первые шесть наименований соответствуют полярным позициям герингового оппонентного процесса (черный/белый, красный/зеленый и синий/желтый). Однако из-за чего цветовые наименования выстроились в определенной последовательности — остается не совсем понятным, но, вероятно, на то есть веские физиологические причины.

Итак, убедительным доказательством концепции Берлина — Кэя является то, что

<sup>2</sup> Эти слова часто употребимы в профессиональной среде, в частности там, где заняты смешением красок тонких, изысканных оттенков, в частности среди интерьер-дизайнеров, художников и т.п. Большинство людей при этом плохо понимают, что имеется в виду или вообще не понимают, о чем речь.

<sup>1</sup> В русском языке очевидно можно выделить термин «голубой», который в английском чаще всего звучит как «blue-green» и, соответственно, не попадает в список Берлина — Кэя. — Прим. пер.



**Рис. 3.4.2** Упрощенное «генеалогическое древо» шести базовых цветовых наименований. Данная последовательность демонстрирует закономерность, с которой базовые цветовые наименования, обнаруживаемые в разных языках формируют комплект из шести наименований. Следует отметить, однако, что из правила имеются исключения.

базовая часть структуры цветового именования фиксирована генетически детерминированной физиологией. Однако данный вывод вовсе не отрицает того, что существуют дополнительные аспекты цветового именования, определяемые все же именно культурой. К примеру, такие слова как «бирюзовый», «золотой» и «серебряный» невозможно соотнести с цветом, если соответствующих предметов нет в культуре данного народа.

### 3.4.2 Фокальные цвета и прототипы

Одной из ключевых проблем, поднятых исследованием Берлина и Кэя, является механизм цветовой категоризации. Классическая точка зрения на этот вопрос — это позиция Аристотеля, утверждавшего, что принадлежность к той или иной категории определяется набором необходимых и достаточных условий. Такая позиция предполагает, что во всех случаях категории равновелико «хороши» (то есть, равновелико репрезентативны), поскольку каждая обладает всеми требуемыми характеристиками. Вывод, который следует из работы Берлина и Кэя, отличен тем, что «лучшие» или «наиболее репрезентативные» примеры, именуемые *прототипами*, играют главную роль в формировании цветowych категорий. Рассмотрим доказательство сего тезиса.

Берлин и Кэй в двадцати исследованных языках выявили, какие цвета с какими базовыми цветовыми наименованиями согласуются. Исследователи применили два метода — *метод граничных* и *метод фокальных цветов*.

В *методе граничных цветов* они предъявляли испытуемым последовательность из 329 максимально насыщенных манселловских образцов (т.е. координаты которых ле-

жат на поверхности цветового тела) и просили своих информантов определить границы категорий, отобрав в эти категории подходящие образцы.

В *методе фокальных цветов* исследователи предъявляли испытуемым те же образцы, но просили их выбрать наиболее подходящее базовое цветовое наименование. Любопытно, что испытуемые с гораздо большей скоростью и достоверностью подбирали подходящее наименование, нежели определяли границы категории. Подходящие названия Берлин и Кэй окрестили *фокальными цветами* и убедительно показали, что цветовые категории структурируются именно вокруг них, нежели за счет сужения категориальных границ.

Идея цветовой категоризации, основанной на прототипах, получила мощное подтверждение в классических исследованиях американского когнитивного психолога Элеонор Рош (в девичестве Хайдер). Рош изучала разные аспекты цветовой категоризации у жителей племени Дани в Новой Гвинее. У дани набор базовых цветовых наименований минимален, всего лишь два: «мола» и «мили», примерно означающие «светло-теплый» и «темно-холодный» соответственно. Рош также изучала цветовую память дани, предъявляя им манселловские образцы и прося выбрать по памяти последний предъявленный образец из большого набора патчей. Опираясь на концепцию прототипов, Рош обнаружила, что дани лучше справляются с заданиями (то есть, совершают меньше ошибок), когда работают с фокальными цветами, нежели когда пытаются очертить границы цветовых категорий (Heider, 1972).

Рош также применила фокальный и граничный методы к исследованию того, как дани усваивают новые цветовые наименования. Она обнаружила, что новые категории они воспринимают легче, когда в центр категории помещается фокальный цвет, название которому имеется в других языках (например «красный», «зеленый», «синий» и «желтый»), нежели когда задаются границы категории. Результаты исследования убедительно показали то, что осваивать новые категории заметно легче, когда категория формируется вокруг фокального цвета (Rosch, 1973a) — что целиком согласуется с прототипической точкой зрения.

В ряде экспериментов Рош опробовала прогностическую пригодность теории цветowych прототипов. К примеру, в одном из них исследователь просила американских студентов так быстро, как только те смогут, ответить «да» или «нет» на вопрос о соответствии предъявляемого им цветового образца названной перед этим категории. Результаты показали то, что испытуемые гораздо быстрее отвечали «да» по фокальным цветам, нежели по остальным (Rosch, 1975a). Было замечено также, что латентные периоды откликов систематически росли по мере удаления цвета образца от прототипа.

В еще одном эксперименте Рош просила студентов разложить цветовые образцы по принципу подобия. Исследователь обнаружила, что человек оценивает нефокальные цвета как чем-то похожие на фокальные, но не наоборот. Дополнительные исследования показали, что дети вначале запоминают названия фокальных цветов, и лишь потом периферийных (Mervis, Catlin, & Rosch, 1975).

В качестве обоснования этих (и ряда других результатов) Рош предложила тезис о том, что в каждой цветовой категории имеется т.н. *когнитивная эталонная точка*, относительно которой прочие цвета осознаются, категоризируются и кодируются в памяти.

Все результаты Рош указывают на то, что прототипы играют ключевую роль во внутреннем представлении цветовых категорий. Членство в цветовой категории вовсе не определяется необходимостью и достаточностью условий, как полагал Аристотель. Как мы покажем в гл. 9, концепция прототипов как стандартная теоретическая кон-



цепция внутренней структуры естественных категорий человеческого разума, распространяется на большое число других естественных категорий и сегодня полностью вытеснила аристотелеву точку зрения.

### 3.4.3 Модель цветовой классификации

Многие научные данные о базовых цветовых категориях и названиях можно истолковать в логике *модели цветовой классификации*, предложенной в 1978 г. американскими исследователями Паулем Кэем и Чадом Мак-Дэниелем (Беркли). В их модели принадлежность к той или иной цветовой категории формализована в показателях т.н. *теории нечетких множеств*, представляющей собой некоторую переработку классической теории нечеткой логики, разработанной Лютфи Заде (также в лоне Беркли — см. гл. 2). Для понимания этой теории мы сделаем короткое «лирическое отступление», объясняющее природу нечетких множеств.

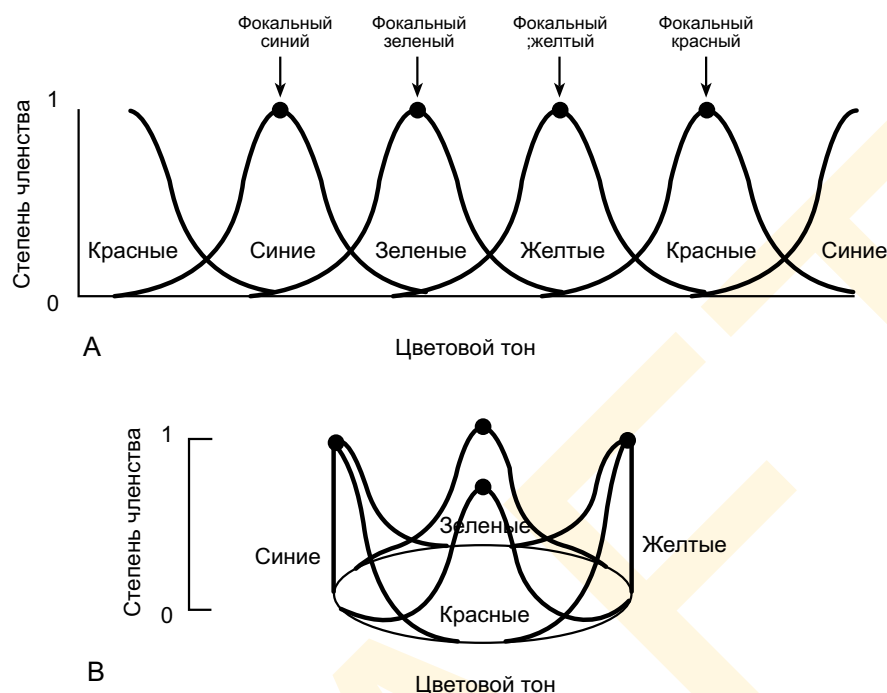
#### 3.4.3.1 Теория нечетких множеств

Теория классических множеств (множеств выбора) согласуется с аристотелевым представлением о категорийной принадлежности, когда каждый объект либо является членом данного множества, либо не является (при этом никаких промежуточных вариантов нет). То есть, принадлежность к множеству выбора можно истолковать как функцию, присваивающую объекту значение 1 — если объект этот удовлетворяет требованию необходимости и достаточности членства в множестве, и 0 — если нет. Теория нечетких множеств, разработанная Заде, несколько видоизменяет теорию классических множеств, позволяя объектам быть членами множеств лишь *до некоторой степени*. Степень членства в категории определяется функцией, присваивающей каждому объекту значение из непрерывного диапазона от 0 (точно вне категории) и 1 (прототип). То есть, в отличие от классической, теория нечетких множеств способна представлять промежуточные уровни членства в множестве.

В логике цветовых категорий сие значит, что фокальный цвет (прототип) в своей категории представлен со степенью членства 1. У прочих, нефокальных цветов степень членства систематически снижается по мере удаления от фокального цвета в цветовом пространстве. В упрощенном виде сказанное проиллюстрирует рис. 3.4.3 по базовым категориям синих, зеленых, желтых и красных. Отметим, что у цветов, расположенных между прототипами, имеется некий промежуточный уровень членства более чем в одной категории, что отражает тот факт, что сине-зеленые (голубые) — это «частичные» члены синей и зеленой категорий со степенью членства, зависящей от их относительной приближенности к синему или зеленому фокальным цветам. Сие естественно в логике нечетких множеств и совершенно неважно множествам классическим. Коль скоро количество времени, которое уходит на определение принадлежности цветового ощущения к той или иной категории, обратно пропорционально *нечеткому значению* этого ощущения по данной категории — высокие уровни членства дают малое время реакции (что подтверждают результаты Рош: фокальные цвета классифицируются наблюдателями много быстрее, чем нефокальные).

#### 3.4.3.2 Категории первичных, производных и композитных цветов

Кэй и Мак-Дэниел (1978) предположили, что шестнадцать базовых цветовых наименований можно разложить на три категории.

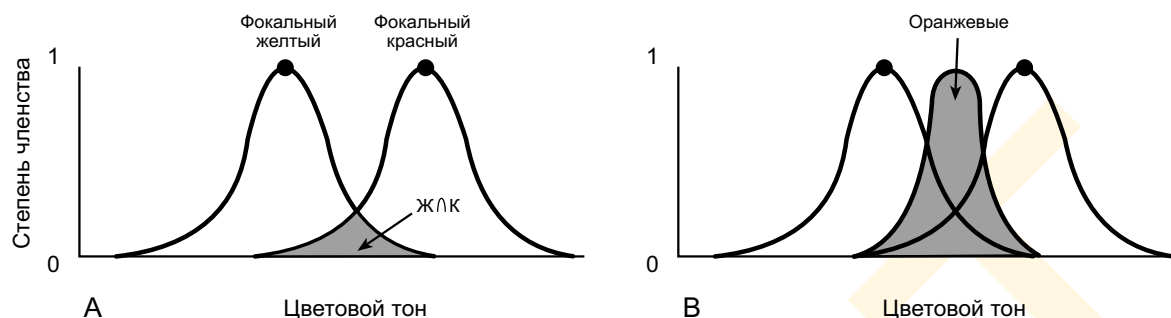


**Рис. 3.4.3** Представление нечеткого множества первичных цветовых категорий. У фокальных цветовых тонов степень членства равна 1.0, а у нефокальных степень членства монотонно снижается с одновременным удалением от «своего» фокального цветового тона. Все цветовые тона (за исключением фокальных) в какой-то мере принадлежат двум первичным категориям. А — размерность цветового тона для простоты, представленная в линейном виде для простоты иллюстрации круговой схемы В.

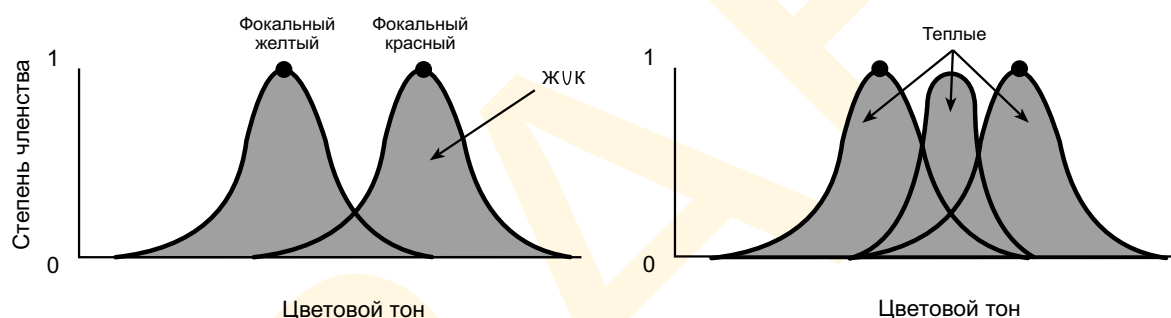
**Категории первичных цветов.** К первой группе категорий — *категории первичных цветов* — отнесли шесть первичных цветовых ощущений, т.е. шесть перцепционных примитивов зрительной системы: *черный, белый, красный, зеленый, синий и желтый* (напомним, что сие в аккурат геринговвы оппозицы).

Прочие базовые цветовые наименования можно проанализировать, применив нечеткую логику к категориям первичных ощущений. Такое применение суть нечеткая версия стандартных логических операций «И» и «ИЛИ».

**Категории производных цветов.** Если  $X$  и  $Y$  — это два нечетких множества, то операция « $X$  И  $Y$ » определена их *нечетким пересечением*. То есть, нечеткое пересечение двух нечетких цветовых категорий определено их перехлестом (рис. 3.4.4 А). В теории Кэя — Мак-Дэниела нечеткое пересечение несколько видоизменено (рис. 3.4.4 В): единичный максимум приходится на точку, где обе членские функции пересекаются и распадаются на обе стороны; там, где они окончательно сходят на нет, там находятся точки ноля. Таким образом, всем членам нечеткой категории оранжевых, которые определены через «красный И желтый», необходима определенная мера красного и желтого, но большинство срединных членов — это «лучшие» члены оранжевой категории, нежели красной или желтой. Согласно анализу Кэя — Мак-Дэниела оранжевый — это нечеткий перекрест «красного И желтого»; серый — нечеткий перекрест «черного И белого»; фиолетовый — «красного И синего»; розовый — «красного И белого»; и, наконец, коричневый — «желтого И черного» (поскольку фактически коричневый — это желтый или оранже-



**Рис. 3.4.4** Нечеткологическое представление вторичных цветовых категорий. Категории, которые меньше первичных (категория оранжевых, например) можно анализировать как зону перекреста первичных категорий (А), но с той поправкой, что они нормированы к единице и со сглаженной формой (В).



**Рис. 3.4.5** Нечеткологическое представление композитных цветовых категорий цветовых. Категории, что больше первичных (скажем, такую как «теплые»), можно рассматривать как нечеткое объединение первичных категорий.

вый низкой светлоты). Исследователи назвали эти категории *категориями производных цветов*, поскольку цветовые ощущения, принадлежащие к этим категориям, получены путем нечеткого перекреста первичных цветовых ощущений.

**Категории композитных цветов.** Оставшиеся базовые цветовые наименования — это наименования из разряда слаборазвитых цветовых именных систем — «теплый», «холодный», «светло-теплый» и «темно-холодный», которые Кэй и Мак-Дэниел называли *категориями композитных цветов*. В нечеткой логике они определены операцией «ИЛИ», которая суть нечеткое объединение двух категорий первичных цветовых ощущений. Нечеткое объединение двух нечетких множеств — это попросту максимум двух компонент членских функций (рис. 3.4.5). То есть, теплый — это нечеткое объединение «красного ИЛИ желтого»; холодный — нечеткое объединение «синего ИЛИ зеленого»; светло-теплый — «белого ИЛИ красного»; и, наконец, темно-холодный — «черного ИЛИ синего ИЛИ зеленого».

Нечеткологическая модель цветовых категорий базируется на большом числе исследований в сфере цветового именования и цветовой категоризации. Более того, она

изящно и очень плотно связывает сии высокоуровневые когнитивные функции с физиологией цветовой обработки в зрительной системе человека.<sup>1</sup>

Нельзя не сказать о том, что разговор о нечеткой логике в контексте беседы о цветовом восприятии — это блестящий пример преимущества междисциплинарного подхода в исследовании: разговор был бы невозможен без данных из физики, психологии, физиологии, компьютерной науки, лингвистики и когнитивной антропологии, объединив которые мы колоссально расширяем и углубляем наше понимание механизмов зрительного восприятия. Все, что мы будем обсуждать далее не столь развито в этом смысле, но успех междисциплинарного подхода в изучении цветового восприятия несомненно принесет плоды и в иных сферах нашей науки (что мы увидим далее).

---

<sup>1</sup> Соответствие феноменов цветового именования обсуждавшимся ранее биологическим механизмам не столь уж строгое, как то следует из теории Кэя — Мак-Дэниела. Ученые утверждали, например, что шесть первичных базовых цветовых наименований соответствуют операциям оппонентной обработки в сетчатке и колленчатом теле. Однако сие совершенно непохоже на те ощущения, которые возникают у человека по причинам, которых мы коснемся в 13-й главе, когда будем говорить о т.н. зрительной осведомленности. Осознанные ощущения в зрительном акте появляются существенно позднее, после того, как достигается цветовая константность восприятия. Но даже в таком случае структура цветового представления на столь высоком уровне, всегда, когда это возможно, походит на красно-зеленую, сине-желтую и черно-белую оппонентную схему ранних этапов зрительного акта.

Часть вторая

**ПРОСТРАНСТВЕННОЕ**

**ЗРЕНИЕ**





Удивительно, но цветовое восприятие окружающего мира (см. гл. 3) — это всего лишь *одна из* потенциалов зрения, и, ежели рассуждать в эволюционном ключе, то далеко-далеко не главная: у многих животных, особенно ночных, например, у кошек и крыс, цветовое зрение почти не развито, что никак не мешает их выживанию как биологических видов. Очевидно, что эволюционно более значимой является способность особи воспринимать *пространственную структуру* окружающего мира: форму, местоположение, размер и ориентацию предметов сцены. Сие тема текущей главы.

О пространственном зрении говорят в контексте детекции и анализа зрительной системой вариаций фотометрической яркости по полю зрения, в частности, когда речь идет о разрешении деталей (острота зрения), а также восприятии краев, яркостных паттернов, форм и текстур.

Разговор о пространственном зрении мы, разумеется, построим на базе четырех этапов восприятия, коротко описанных нами во второй главе. Напомним: *оптический этап, фасетный, предметный и категорийный*. Однако материал, касающийся пространственной обработки столь обширен, что расположится на шести главах. В текущей речь пойдет о сугубо об оптическом этапе, т.е. о тех зрительных процессах, что касаются начального анализа двумерной стимуляции сетчатки, но не психологической реконструкции трехмерных свойств поверхностей и объектов. Исследователи зрения в целом единодушны в том, что необходим ряд предварительных процессов, прежде чем ретинальное изображение окажется интерпретированным в логике экологических сущностей, таких как поверхности и объекты.

В науке о зрении существовали три принципиальных подхода к изучению процессов оптической пространственной обработки: *количественный, психофизический и физиологический*. Подходы эти основывались на принципиально отличных друг от друга методах (см. гл. 2), но имеющих весьма сходные цели. Согласно марровскому анализу трех уровней обработки информации (см. раздел 2.3) идеальный способ воедино увязать эти подходы — это начать с общего описания на количественном уровне, продолжить в аспекте психофизическом (ограничившись при этом алгоритмами обработки ретинального изображения) и закончить объяснением физиологической организации этих алгоритмов на уровне нейронном. Такой схемы мы придерживались в разговоре о цветовом зрении (и она отлично работала, поскольку механизмы цветового восприятия логично рассматривать именно во временном домене).

Однако в данной главе мы отойдем от принятого ранее правила и начнем с разбора физиологических аспектов пространственной обработки ретинального изображения, поскольку к тому нас подталкивает хронология научных событий: первый шаг был сделан в 50-е и 60-е годы прошлого века, когда на свет появились методы экстраклеточной регистрации активности единичных нейронов — методы сугубо физиологиче-

ские. Как ни печально звучит, но даже спустя почти 50 лет и, несмотря на колоссальные научные усилия, функциональное назначение нейронов многих типов так и остается невыясненным. Впрочем, сие закономерно, и недаром Марр настаивал на том, что прежде всего должен выполняться количественный анализ происходящего.

Сегодня существует несколько параллельных теорий, толкующих функцию нейронов — этих загадочных крошечных клеток. В текущей главе мы коротко опишем многотрудный путь, который пришлось пройти, прежде чем стала более или менее понятной их роль в зрительном акте. Исчерпывающего ответа, разумеется, нет до сих пор, но определенный прогресс налицо. Отметим, наконец, что фрагменты единой научной картины зрения собираются исследователями порой посредством принудительно разных научных подходов.

Заканчивается текущая глава разговором о самых последних исследованиях в сфере нейронной обработки ретинального изображения. В разделе 2.3.3 мы коротко коснулись фодоровской модулярной гипотезы, гласящей, что восприятие строится на совокупной работе независимых (или почти независимых) подсистем, специализирующиеся на обработке разных видов информации. В контексте рассуждений о работе зрительной системы эта гипотеза звучит так, что оптический этап восприятия окраски объекта, его формы, движения и объема реализуется в разных зрительных путях (Livingstone & Hubel, 1987). Физиология зрительных путей до крайности сложна и путана, психофизические ее толкования противоречивы, но при этом вполне вероятно, что общая структура оптического этапа строится именно на работе специализированных проводящих путей.

## 4.1 АНАТОМИЯ И ФИЗИОЛОГИЯ ПРОСТРАНСТВЕННОГО ЗРЕНИЯ

Как было показано в гл. 1, первые знания о биологической подоплеке зрения человечество добыло при изучении различных видов патологии, возникшей вследствие повреждения головного мозга при травмах, опухолях и огнестрельных ранениях (Ferrier, 1878; Glickstein, 1988). Обнаружилось, что участки мозга, отвечающие за зрительное восприятие, расположены в районе затылочной коры (или где-то поблизости). Однако, несмотря на это важнейшее открытие, почти ничего не было известно о специфике нейронных событий при зрительном акте. С появлением в 50-х годах прошлого века техники *экстраклеточной регистрации активности единичных нейронов* (см. раздел 2.2.3) стали поступать первые сведения о характере нейронной информационной обработки и появилось новое научное направление, клетка за клеткой постигающее механизмы зрения.

Итак, нам предстоит познакомиться с материалом, который дали нам перечисленные исследования в плане обработки визуальной информации на оптическом уровне.

### 4.1.1 Нейрогистология сетчатки

Разговор о физиологии зрительного акта мы начнем с пространственной обработки на уровне сетчатки. Напомним, что в первой главе мы коротко описали различные слои ретинальных нейронов, выполняющих пространственную и временную обработку информации, прежде чем нейронный сигнал выйдет из глаза по зрительному нерву. Сегодня настало время рассмотреть эти слои детально. Сие необходимо не только потому, что дает понимание происходящего на ранних этапах зрительного акта, но еще и

по той причине, что открытие нейронных слоев — это предтеча эпической работы Хьюбела и Визеля по изучению корковых механизмов пространственного зрения.

Итак, сетчатка — это тонкий слой нервной ткани, выстилающей внутреннюю поверхность глаза, по ряду причин крайне интересная исследователям и клиницистам. Ретинальные клетки развиваются из тех же клеток, что и клетки мозга, следовательно, сетчатка является составной частью центральной нервной системы: действительно, сетчатку можно рассмотреть с помощью офтальмоскопа, то есть, это единственный фрагмент центральной нервной системы, который можно непосредственно наблюдать без хирургического вмешательства. По этой причине сетчатка весьма информативна при визуальной диагностике различных заболеваний нервной системы, а также сердечнососудистых заболеваний.

Разумеется, сетчатка, в первую очередь, выполняет множество зрительных функций. В предыдущих главах мы коснулись путей кодирования светового сигнала фоторецепторами, а в данной главе поговорим о том, как ретинальные нейроны конвертируют рецепторный сигнал в нервные импульсы, отправляемые в мозг.

Межклеточные связи и нейронная специализация сетчатки таковы, что образуется множество параллельных нейронных проводящих путей, несущих в мозг различные виды информации о ретинальном изображении.

Наконец, чувствительность фоторецепторов и прочих ретинальных нейронов меняется в зависимости от изменений в окружающем освещении. Поскольку отдельные нейроны управляют своей чувствительностью при изменении уровня внешнего освещения, сетчатка эффективно передает визуальную информацию в широком диапазоне условий: от лунного света до яркого солнечного дня, то есть в диапазоне, превышающем шесть порядков.

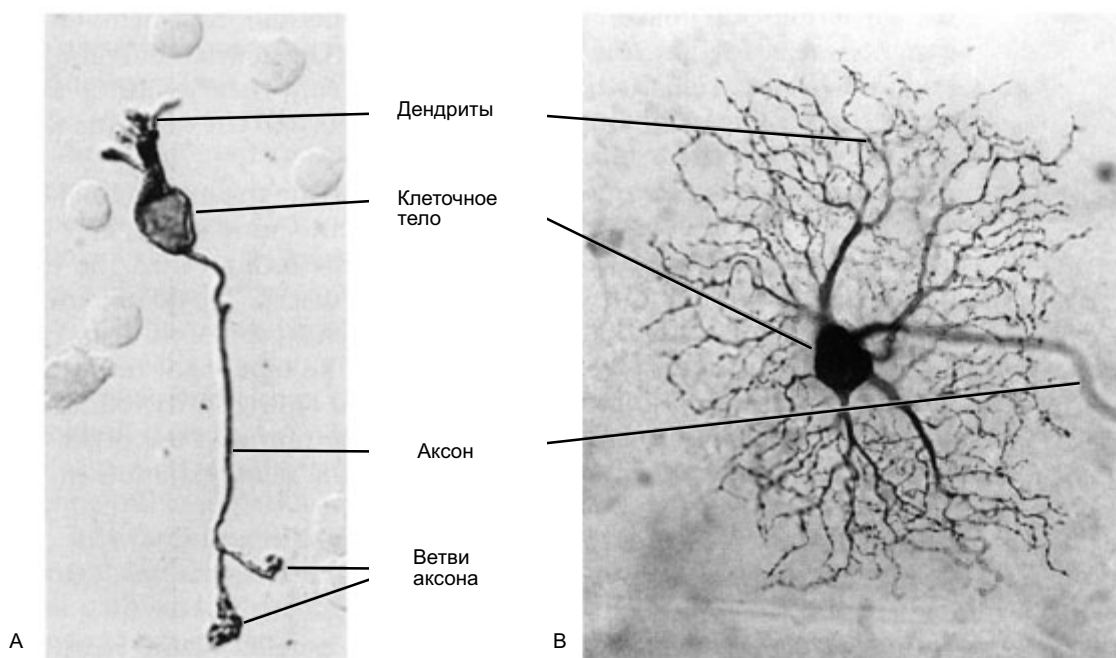
Изучив то, как сетчатка организует информацию о проекционном ретинальном изображении наблюдаемой сцены и отвечает на изменения в условиях просмотра, мы очень многое узнаем о том, каким образом мозг организует и интерпретирует эту информацию.

Почти на всем своем протяжении сетчатка приматов в толщину примерно 0.5 мм и вспомним, что состоит она из трех клеточных слоев и двух слоев, содержащих синаптические связи между нейронами (рис. 1.3.7 и 1.3.8). Возле оптической оси глаза в сетчатке приматов, напомним, имеется специальный участок — центральная ямка (fovea) — состоящий лишь из одного слоя нейронов — колбочковых фоторецепторов — колбочек (рис. 1.3.13 и 1.3.14).

На рис. 4.1.1 показаны ретинальные нейроны двух типов, а также их составные элементы: *дендриты*, *клеточные тела* и *аксоны*. Кроны дендритов, принимают сигнал от других нейронов; аксон (который иногда ветвится) несет выходной сигнал нейрона к нейрону целевому. Форма дендритовой кроны и ветвей аксонов — важнейший признак принадлежности нейрона к тому или иному классу, для классификации ретинальных нейронов мы будем использовать и другие признаки — положение клеточного тела, дендритов и аксонов; размер и форму клеточного тела и дендритовой кроны; характер связи с другими нейронами.

#### 4.1.1.1 Нейронный состав сетчатки

Вспомним, что существует пять базовых категорий ретинальных нейронов, но у каждой имеется несколько подкатегорий. Главные категории отличаются друг от друга местоположением клеточных тел нейронов, дендритовых крон и терминалей аксонов.



**Рис. 4.1.1** Ретинальные нейроны различаются по форме и размеру. А — клеточное тело биполярных клеток располагается во внешнем ядерном слое. Их дендриты могут контактировать с фоторецепторами и горизонтальными клетками, а их аксоны несут выходной сигнал биполяров во внутренний плексиформный слой, где контактируют с кронами дендритов ганглиозных клеток. В — клеточные тела ганглиозных клеток залегают в соответствующем слое (слой ганглиозных клеток). Аксоны ганглиозных клеток формируют оптический нерв. Ганглионары различаются по свойствам кроны дендритов, межсоединениям и форме клеточных тел. Клетки, показанные здесь, названы в свое время Стивеном Поляк «зонтичными» или «парасоль» (1941, 1957).

**Фоторецепторы.** Клеточные тела фоторецепторов расположены во *внешнем ядерном слое* сетчатки, а их синаптические терминалы контактируют с дендритами биполярных и горизонтальных клеток во *внешнем плексиформном слое*.

**Биполярные и горизонтальные клетки.** Клеточные тела биполярных и горизонтальных клеток расположены во *внутреннем ядерном слое*. И дендриты, и терминалы аксонов горизонтальных клеток образуют связи с клетками *внешнего ядерного слоя*. Биполярные клетки, однако, образуют связи с дендритами ганглиозных клеток *внутри внутреннего плексиформного слоя*. Поскольку только биполярные клетки связывают сигналы внешнего и внутреннего плексиформных слоев, все визуальные сигналы в обязательном порядке проходят через биполярные клетки.

**Амакриновые клетки.** Тела амакриновых клеток также расположены во *внутреннем ядерном слое*. Амакриновыми эти клетки назвал великий испанский физиолог, лауреат нобелевской премии Сантьяго Рамон-и-Кахаль, дабы подчеркнуть отсутствие у них аксонов и наличие одних лишь дендритов. Дендриты амакриновых клеток образуют связи с дендритами ганглиозных клеток во *внутреннем плексиформном слое*.

**Ганглиозные клетки.** Тела ганглиозных клеток расположены в слое ганглиозных клеток и кроны их дендритов соединены с терминалами аксонов биполярных клеток и дендритами амакриновых клеток.



Аксоны ганглиозных клеток сетчатки дают только ее выходной сигнал и, как мы уже говорили выше, образуют оптический нерв и покидают сетчатку в точке, именуемой *оптическим диском* или *слепым пятном* (поскольку в нем нет фоторецепторов).

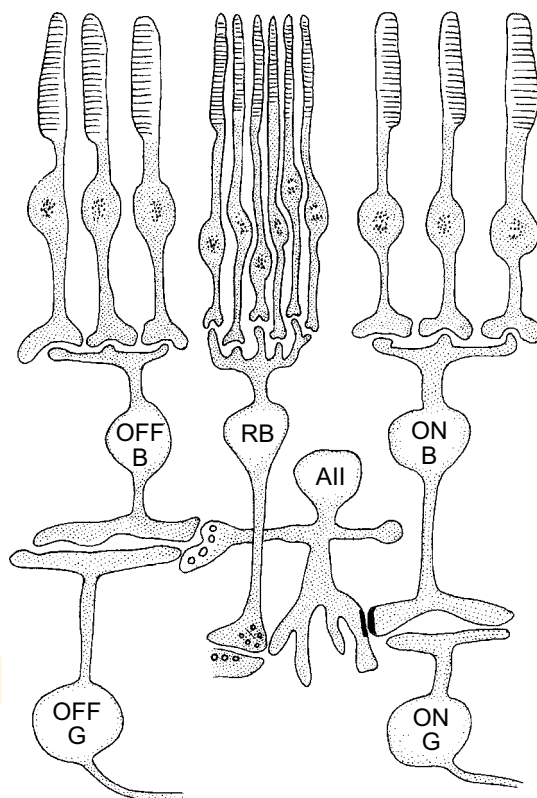
#### 4.1.1.2 Специализация ретинальных функций

Сетчатка направляет зрительную информацию по параллельным нервным проводящим путям, специализированным на решении разных зрительных задач. Ранее мы уже рассматривали один из примеров нейронной специализации в сетчатке: вспомним, что существуют два типа фоторецепторов, реагирующих на проекционное оптическое изображение сцены, — палочки и колбочки. Рецепторы кодируют это изображение в разных диапазонах его интенсивности.

Сортировка палочковых и колбочковых сигналов реализуется несколькими синаптическими связями.

В 1994 г. профессор Ютского университета Хельга Кольб с сотрудниками показала, что палочковые сигналы идут по отдельному *палочковому проводящему пути сетчатки*, пока не достигнут ганглиозных клеток (рис. 4.1.2). Палочки образуют связи биполярными клетками определенного класса — *палочковыми биполярами*, собирающими отклики от множества палочек. Судя по всему, объединение сигналов от множества палочек повышает чувствительность палочкового пути. Палочковые биполяры образуют синапсы непосредственно с *амакриновыми клетками типа АII*, но в отличие от прочих биполяров палочковые не образуют прямых синапсов с ганглиозными клетками.

Синапсы палочковых биполяров формируют очень тонкий подслой внутри внутреннего плексиформного слоя. Наконец, АII-амакриновые клетки образуют синапсы с ганглиозными клетками на том же уровне во внутреннем плексиформном слое, а так-



**Рис. 4.1.2** Палочковый проводящий путь в сетчатке позвоночных. Аксоны более чем 20 палочек сходятся на одной палочковой биполярной клетке (RB). Палочковые биполяры посылают выходной сигнал в специальные амакриновые клетки (АII), расположенные во внутреннем плексиформном слое. Клетки АII связывают палочковый сигнал с ганглиозными клетками двух типов — дендриты ганглионаров одного типа находятся во внутреннем плексиформном слое, другого — в подлежащем слое. Ганглиозные клетки с кронами дендритов в верхнем слое (OFF G) также получают входной сигнал от биполярных клеток, чей отклик подавлен световой стимуляцией (OFF B); ганглиозные клетки, крона дендритов у которых расположена в подлежащем слое (ON G), тоже получают входной сигнал от биполярных клеток, возбужденных световой стимуляцией палочек (ON B).

По материалам Wassle & Boycott, 1991.

же на втором уровне в этом же слое. Те ганглиозные клетки, что соединены с АП-клетками, получают сигнал и от колбочкового проводящего пути. Следовательно, палочковый проводящий путь в данной точке стыкуется с колбочковым и теряет свою самостоятельность.

Исследуя свойства палочкового пути, мы легко можем отследить основные особенности организации прочих зрительных путей:

— во-первых, палочковый путь простирается всего на несколько синапсов, объединяющих затем свой сигнал с главным зрительным сигналом. В результате, вышестоящие отделы могут обрабатывать колбочковые и палочковые сигналы, не нуждаясь в их сепарации. Таким образом, перед нами пример того, как зрительный путь решает визуальную задачу на «расстоянии» нескольких синапсов, а затем пропадает;

— во-вторых, мы видим, что зрительные пути могут строиться для выполнения в определенном смысле элементарных функций, в частности таких как расширение некоторых аспектов информации в ретинальном изображении. В палочковом проводящем пути весьма высока конвергенция сигналов: единичный палочковый биполяр может объединить сигналы от 1500 палочек, делая тем самым палочковый путь весьма и весьма подходящим к захвату визуальной информации на низких уровнях освещенности сцены, но проигрывая при этом в показателях остроты зрения. Сей пример демонстрирует то, что анатомические связи в отдельных зрительных путях специализируются на решении сугубо количественных задач. В данном случае конвергенция связей в единый палочковый путь интегрирует световую энергию, рассеянную по сетчатке, и повышает абсолютную светочувствительность системы.

Также, некоторые виды поведения могут потребовать точной зрительной информации определенного рода, скажем, при слежении за движущимся объектом. Чтобы повысить эффективность механизма нам может потребоваться проводящий путь с исключительным высоким временным разрешением. Другие психофизические реакции требуют зрительной информации иного сорта, скажем, для идентификации текстурных паттернов. Сие может потребовать запуска в работу проводящего пути с высоким пространственным разрешением.

В 1993 г. профессор Роберт Родик (Вашингтонский университет, 1993) установил, что внутри сетчатки берут начало до 20 проводящих путей и что каждый из этих путей посылает сигналы в разные отделы нервной системы. При этом подобно палочкам и колбочкам каждая субкатегория ганглиозных клеток несет в себе полную копию ретинального изображения. То есть, у нас есть основания предполагать, что каждая субкатегория ганглиозных клеток специализируется на передаче зрительной информации определенного вида.

Итак, о сериях связанных друг с другом нейронов, параллельно несущих определенную визуальную информацию, мы будем говорить как о *зрительных проводящих путях* и текущих по ним *зрительных потоках*. Точное место в сетчатке, где информация делится на зрительные потоки неизвестно, но, судя по всему, такое деление начинается немедленно на синапсах палочек и колбочек: фактически существует порядка 15-20 различных биполярных клеток, контактирующих с каждой колбочкой, информация, закодированная каждым биполяром, может служить стартовой точкой зрительного потока, текущего по тому или иному зрительному пути (Rodieck et al., 1993; Boycott & Wassle, 1991).

Несомненно, что одной из важнейших функций сетчатки является распределение по зрительным потокам кодированной фоторецепторами информации. Полагают, что целью деления на потоки является эффективная связь информации, заложенной в ре-

ринальное изображение, с соответствующими областями мозга, отвечающими за тот или иной вид зрительной обработки.

Мы предполагаем, что каждый зрительный поток несет эффективное представление пространственно-временного компонента ретинального изображения, наиболее полно отвечающее зрительным задачам, решаемым той или иной областью коры, принимающей сигнал от этого потока. Следовательно, исследовав т.н. *реактивную чувствительность* нейронов внутри зрительного потока к стимулам различных видов, мы можем многое узнать о функциональной роли этого потока.

#### 4.1.1.3 Контраст ретинального изображения и адаптация

Существует несколько зрительных аспектов общих всем специализированным зрительным потокам, и есть смысл разобраться с ними на этапе исходного кодирования зрительного сигнала в сетчатке, то есть *до того*, как сигналы отправятся в разные целевые отделы мозга. Одним из аспектов является то, что ретинальные нейроны должны оставаться чувствительными к паттерну ретинального изображения в широком диапазоне интенсивностей внешнего освещения — от тусклого вечернего до яркого дневного, то есть в диапазоне шести порядков.

Нейроны периферической зрительной системы (т.е. глаза) частично решают эту проблему, реагируя на *локальный контраст* в ретинальном изображении, но не на абсолютную величину стимуляции.

⇒ *Локальный контраст* — это процентная доля отклонения фотометрической яркости того или иного участка изображения относительно средней фотометрической яркости этого изображения.

Диапазон контрастов в типичном ретинальном изображении остается более или менее постоянным и редко когда выходит за границы двух порядков. Кодируя контраст, а не абсолютную величину стимуляции, нейроны с весьма скромным выходным динамическим диапазоном (т.е. диапазоном выходного сигнала) могут передавать очень точную информацию о ретинальном изображении, невзирая на огромные перепады интенсивности внешнего освещения. В дальнейшем мы дадим количественное доказательство сего тезиса и покажем, что сигнал контраста — это фундаментально важный сигнал. Сигнал контраста тесно связан со свойствами поверхностей и вспомним, что именно поверхности — это те зрительные сущности, которые мы хотим идентифицировать или распознать.<sup>1</sup>

#### 4.1.2 Зрительные потоки и рецептивные поля

Разговор о зрительных путях и зрительных потоках мы начнем с принципов классификации ретинальных нейронов, а затем приступим к разбору характера информации, проходящей по двум главным зрительным путям: *парвоклеточному* и *магнокле-*

<sup>1</sup> В некоторых источниках вопрос динамического диапазона толкуется через то, что отклик фоторецепторов идет по компрессивной функции, в частности логарифмической или степенной. Компрессивная функция сжимает шесть порядков световых интенсивностей в диапазон отклика нейронов от одного до двух порядков от их естественной вариабельности. В современной литературе сия точка зрения частично заменена рассуждениями о контрасте стимула и контрасте нейронного отклика.

*точному.* Мы уделим этим путям особое внимание по двум причинам: во-первых, потому что они лучше всего изучены; во-вторых, потому что на выходе они представляют большую часть общего выходного сигнала сетчатки.

#### 4.1.2.1 Классификация нейронов

Форма и структура нейрона, включая крону его дендритов, клеточное тело и аксон — именуются *морфологией* этого нейрона. Главный критерий категорийной классификации нейронов — это их морфология.

Второй критерий — это нейронная электрическая реакция на разного рода сигналы, то есть — *электрофизиология* нейрона.

Третий критерий — это химический состав веществ, наполняющих нервную клетку, то есть — *биохимия* нейрона.

Четвертый — это *паттерн связей* нейрона с другими нейронами.

Понятно, что наиболее полная категоризация лежит на стыке перечисленных подходов.

Вспомним, что мы использовали все перечисленные методы при делении фоторецепторов на палочки и колбочки: палочки и колбочки можно классифицировать, основываясь на их морфологии (палочкообразная форма клетки, колбочкообразная форма), характере электрического отклика на световое воздействие, типе фотопигмента и характере связей с другими нейронами (у палочек нет связей с нейронами центральной ямки). Все вместе указывает на то, что палочки и колбочки различаются функционально: колбочки несут визуальную информацию высокого пространственного разрешения, а палочки — информацию о стимулах очень низкой интенсивности (скотопические условия) и низкого пространственной частоты.

Выработанный принцип мы будем применять к дальнейшему анализу функции нейронов, и разбор функции ганглиозных клеток сетчатки мы начнем с того, что рассмотрим, каким образом мы можем использовать описанные выше методы при категоризации этих клеток на функциональные группы.

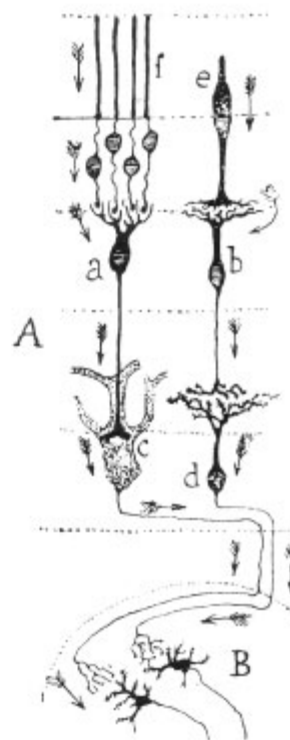
#### 4.1.2.2 Морфология макро- и парвоклеток

Если поглядеть на слой ретинальных ганглионаров с помощью обычного оптического микроскопа, то можно увидеть множество клеток различных размеров и форм и с разнообразными кронами дендритов. В своих выдающихся исследованиях упомянутый нами выше Сантьяго Рамон-и-Кахаль (рис. 4.1.3) исследовал виды ретинальных клеток в сетчатке многих млекопитающих (но не приматов) и выявил ее основные гистологические структуры. Для классификации нейронов Кахаль использовал различные морфологические свойства, в основном локализацию окончаний дендритовой кроны.

Новая эра в гистологических исследованиях сетчатки приматов началась с работы профессора Чикагского университета Стивена Поляк (1941, 1957), написавшего две монументальных монографии об исследовании зрительной системы приматов. Поляк описал множество аспектов ее анатомии и гистологии, в частности гистологической структуры сетчатки.

Автор выделяет пять различных категорий ганглионаров, исходя из размера их клеточных тел и свойств их дендритовых крон. Один из принципиальных и в данный момент нам интересных критериев его классификации — это размер и распространенность дерева дендритов ганглиозных клеток: в большинстве участков сетчатки мы мо-





**Рис. 4.1.3** Сантьяго Рамон-и-Кахаль в своей лаборатории. Справа — рисунок, иллюстрирующий направление зрительных сигналов в сетчатке млекопитающих, выполненный Кахалем. В части «А» рисунка, клетки, отмеченные как «а» — палочковые биполяры; b — колбочковые биполяры; «с» и «d» — ганглиозные клетки; «е» и «f» — колбочковые и палочковые фоторецепторы соответственно. Связи с подкорковым центром показаны в части «В».

По материалам Rodieck, 1973.

жем идентифицировать нейроны по относительным плотности и компактности дендритовых крон в сравнении с таковыми у прочих ганглиозных клеток.

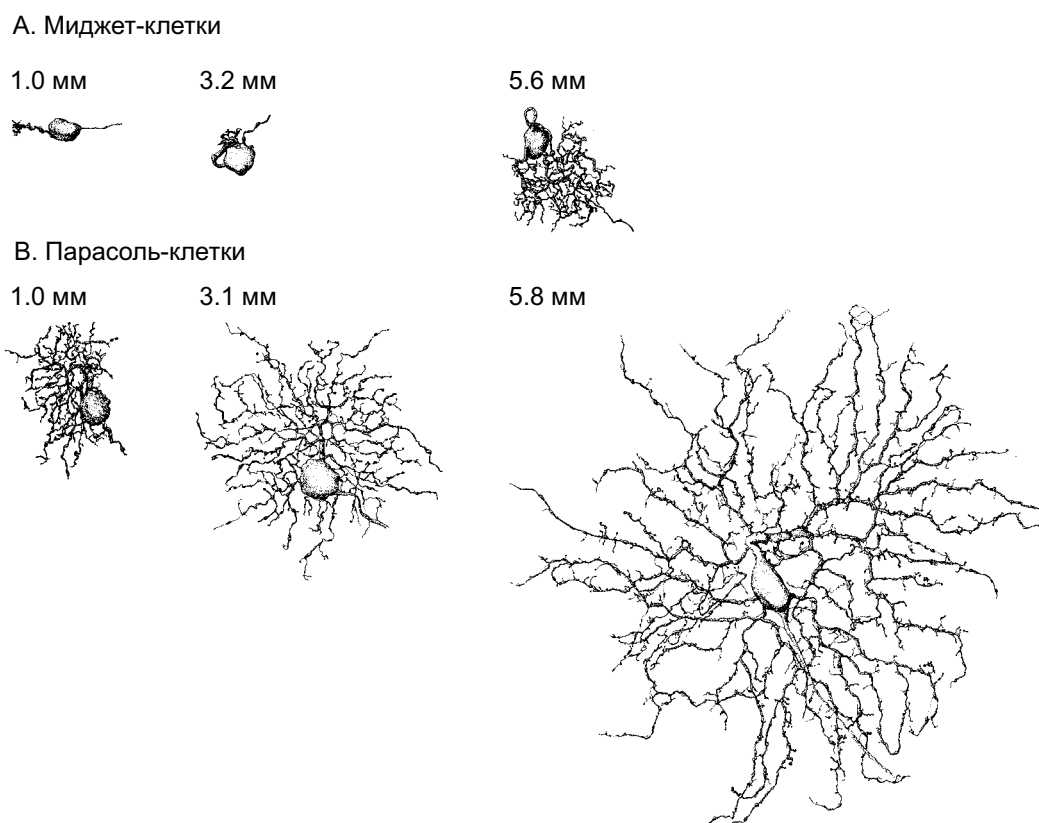
Вблизи центральной ямки эти нейроны идентифицируются благодаря тому, что образуют связи только с единичной биполярной клеткой, которая, в свою очередь, образует связь с единичной колбочкой. Поляк первым выявил этот тип ганглиозных клеток (коих в изобилии у приматов, но которых нет у других млекопитающих), и назвал их *миджет-клетками* (карликовые ганглионары) — 4.1.4 А.

Морфология миджет-ганглиозных клеток контрастирует с вторым классом ганглиозных клеток, показанных на рис. 4.1.4 В и названных Поляк *парасоль-клетками* (*зонтичные ганглионары*).

#### 4.1.2.3 Ретинальная эксцентрика

Как отмечал Поляк, размер ретинальных нейронов многих типов растет по мере удаления от центральной ямки, к примеру: клеточные тела и поля дендритов, которые относительно компактны в центральной ямке — относительно велики на периферии сетчатки. Вместе с тем, оказалось, что из-за ретинальной неоднородности миджет-клетки на периферии сетчатки оказываются крупнее парасоль-клеток возле центральной ямки.





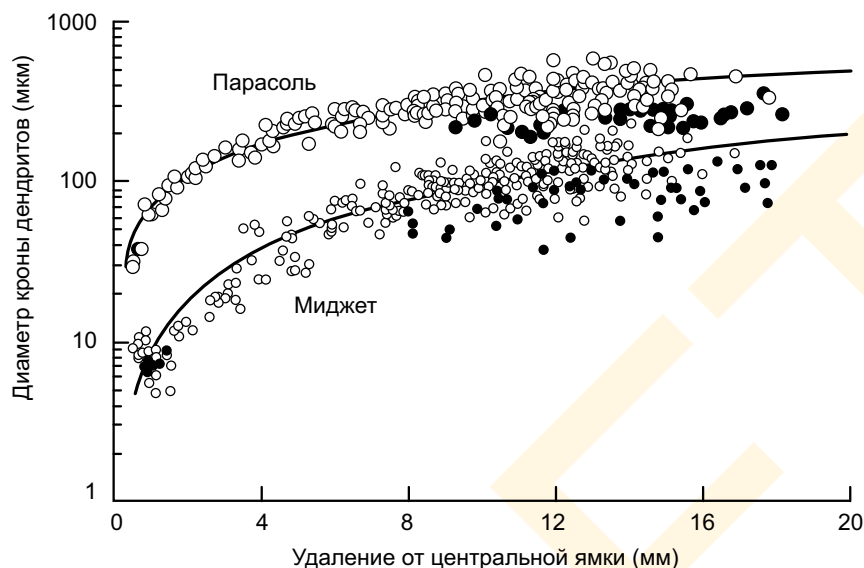
**Рис. 4.1.4** Сравнительная морфология миджет- и парасольганглиозных клеток на разных удалениях от центральной ямки. С помощью камеры обскура были получены изображения миджет(А) и парасоль-клеток (В) на разных участках сетчатки. В сравниваемых позициях кроны дендритов миджет-клеток меньше и плотнее, чем у клеток парасоль. У клеток обоих типов абсолютный размер кроны дендритов растет по мере удаления от центральной ямки.

По материалам Watanabe & Rodieck, 1989.

Однако ежели абсолютный размер не является достоверным индикатором, то как мы можем знать, что клетки, по-разному удаленные от центральной ямки, то есть, обладающие, как говорят, разной мерой *зрительной эксцентрики*, относятся к одному и тому же типу?

В 1974 г. в своей эпохальной статье «Морфологические типы ганглиозных клеток сетчатки домашней кошки» Брайан Бойкотт (Великобритания) и Хайнц Вессле (Германия) показали, как выполнить такую идентификацию: идея проста и изящна — выполняются измерения, охватывающие широкий диапазон ретиальной эксцентрики и сравниваются отклонения внутри популяции нейронов определенного клеточного типа. Отметим, что в дальнейшем методы Бойкотта и Вессле были применены к приматам, в частности к человеку (Levintal et al., 1981; Pery and Cowey, 1981, 1984; Rodieck et al., 1985; Watanabe & Rodieck, 1989; Dacey & Petersen, 1992).

На рис. 4.1.5 показано, что крона дендритов у миджет- и парасоль-клеток растет в размере по мере удаления от центральной ямки. Несмотря на то, что клетки обоих ти-



**Рис. 4.1.5** Диаметр кроны дендритов как функция от ретиальной эксцентрики в глазу человека. На графике показаны размеры дендритовой кроны миджет- и парасоль-нейронов. Сплошные кружки — удаление от центральной ямки в назальном направлении; полые — в темпоральном. У обоих типов нейронов размер кроны дендритов растет по мере удаления от центральной ямки, но в любом положении клетки легко идентифицируются по *относительному* диаметру их крон.

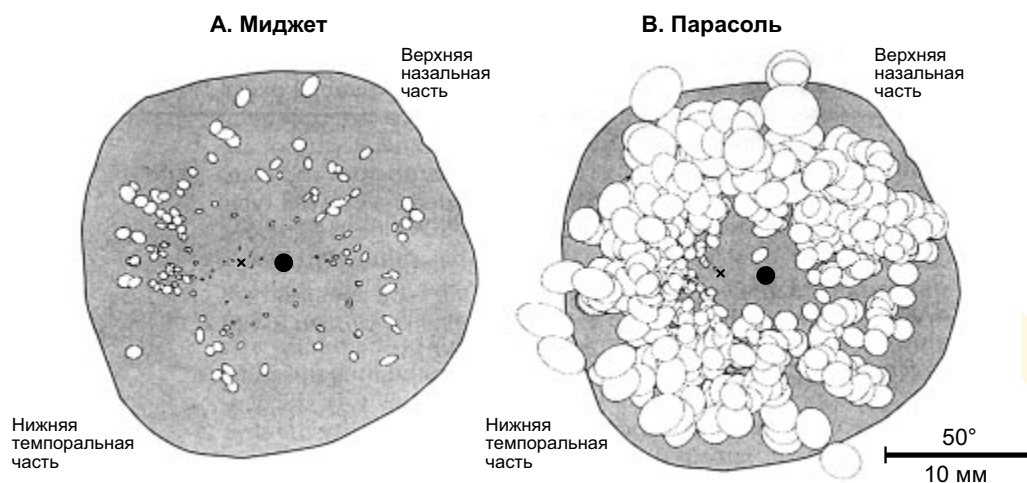
По материалам Dacey & Petersen, 1992.

пов растут в размере внутри каждого клеточного типа, относительный диаметр дендритовых крон плавно меняется на каждом шаге ретиальной эксцентрики.

Рис. 4.1.5 демонстрирует нам также и то, что сигналы от миджет-клеток образуют единый зрительный поток, а от клеток парасоль — другой. Дальнейшее доказательство того, что данные два класса нейронов образуют независимые друг от друга зрительные потоки, находим в покрытии ими ретиального изображения. Каждая точка этого изображения кодирована как минимум одной миджет- и одной парасоль-ганглиозной клеткой. То есть, нейроны обоих классов кодируют полную копию ретиального оптического изображения, сигнал о котором пришел от фоторецепторов (рис. 4.1.6).

Однако же, несмотря на то, что обе популяции клеток самостоятельно и целиком кодируют ретиальное изображение, выполняют они это с разным пространственным разрешением: в центральной ямке миджет-клетка получает сигнал от единичной колбочки (при этом особо отметим, что сказанное вовсе не значит, что колбочка посылает сигнал в единичную биполярную или ганглиозную клетки!); по распространенности дендритовой кроны клеток парасоль мы видим, что они получают входной сигнал от намного более крупных областей сетчатки.

Высокое пространственное разрешение, достигаемое миджет-клетками, означает, что для кодирования ретиального изображения этими клетками требуется их большее количество. Действительно, миджет-клеток примерно в 7-9 раз больше, чем клеток парасоль (Perry et al., 1984; Sterling et al., 1994). Миджет-клетки кодируют ретиальное изображение с полным пространственным разрешением фоторецепторов, что составляет приблизительно 60 циклов на градус. Численно меньшая популяция ганг-



**Рис. 4.1.6** Кроны дендритов миджет- и парасоль-клеток, представленные на данной иллюстрации белыми овалами, несут на себе полные копии ретинального изображения. Кроны дендритов миджет-клеток (А) ощутимо меньше таковых у клеток парасоль (В). Кроны дендритов у парасоль-клеток ощутимо больше в каждой локали сетчатки, поэтому их нужно меньше, чтобы покрыть всю сетчатку. Черная точка — оптический диск; крестик — центральная ямка.

По материалам Watanabe & Rodieck, 1989.

лиозных клеток парасоль способна закодировать ретинальное изображение с пространственным разрешением лишь в 20 цикл/град.

Любопытно и то, что парасоль-клетки чувствительнее к хроматическому воздействию, нежели к яркостному контрасту, тогда как миджет-клетки напротив — чувствительнее к яркостным отличиям и менее чувствительны к хроматике. Доказательством отличия в отклике на хроматическое воздействие у этих клеток является отличие в паттернах связей с фоторецепторами: парасоль-клетки получают входной сигнал только от колбочек, тогда как миджет-клетки — как от палочек, так и от колбочек. Чуть позднее мы увидим, как функциональные отличия между миджет- и парасоль-проводящими путями распространяются и на корковые отделы зрительной системы за счет магно- и парвослоев коленчатого тела.

#### 4.1.2.4 Центральные проекции

Второй метод идентификации зрительных потоков, исходящих из слоев ганглиозных клеток, состоит в исследовании того, как ретинальные ганглиозные клетки разных типов соединяются с нейронами мозга. Путем введения т.н. *трэйсеров* — *трассирующих веществ*, текущих из мозга назад сетчатке — мы можем отследить то, куда ганглиозная клетка посылает свой выходной сигнал.

Нейробиолог Хью Перри (Оксфорд), занимавшийся этим вопросом, в качестве трэйсера использовал фермент пероксидазу, выделенную из хрена обыкновенного. Пероксидаза, поглощенная нейроном, всегда в движении: если вещество было поглощено аксоном, то через некоторое время оно оказывается в клеточном теле, и наоборот — если фермент поглощен клеточным телом, то через некоторое время он достигает терминалей аксона. Присутствие пероксидазы в клетках выявляется специальными гистохимическими методами.

Вначале Перри вводил фермент в оптический нерв неподалеку от сетчатки, и таким манером ему удалось идентифицировать все типы ганглиозных клеток и то, как они выглядят, когда прокрасятся трэйсером. Затем он ввел пероксидазу в латеральное коленчатое тело таламуса, которое, напомним, принимает на себя большую часть аксонов сетчатки, благодаря чему пероксидаза пошла по многим ганглиозным клеткам. В итоге, Перри установил, что 90% ганглиозных клеток сетчатки макаки посылают свои аксоны в различные слои ЛКТ. Помимо того, что степень насыщения пероксидазой ганглиозных клеток позволяет классифицировать их как клетки миджет или парасоль, исследователь убедительно показал, что как минимум еще два других типа ретинальных ганглионаров посылают сигнал в ЛКТ (Rodieck & Watanabe, 1993).

**Слои ЛКТ.** Клеточная архитектура латерального коленчатого тела очень интересна и дает исследователям ключ к пониманию функций главных компонентов зрительной системы. Как и многие нейронные образования головного мозга латеральное коленчатое тело *ламिनарно* (слоисто), т.е. построено из множества нейронных листов. Оба коленчатых тела головного мозга состоят из шести самостоятельных слоев клеток (рис. 4.1.7): нижние два слоя именуются *магноклеточными* (от лат. «magnus» — большой, крупный), поскольку образованы крупными клетками; верхние четыре слоя называют *парвоклеточными* (от лат. «parvus» — мелкий), поскольку образованы эти слои мелкими нейронами.

Магно- и парвоклетки разнятся не только *морфологически* (т.е. по форме и размеру), но и функционально: магноклеточные клетки весьма чувствительны к яркостным отличиям внутри сложного стимула, но почти не реагируют на хроматическое воздействие; у магноклеток относительно большие рецептивные поля и кратковременный



**Рис. 4.1.7** Поперечный разрез латерального коленчатого тела макаки. Латеральное коленчатое тело (ЛКТ, LGN) расположено в области мозга, именуемой *таламусом*. У приматов аксоны ганглиозных клеток по большей части заканчиваются в ядрах коленчатого тела. При окраске по методу Гольджи ядра коленчатого тела выглядят состоящими из шести слоев: у нейронов четырех верхних слоев клеточные тела мелкие, из-за чего слои эти получили название *парвоклеточных*; клеточные тела нейронов двух нижних слоев много крупнее, благодаря чему эти слои получили название *магноклеточных*. Также в коленчатом теле находятся нейроны, тела которых расположены между означенными слоями, т.е. в областях, именуемых *вставками*. Каждый слой несет в себе ретинотопическую карту поля зрения. Аксоны нейронов ЛКТ идут в V1 зрительной коры, образуя коленчато-затылочный путь (см. рис. 1.3.1)

По материалам Hubel & Wiesel, 1977.



отклик на быстрые изменения в ретиальной стимуляции; парвоклеточные клетки относительно нечувствительны к яркостным изменениям, высокочувствительны к хроматическому воздействию, имеют относительно небольшие рецептивные поля (обычно в половину от размера таковых у магноклеток) и дают весьма продолжительный отклик на изменения в ретиальной стимуляции.

Функциональные отличия между магно- и парвоклетками коленчатого тела сведены в таблицу 4.1.1:

ТАБЛИЦА 4.1.1 Функциональная дифференцировка нейронов магно- и парвослоев ЛКТ

	Парвонероны	Магнонероны
Хроматическая чувствительность	Высокая	Низкая
Чувствительность к яркостному контрасту	Низкая	Высокая
Чувствительность к пространственной частоте	Высокая	Низкая
Временное разрешение	Медленно	Быстро
Размер рецептивных полей	Малый	Большой

Стоит отметить, что большинство этих отличий нечетки и статистика показывает, что имеют место выраженные функциональные наложения. Тем не менее, функциональное деление привело некоторых исследователей к мысли, что магноклетки образуют специализированный проводящий путь, отвечающий за обработку информации о движении объекта и его объеме, тогда как парвоклетки — отдельный путь обработки хроматической информации и информации о форме объектов (DeYoe & Van Essen, 1988; Livingstone & Hubel, 1988).

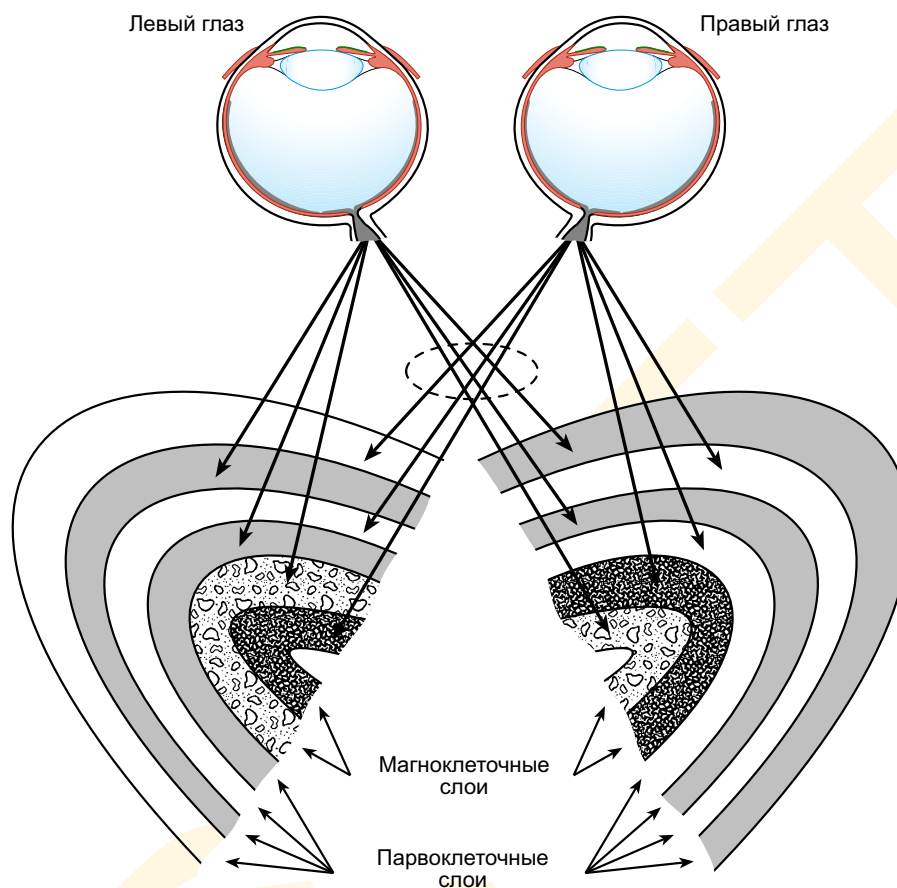
Несмотря на то, что полный сигнал, поступающий в коленчатое тело — это сигнал от сетчатки обоих глаз, каждый слой в отдельности получает вход только от одного глаза: в верхний парвоклеточный слой от контралатерального глаза (т.е. глаза противоположной стороны), во второй парвослой — от глаза своей стороны, в третий слой — с противоположной и в четвертый (нижний парвоклеточный) — со своей; у магнослоев наоборот — в верхний слой поступает сигнал от глаза своей стороны, в нижний — противоположной (рис. 4.1.8).

В коленчатом теле головного мозга приматов в распределении аксонов, идущих от миджет- и парасоль-клеток существует определенная закономерность. Аксоны парасоль- и миджет-клеток сетчатки образуют связи в разных слоях ЛКТ: у миджет-клеток аксоны заканчиваются в парвоклеточных слоях, а аксоны клеток парасоль — в магноклеточных слоях коленчатого тела.

Наряду с описанными, напомним, имеются клетки, залегающие во вставках (Fitzpatrick et al., 1983). Отметим, что вставки могут получать сигналы от ганглиозных клеток других классов.

**Парво- и магноклеточные проводящие пути.** Форма ретиальных ганглиозных клеток и проекции в центральную нервную систему (центральные проекции) указывают на то, что миджет- и парасоль-клетки формируют отдельные зрительные потоки. Проводящий путь, который берет начало от миджет-клеток и заканчивается в парвоклеточных слоях ЛКТ — именуется *парвоклеточным проводящим путем*; путь, идущий





**Рис. 4.1.8** Схема маппинга ретиальной информации на слои ЛКТ. Верхние четыре слоя — парвоцеллюлярные; нижние два — магноцеллюлярные. Слои ЛКТ, получающие сигнал от сетчатки левого глаза покрашены серым, а от правого — белым. Верхние слои получают сигнал от контралатерального (оппозитного) глаза и по ходу к низу идет перемежение с парвоцеллюлярными слоями, принимающими сигнал от сетчатки своего глаза. В магноцеллюлярных слоях верхний принимает сигнал от гомолатерального (своего) глаза, а нижний — от контралатерального.

Пунктирная линия — хиазма.

от клеток парасоль и оканчивающийся магноклеточными слоями — именуется *магноклеточным проводящим путем*.

Роль проводящих путей в зрительном акте — предмет активных исследований и гипотез. Какова мера сегрегации (разделения) сигналов по зрительным путям? Несут ли эти сигналы специализированную информацию? В следующих разделах мы познакомимся с отличиями в откликах нейронов этих путей на световую стимуляцию фоторецепторов, а затем еще раз коснемся этого вопроса.

Итак, несмотря на то, что большая часть ретиальных ганглиозных клеток посылает выходной сигнал в латеральное коленичатое тело таламуса, у волокон оптического тракта имеется множество иных целевых участков в мозге. К примеру, в 1980 г. Х. Перри и Ф. Кови (1980) провели эксперименты с введением пероксидазы в верхние бугры четверохолмия (отдел среднего мозга), о которых известно, что они получают сигнал от ганглиозных клеток сетчатки. Было выявлено, что примерно 10% ганглиоз-

ных клеток посылают аксоны именно в эту область мозга, но ни одна из них не относится к классу миджет- или парасоль-клеток.

В 1993 г. упомянутый выше американец Роберт Родик и японский нейрофизиолог Масами Ватанабе провели совместную масштабную серию измерений, касающую работы зрительных потоков. Исследователи пришли к выводу, что ганглиозные клетки каждой из подкатегорий посылают сигнал только в одно целевое место мозга, благодаря чему морфология ганглиозных клеток становится наиважнейшим признаком, определяющим организацию зрительных потоков, исходящих из сетчатки.

Однако же вновь повторим: бо́льшая часть ретинальных сигналов направлена в коленчатое тело, но при этом ретинальные синапсы составляют лишь 10% от общего числа синапсов в ЛКТ — порядка 60% синапсов коленчатого тела отвечают за сигналы, идущие из коры, а оставшиеся 30% отведены связям с другими отделами мозга (Sherman & Koch, 1990).

Аксоны ганглиозных клеток, выходящие из глаза, формируют оптический нерв. В определенном месте нерва они образуют т.н. *хиазму* — *зрительный перекрест*, в котором часть волокон перебрасывается на противоположную сторону — и заканчиваются синапсами на клетках коленчатого тела таламуса (см. рис. 1.3.30). Рецептивные поля клеток коленчатого тела также организованы по принципу центр-окружение и очень схожи с рецептивными полями ганглиозных клеток, но при этом несколько крупнее и с более жестким торможением в окружении.

Наибольшее отличие между последним ретинальным слоем и коленчатым телом состоит в клеточной архитектуре: ганглиозные клетки образуют двухмерный тонкий листок параллельный рецепторной поверхности сетчатки и принимают сигнал только от соседних клеток своего глаза; клетки коленчатого тела образуют трехмерные структуры и получают сигнал от обоих глаз. Однако каждая отдельно взятая клетка коленчатого тела *монокулярна*, т.е. дает отклик на стимулирующий ее сигнал только от одного глаза. Т.н. *бинокулярные клетки*, т.е. клетки, принимающие сигнал от обоих глаз, не встречаются вплоть до зрительной коры.

**Ретинотопический мэппинг.** Пространственно каждый слой ЛКТ организован также как и сетчатка того глаза, от которого пришел сигнал, то есть, мы говорим, что имеет место т.н. *ретинотопический мэппинг* (*топографический мэппинг*), т.е. сохраняющий относительное взаиморасположение нейронов от сетчатки до ЛКТ: соседние участки сетчатки проектируются на соседние участки коленчатого тела, сохраняя тем самым количественные (но не качественные) пространственные отношения на т.н. *ретинотопической карте*. Сей вид пространственного мэппинга — это общее свойство высших отделов зрительной системы, обнаруживаемое во многих участках зрительной коры (что мы увидим далее). Ретинотопические карты во многих слоях коленчатого тела расположены так, что соответствуют участкам сетчатки, находящимся друг над другом, но при этом, судя по всему, никаких межслойных взаимодействий не возникает.

Наконец отметим, что каждая карта представляет лишь половину сетчатки на ее *ипсилатеральной* (*гомолатеральной*, своей) стороне от центральной ямки, что соответствует *контралатеральной* (противоположной) стороне поля зрения. Так получается в результате наличия хиазмы — зрительного перекреста (рис. 1.3.30). Далее неизменность проекции своего полуполя обеспечивают аксоны нейронов коленчатого тела, идущие в стриарную кору по т.н. *коленчато-затылочному пути* (*пучок Грасиоле*).

#### 4.1.2.5 Проводимость возбуждения и усиление контраста

Существуют некоторые отличия в способах кодирования визуальной информации нейронами магно- и парвоклеточных путей. Отличия эти — ключ к ответу минимум на два вопроса: «Информация каких видов представлена в этих зрительных потоках?» и «Какую роль играют эти потоки в зрительном акте?».

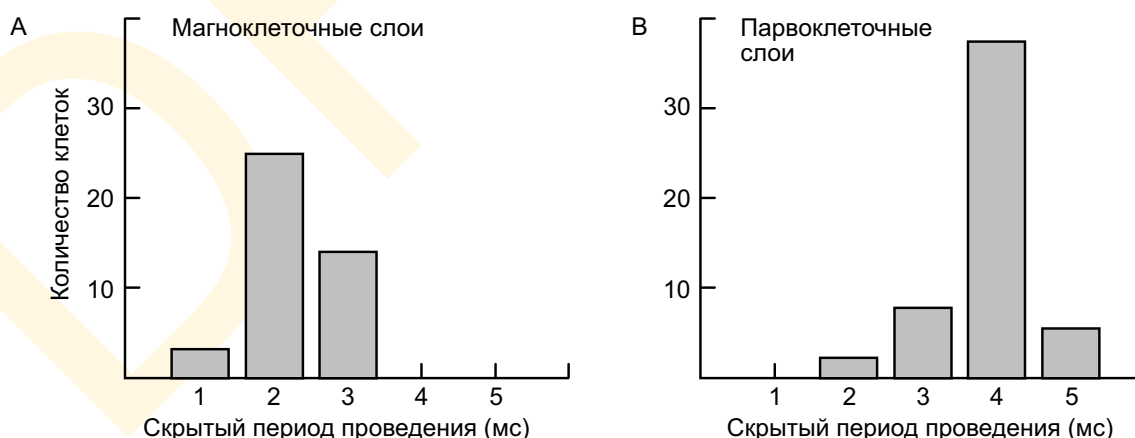
Время проведения электрических сигналов от сетчатки в парвоклеточные слои коллатерального тела больше, чем таковое в слоях магноклеточных. В хиазме аксоны оптических нервов обеих сетчаток объединяются, а затем реорганизуются на две отдельных группы, кодирующих информацию о правом и левом полях зрения (эти две группы образуют связи с разными полушариями мозга — см. рис....).

В 1978 г. американцы Петер Шиллер и Джозеф Мальпели (Массачусетский технологический институт, 1978) измерили проводимость сигнала от электрической стимуляции хиазмы в два разных слоя коллатерального тела: ученые обнаружили, что в парвоклеточный слой сигнал приходит позднее, нежели в магноклеточный — см. рис 4.1.9.

Во-вторых, отклик нейронов в обоих путях на контрастные паттерны стимуляции отличается весьма устойчиво.

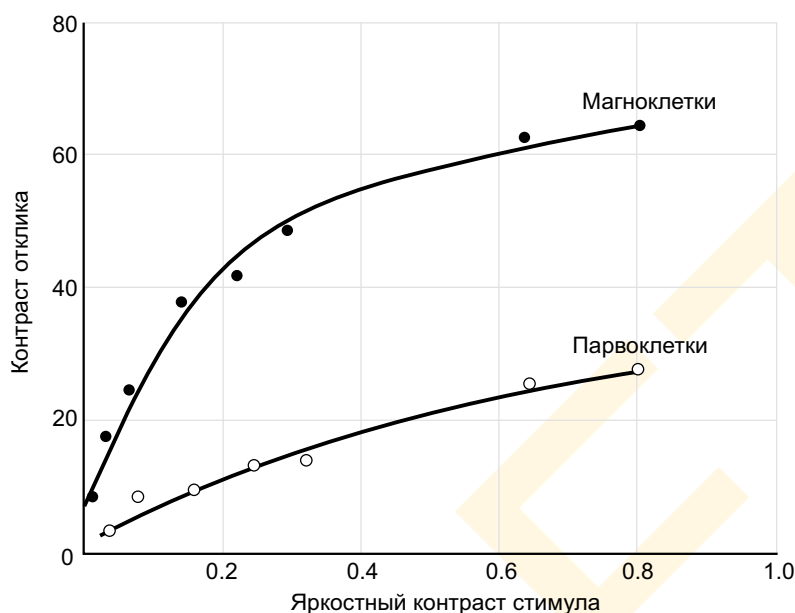
В 1982 г. американцы Иегуд Каплан и Роберт Шепли (университет Рокфеллера, 1982, 1986) наблюдали как при росте контраста синусоидальной решетки (см. раздел 4.2.1) отклик нейронов в магноклеточном пути менялся быстрее, чем в парвоклеточном. На рис. 4.1.10 дано сравнение кривых отклика на контрастный стимул нейронов парвоклеточного и магноклеточного проводящих путей: по оси абсцисс отложен фотометрический контраст стимула, а по оси ординат — отклик нейронов как процентная доля от уровня спонтанного отклика. Мы видим, что кривая отклика в магноклеточном пути растет быстрее при росте контраста стимула и также быстрее насыщается на низком контрасте. Наклон кривой отклика на контрастный стимул именуется *нейронным усилением контраста*. Глядя на рис. 4.1.10, сказав, что магноклеточные нейроны ЛКТ усиливают контраст примерно в 8 раз больше парвоклеточных.

Отметим наконец, что хотя исходно Каплан и Шепли наблюдали разницу в усилении контраста нейронами ЛКТ, ученые показали, что сие отличие можно отследить по



**Рис. 4.1.9** Время проведения возбуждения от хиазмы (при ее электрической стимуляции) к магноклеточным и парвоклеточным слоям латерального коллатерального тела таламуса. Отклики нейронов в магноклеточных слоях (A) возникают быстрее, чем в парвоклеточных (B).

По материалам Schiller & Malpeli, 1978.



**Рис. 4.1.10** Функции отклика нейронов ЛКТ на контрастную стимуляцию. Контраст отклика магноклеток (сплошные кружки) по мере роста контраста стимуляции растет быстрее, чем контраст отклика парвоклеток (открытые кружки)

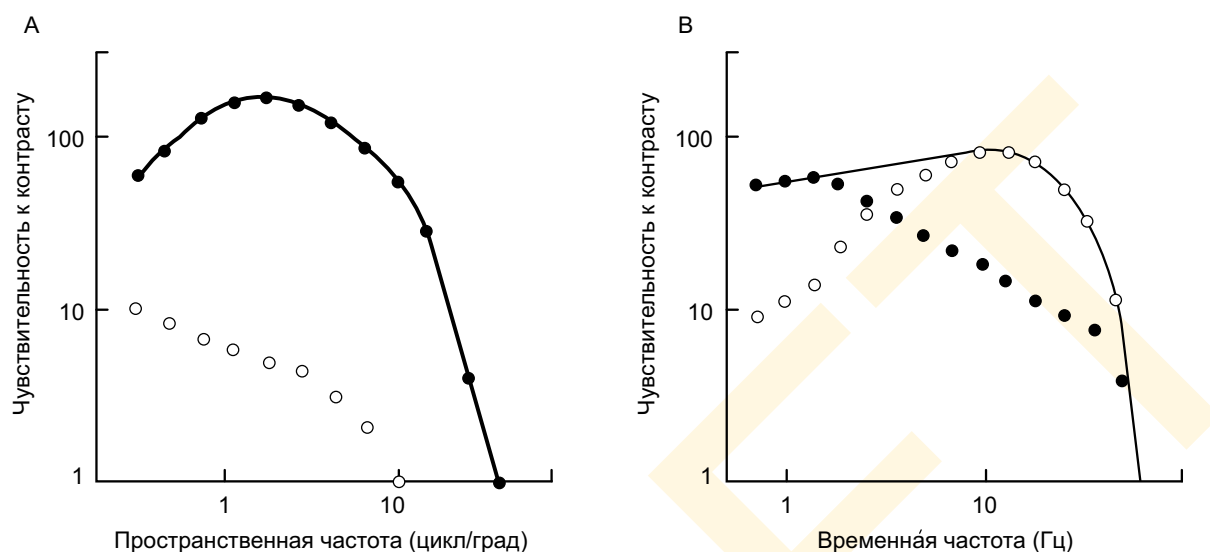
По материалам Shapley, 1990.

отличиям в величинах откликов миджет- и парасоль-нейронов в сетчатке приматов. То есть, вполне возможно, что отличия в сигналах начинаются в биполярных связях колбочек друг с другом.

#### 4.1.2.6 Зрительная информация, кодируемая в парво- и магноклеточном путях

Гистологические и физиологические исследования свидетельствуют о том, что парвоклеточные и магноклеточные проводящие пути несут в мозг информацию различных видов. Мы попытаемся в этом убедиться с помощью блокады проводящих путей у животных и наблюдения за дальнейшим состоянием последних.

Выполняя блокаду, мы должны быть очень осторожны, чтобы не повредить те волокна, что идут к, либо от интересующих нас участков. Весьма популярным веществом, позволяющим поставить данный опыт, является иботеновая кислота, разрушающая клеточное тело нейронов, но интактная к их аксонам (что и используется для воздействия на магно- и парвослои ЛКТ). Наблюдая изменения в состоянии животных после иботеновой блокады парво- и магноклеточных путей, мы сможем сделать кое-какие выводы о той информации, что представлена этими путями (Schiller & Logothetis, 1990; Merigan et al., 1991a, b; Lynch et al., 1992), в частности: при разрушении парвоклеточных слоев колленчатого тела макаки резко страдают функции различения предметов по их окраске и функция паттерн-детекции. Поскольку на парвоклеточный путь приходится порядка 70% ганглиозных клеток сетчатки, сей результат, пожалуй, не столь уж и неожиданный; при разрушении нейронов магноклеточных слоев ЛКТ многие зрительные функции остаются неповрежденными. Результаты некоторых психофизических тестов до и после блокады подытожены рис. 4.1.11 (Merigan et al., 1991 a): итог таков, что когда нейроны магноклеточных слоев разруше-



**Рис. 4.1.11** Психофизические эффекты блокады парвовклеточных и магноклеточных слоев ЛКТ макаки. А — гладкая кривая — психофизический отклик здорового животного при детекции стационарной решетки; сплошные кружки — реакция макаки, у которой выполнена блокада магноклеточного проводящего пути; полые кружки — парвовклеточного. В — гладкая кривая психофизический отклик здорового животного при детекции пространственно-низкочастотной решетки, мигающей с разными частотами. Сплошные кружки — реакция обезьяны после блокады магноклеточного пути; полые кружки — парвовклеточного.

По материалам Merigan & Maunsell, 1993.

ны, животное оказывается менее чувствительным к мигающим решеткам низкой пространственной частоты. Потеря чувствительности свидетельствует о том, что магноклеточный путь несет в себе большую часть информации по данному аспекту ретинального изображения и указывает на то, что путь этот специализирован на повышение способности особи к решению задач, требующих информации высокой временной частоты.

Какие же виды психофизических реакций зависят от стимуляции низкой пространственной, но высокой временной частоты? Рассмотрим два примера зрительной задачи, требующей точной и скоростной передачи данных о быстро меняющихся сигналах ретинального изображения: распознавание движения объекта и слежение за ним.

По центральным проекциям магноклеточного проводящего пути мы можем судить о том, что магноклеточный проводящий путь играет одну из ведущих ролей в проведении высококачественной информации, необходимой для восприятия движения. Однако сигналы этого пути не являются абсолютно необходимыми для решения такого рода задач.

В 1991 г. профессор офтальмологии Рочестерского университета Вильям Мериган (1991 а), изучавший восприятие движения с помощью блокады магноклеточного пути обезьян, обнаружил, что они могут компенсировать наведенную экспериментатором функциональную недостаточность простым нейронным усилением контраста. Иными



словами, мы можем компенсировать нехватку информации, поставляемой по магноклеточному пути, повышая качество информации, идущей по пути парвоклеточному.

#### 4.1.2.7 Ганглиозные клетки

Вспомним, что ганглиозные клетки сетчатки — это исток нескольких проводящих путей, в которых информация о свете конвертируется в темпоральные (временные) пакеты дискретных электрических импульсов, названных *потенциалами действия* или *спайками*. Мы измеряем отклики этих клеток, записывая временную картину потенциалов действия, вызванных, в конечном счете, световой стимуляцией фоторецепторов. Можно уверенно сказать, что **свойства нейронной трансформации световых сигналов в паттерн потенциалов действия — это наипервейшее наше знание о функциональном назначении нейронов и, следовательно, зрительной роли проводящих путей.**

В электрофизиологическом аспекте участок преобразования оптической информации в электрический импульсный сигнал, ассоциированный с данным нейроном, именуется *рецептивным полем* этого нейрона.

Идея рецептивных полей, впервые предложенная американским физиологом Хельденом Киффером Хартлайном<sup>1</sup> (1940), — это краеугольный камень электрофизиологического описания функции зрительных нейронов. Однако концепция эта — общепсихологическая, и понятие «рецептивного поля» используется для описания нейронных свойств и в иных сенсорных и моторных модальностях (Mountcastle, 1957).

Классически рецептивное поле ретинального нейрона определено так:

⇒ Рецептивное поле ретинального нейрона — это пространственный участок сетчатки, световое воздействие на который вызывает отклик этого нейрона.

Рецептивное поле можно диагностировать подачей на сетчатку мелких световых пятен переменного размера и простых движущихся штрихов, ловя при этом границу начала отклика и его пропадания. Отклики множества нейронов в зрительном пути возникают в ответ на свет, действующий на очень маленькие фрагменты сетчатки и, следовательно, от очень небольших участков поля зрения. На первых парах приведенное определение рецептивного поля вполне подойдет для получения и реализации адекватного предварительного описания нейронных преобразований.

Несмотря на то, что мы говорим о рецептивном поле ганглиозной клетки, поле это фактически определено свойствами целого зрительного пути, начинающегося в сетчатке и включающего в себя преобразования, исполняемые самим нейроном. В некоторых случаях, когда имеет место обратный нисходящий сигнал на нейрон из центральных отделов мозга, рецептивное поле мы измеряем с учетом вклада различных отделов проводящего пути (полагая при этом, что никакого обратного сигнала на собственно ганглиозные клетки нет).

В данном разделе мы обсудим способ оценки такого важнейшего свойства рецептивных полей как их организация по принципу «центр/окружение».

Существует два места в зрительных путях, где мы можем более или менее четко зафиксировать спайковую активность ретинальных ганглионаров (рис. 4.1.12), а именно у тел ганглиозных клеток, расположенных на поверхности сетчатки (войдя микроэлектродом вплотную к роговице), либо же в оптическом нерве, состоящем, напо-

<sup>1</sup> В 1967 г. Х. К. Хартлайн получил нобелевскую премию по медицине за исследование нейрофизиологических аспектов зрения. — *Прим. пер.*

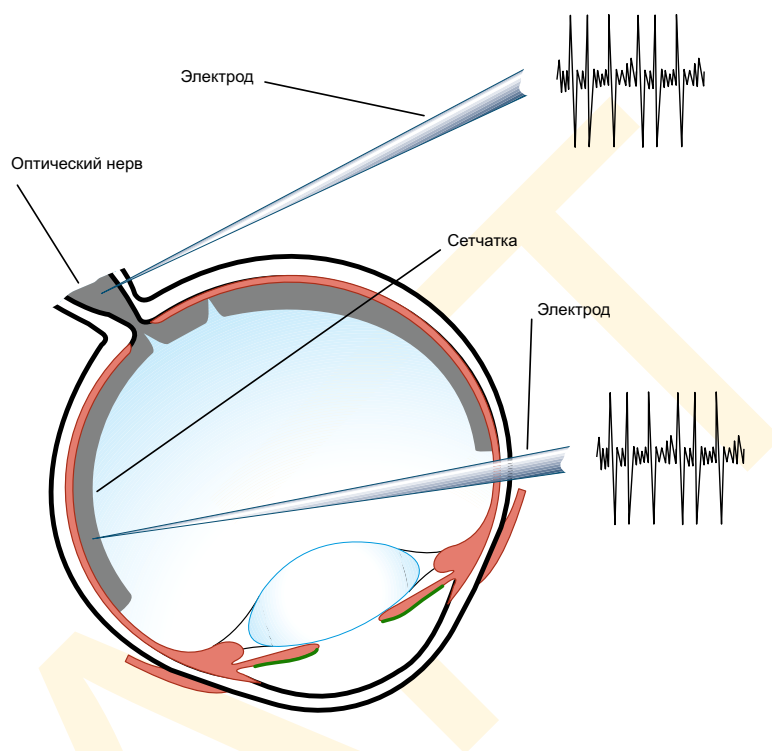
ним, из аксоновых волокон ганглиозных клеток, несущих сигнал в коленчатое тело таламуса.

В ответ на ганцфельд (см. раздел 3.2.1.8) — т.е. крупное яркостно-гомогенное поле — большинство ганглиозных клеток выдают поток хаотических, случайных потенциалов порядка 50 спайков в секунду. У большинства ганглиозных клеток временная последовательность спайковой активности несистематична, поэтому вероятность появления спайка в следующий короткий интервал времени примерно постоянна. Мы назовем среднее число потенциалов действия в единицу времени в ответ на ганцфельд *частотой спонтанной пульсации* ретиальной ганглиозной клетки.

Первыми нейронами сетчатки, чьи пространственные свойства поддались научному анализу, стали ганглиозные клетки. Примерно в одно и то же время венгро-американец Стефан Вильям Кюфлер (которого часто называют отцом современной нейронауки) и британец Хорэйс Базиль Барлоу начали записывать частоты пульсации отдельных ганглиозных клеток с помощью упомянутых уже микроэлектродов — сверхтонких стеклянных (или металлических) иглоочек диаметром порядка 0.001 см. Метод получил название *экстраклеточной регистрации активности единичных нейронов* или коротко — *экстраклеточной регистрации*.

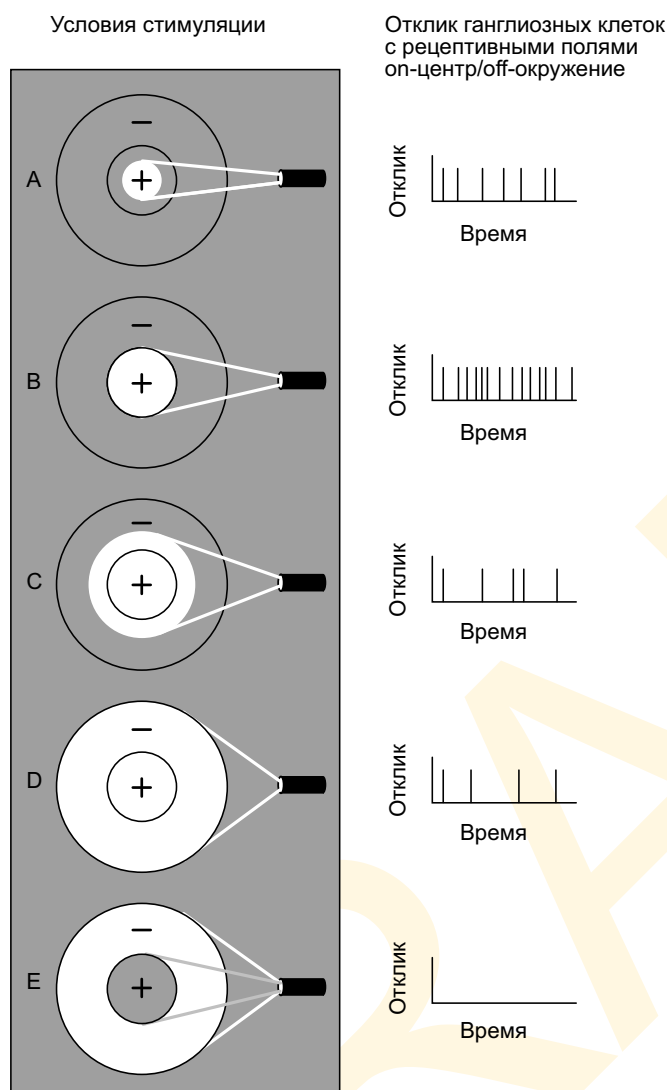
Запись сигнала проводилась возле нейрона, но при этом спайковый разряд фиксировался почти полностью. Чтобы понять, какие клетки отвечают, какие нет, в момент регистрации клеточной активности животным, разумеется, предъявлялись различные зрительные стимулы. Кюфлер исследовал ганглиозные клетки кошек, а Барлоу — кроликов, но оба экспериментатора обнаружили один и тот же факт — максимум пульсации клетки выдают на точечный стимул определенного размера, направленный в определенную позицию сетчатки. Если отклониться от этого размера в большую или меньшую стороны — отклик нейронов прогрессивно падает (см. рис. 4.1.13). Так были открыты рецептивные поля ганглиозных клеток. Результирующая картина откликов указала на определенного рода антагонизм между внутренним кругом этих полей и их *кольцом* (окружением).

В дальнейшем исследователи построили множество карт рецептивных полей ганг-



**Рис. 4.1.12** Потенциалы действия ретиальных ганглиозных клеток можно зафиксировать с помощью микроэлектрода, установленного возле клеточных тел с слое ганглиозных клеток сетчатки, либо возле оптического нерва.

По материалам Enroth-Cugell & Robson, 1984



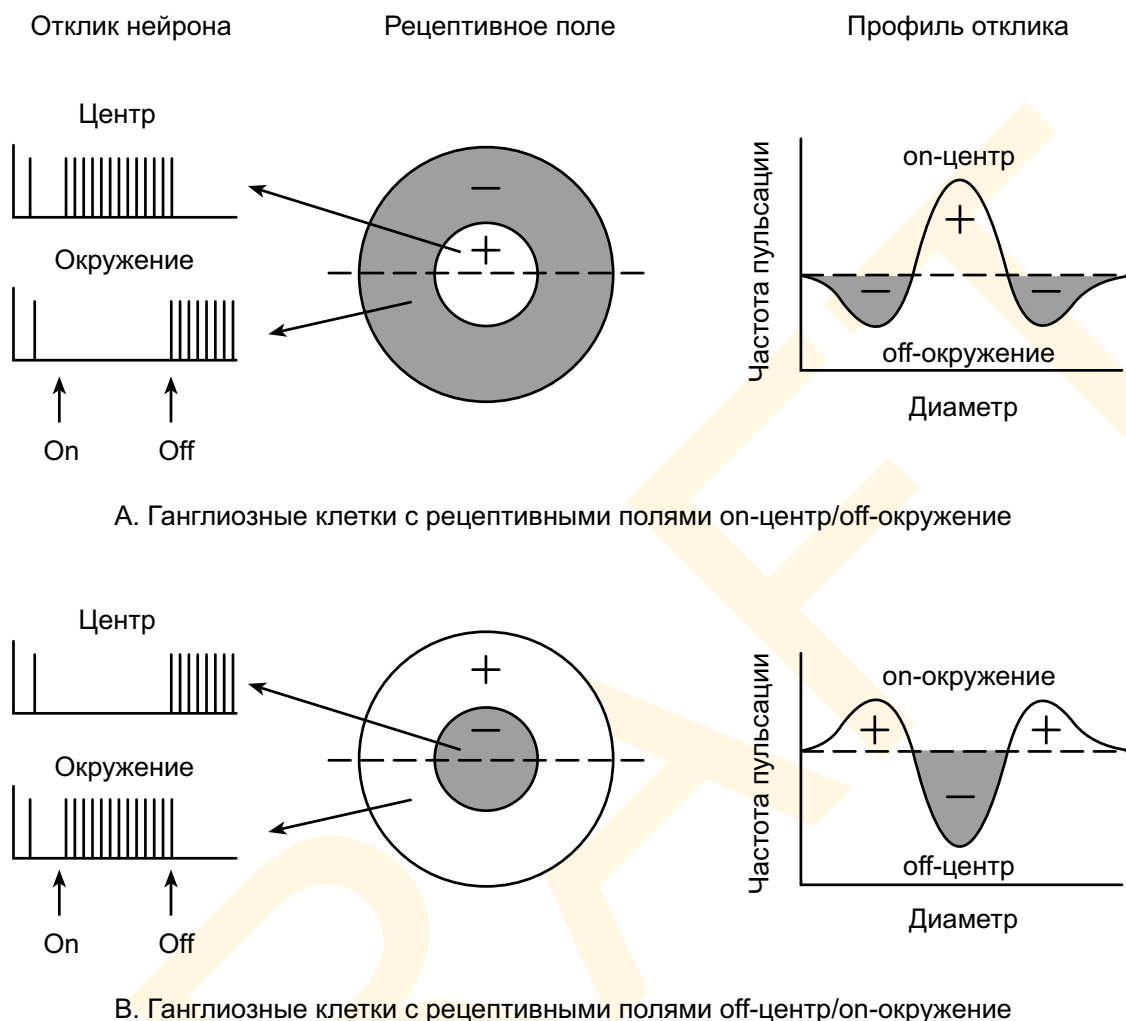
**Рис. 4.1.13** Отклики ганглиозных клеток с рецептивными полями организованными по принципу оп-центр/офф-окружение. По мере роста диаметра световой точки отклик вначале растет (А), достигая максимума, когда в точности покрывает возбуждающий центр рецептивного поля (В), а затем начинает падать, поскольку стимул выходит в тормозящее окружение (С и D). Минимум отклика (Е) наблюдаем тогда, когда стимул воздействует только на тормозящее окружение.

часто называют «сомбреро», поскольку его график напоминает срез мексиканской шляпы. Рецептивное поле off-центр-клетки (рис. 4.1.14 В) называют «сомбреро наизнанку».

Кюффлер (1953) первым диагностировал рецептивные поля ганглиозных клеток сетчатки млекопитающих, стимулируя ее в разных позициях мелкими световыми точками и записывая разность между частотой спонтанной пульсации и откликом на

лиозных клеток и выделили два типа этих клеток: *оп-центр-клетки*, возбуждающиеся при световой стимуляции центра рецептивного поля (рис. 4.1.14 А) и *офф-центр-клетки*, выдающие разряд при прекращении его стимуляции (рис. 4.1.14 В). Иными словами, отличие между двумя типами клеток в том, что одни отвечают возбуждением на световую стимуляцию центра их рецептивного поля (т.е. частота пульсации увеличивается) — *оп-центр-клетки*, а другие — торможением (т.е. частота пульсации падает) — *офф-центр-клетки*. При этом область, окружающая центр рецептивного поля, всегда имеет оппозитные характеристики (рис. 4.1.2): у *оп-центр-клеток* всегда *офф-окружение*, а у *офф-центр-клеток* всегда *окружение оп*. То есть, мы говорим об организации рецептивных полей по типу «центр-окружение», описанному нами в главе 3 в контексте латерального торможения и эффекта маховых полос (см. рис. 3.2.31).

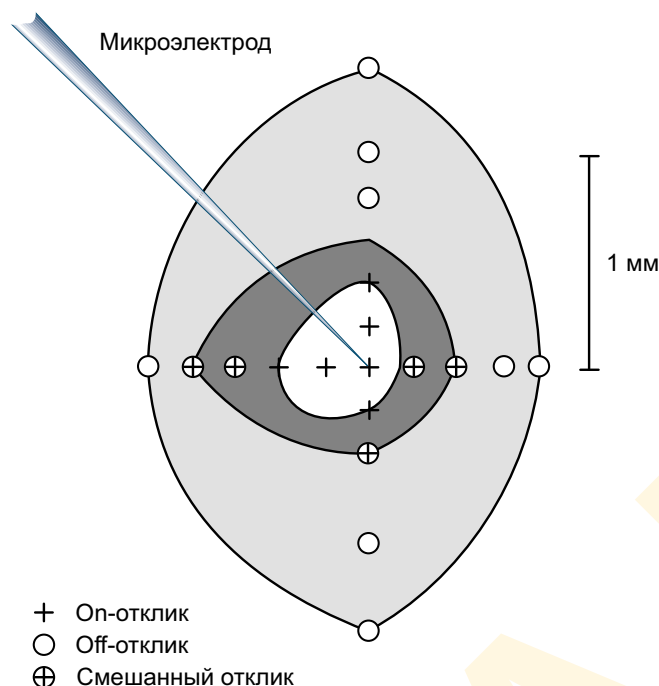
Тщательное исследование рецептивных полей позволило характеризовать структуру *оп-центр-клеток* — см. рис. 4.1.14 А, на котором степень возбуждения/торможения клетки в ответ на световое воздействие внутри рецептивного поля отложено по оси ординат. Рецептивное поле данного типа



**Рис. 4.1.14** Структура рецептивных полей ганглиозных клеток. Клетки с рецептивными полями класса оп-центр/оф-окружение (А) возбуждаются в ответ на световое воздействие в оп-центр рецептивного поля и тормозятся в ответ на прекращение этого воздействия; и наоборот — прекращают пульсацию в ответ на световую стимуляцию в оф-окружение и возбуждаются в ответ на прекращение воздействия. Собственно рецептивные поля двух видов показаны в центре, а справа даны «сомбреро» отклика как функция от диаметральной позиции рецептивного поля (пунктирная линия). Клетки с рецептивными полями класса оф-центр/оп-окружение (В) реагируют оппозитным образом.

точечную стимуляцию. Общая схема кюфлерового опыта дана на рис. 4.1.15: мелкая световая точка, вспыхнувшая в центре рецептивного поля вызывает усиление пульсации над спонтанной активностью. Когда точка попадает в окружение, клеточная активность падает. Стимуляция промежуточного участка приводит к возбуждению нейрона в начале его стимуляции и некоторому торможению при затухании стимула (Kuffler, 1953; Barlow et al., 1957).

Отклики, показанные на рис. 4.1.15, определяют рецептивное поле класса *оп-центр/оф-окружение* (примерно половина ганглиозных клеток сетчатки отвечают по такой схеме). Остальные тормозятся светом, падающим в центр их рецептивных по-



**Рис. 4.1.15** Организация «центр/окружение» рецептивных полей ретиальных ганглиозных клеток млекопитающих. В небольшом центральном участке сетчатки световая стимуляция фоторецепторов может как возбуждать, так и подавлять активность нейрона. В окружающем участке световая стимуляция фоторецептора всегда дает оппозитный эффект. Рецептивное поле, показанное на данном рисунке организовано по типу оп-центр/офф-окружение.

По материалам Kuffler, 1953

ем научиться прогнозировать картину откликов на любой зрительный стимул. Чтобы полнее описать это преобразование, нам необходимо разработать системный метод измерения нейронного отклика. Поскольку число возможных зрительных стимулов весьма и весьма велико, нам понадобится такой метод, который позволит на малом числе исходных измерений строить прогнозы откликов на все прочие стимулы.

Один из таких методов предлагает т.н. *теория линейных систем*: если отклики нейрона удовлетворяют т.н. *принципу суперпозиции* (см. Приложение...), то для прогнозирования того, как этот нейрон будет отвечать на многие другие зрительные стимулы, мы можем ограничиться всего лишь несколькими измерениями. Как мы увидим в дальнейшем, линейный принцип позволяет строить вполне адекватные прогнозы откликов большинства нейронов сетчатки и латерального колленчатого тела, по крайней мере у приматов.

Как же нам убедиться в том, что взаимоотношения входного и выходного сигналов ретиальных ганглионаров удовлетворяют принципу линейной суперпозиции?

В исследовании, о котором в 1970 г. доложили американцы Кристина Энрот-Кугель и Лоренс Пинто, изучался антагонизм центра и окружения рецептивных полей

лей и возбуждаются, когда свет действует на окружение. То есть, мы говорим о рецептивных полях класса *офф-центр/оп-окружение*.

Дендритовые кроны ганглиозных клеток сетчатки с оп-центр-рецептивными полями заключены во внутреннем плексиформном слое и тем отделены от ганглионаров с рецептивными полями типа *офф-центр*. Ганглиозные клетки, образующие связи в верхней части внутреннего плексиформного слоя, имеют *офф-центр-рецептивные* поля, тогда как нейроны, синаптирующие в нижней половине этого слоя имеют поля типа *оп-центр*. Следовательно, гистология и электрофизиология свидетельствуют о том, что существуют как минимум два типа зрительных путей, берущих начало в сетчатке: *офф-* и *оп-центр-пути*.

**Диагностика рецептивных полей ганглиозных клеток.** Классическое рецептивное поле лишь частично описывает реактивные свойства нейронов. Чтобы лучше понять суть процесса преобразования оптической информации в нейронный отклик, мы попробуем



ганглиозных клеток сетчатки кошек. Суперпонировались два стимула: точка, помещенная в центр рецептивного поля ганглиозной клетки и кольцо, помещенное в антагонистичное окружение (рис 4.1.16). Если отклик ганглиозной клетки подчиняется принципу суперпозиции, мы сможем предсказать темпоральный отклик, когда предъявим точку и кольцо одновременно. Мы видим, что по данной паре стимулов и данной ганглиозной клетке принцип суперпозиции устойчиво прогнозирует результирующий отклик нейрона.

**Измерения устойчивого сигнала.** Чтобы дать исчерпывающее описание параметров нейронного отклика, мы должны включить в характеристику рецептивного поля временной показатель. Однако о пространственно-временном описании рецептивных полей мы будем говорить чуть ниже, а пока, для простоты, удалим переменную времени из наших рассуждений и будем рассматривать отклик только после нескольких секунд экспозиции стимула и стабилизации отклика нейрона. Сей, как говорят, асимптотический отклик именуют *устойчивым откликом* нейрона. Еще раз подчеркнем, что измеряя лишь устойчивый отклик, мы исключаем временной фактор из нашего анализа.

Итак, как всегда при линейном подходе, мы должны дать полную спецификацию входного и выходного сигналов. Отметим, что одним из наибольших достижений последних лет является четкая формулировка определений входа и выхода, которую мы и используем при диагностике рецептивных полей. Оба этих определения основаны на понятии «контраст».

**Стимул (вход).** В качестве стимула мы рассматриваем пространственное оптическое изображение на поверхности сетчатки. Изображение это мы полагаем состоящим из двух основных компонентов: первый компонент — это средняя фотометрическая яркость; второй — это колебания яркости вверх-вниз от среднего значения. Общий уровень фотометрической яркости стимула всегда имеет положительное значение. Отклонения от среднего уровня — суть *контраст* стимулирующего паттерна, несущий в себе как положительные, так и отрицательные значения.

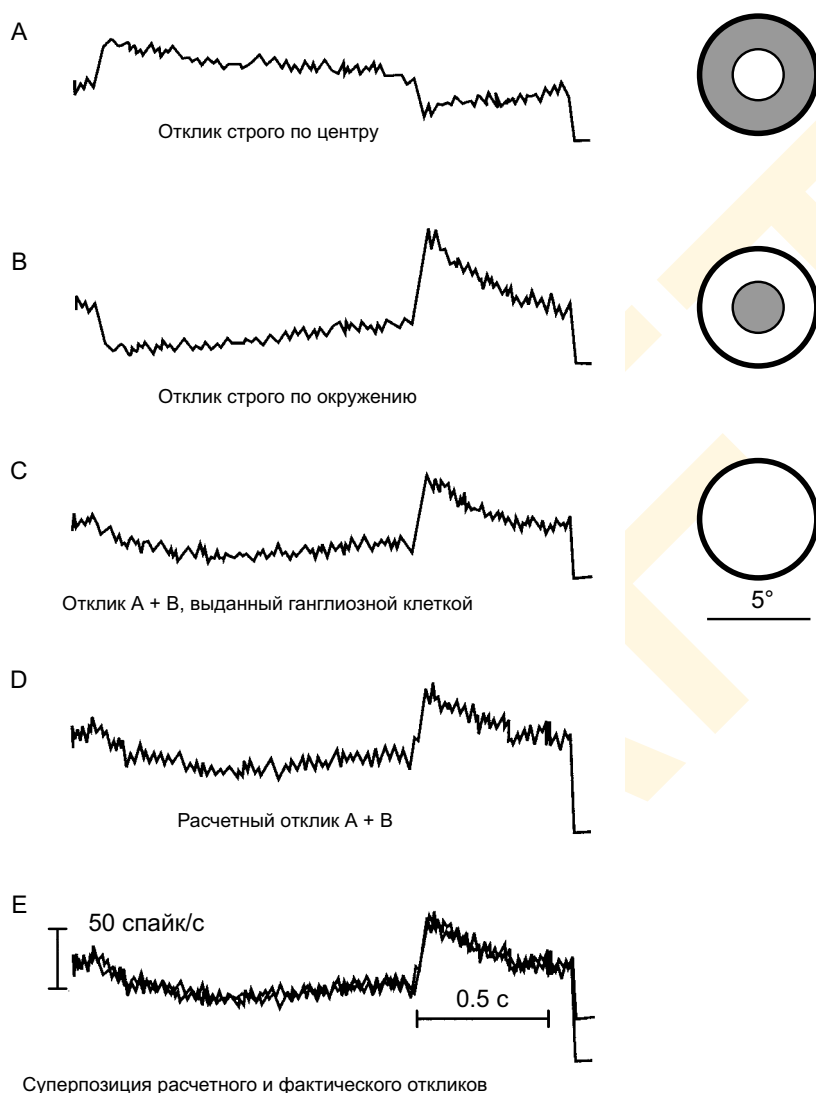
Рассмотрим простейший случай — одномерный стимул, т.е. варьирующий по одной пространственной размерности и константный по второй. Мы можем охарактеризовать контраст этого стимула одной переменной, скажем —  $x$ . Уравнение, соотносящее интенсивность стимула со средней интенсивностью фона и контрастом стимула звучит так:

$$i_x = [1 + a_x] m,$$

где  $i_x$  — фотометрическая яркость стимула,  $m$  — средняя фотометрическая яркость,  $a_x$  — контраст стимула.

**Отклик (выход).** Обозначим частоту спонтанной пульсации нейрона как  $r_0$ . Теперь предположим, что мы предъявляем зрительной системе стимулирующий паттерн с контрастом  $a$ . При этом устойчивый отклик нейрона равен  $r$  спайков в секунду (спк/с). Изменение частоты пульсации ганглиозной клетки ( $\Delta r = r - r_0$ ) мы положим откликом нейрона. Таким образом, если мы давеча рассматривали входной сигнал как *отличие* в фотометрической яркости стимула от средней яркости фона, то выходной сигнал нейрона будет представлять собой *отличие* от среднего спонтанного отклика.

**Проверка линейности.** Дабы убедиться в исполнении принципа суперпозиции, мы должны измерить отклики нейрона на стимуляцию двумя паттернами разного контраста и отклики на суммарную стимуляцию, то есть стимуляцию обоими паттернами



**Рис. 4.1.16** Тест на суперпозицию отклика ретиальной ганглиозной клетки. А — усредненный по времени отклик на мигающий диск в центре рецептивного поля. В — отклик на мигающее кольцо по периферии рецептивного поля (окружение). С — отклик на одновременную стимуляцию в центре и в окружении. D — расчетная сумма откликов по А и В. Е — сравнение фактического и расчетного откликов.

По материалам Enroth-Cugell & Pinto, 1970.

одновременно. Предположим, что мы используем два паттерна с контрастами  $a_x$  и  $b_x$  (соответствующие изменения в стационарном отклике обозначим как  $\Delta r_a$  и  $\Delta r_b$ ). Для теста на суперпозицию мы объединим два контрастных стимулирующих паттерна и получим новый стимул:

$$[1 + a_x + b_x]m.$$

Мы ожидаем при этом, что устойчивый отклик на новый стимул окажется равным

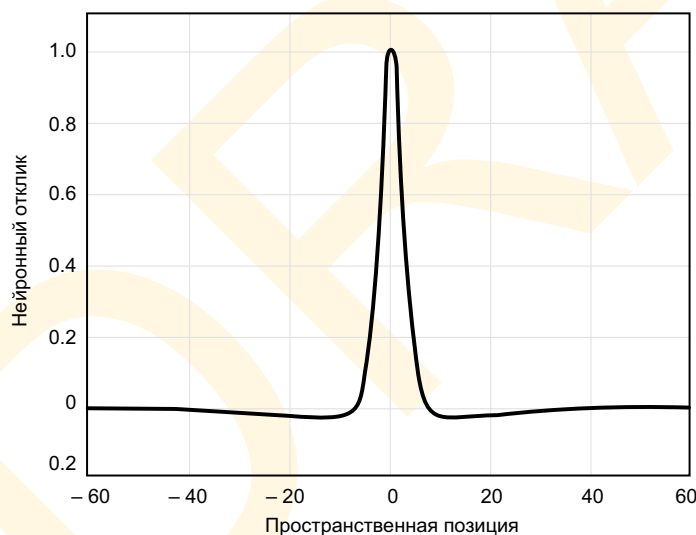
$\Delta r_a + \Delta r_b$ . Отметим, что новый стимулирующий паттерн получен сложением показателей контраста ( $a_x$  и  $b_x$ ), но не фотометрических яркостей стимулов. Если отклик нейрона удовлетворяет принципу суперпозиции, мы можем прогнозировать отклик этого нейрона на любой контрастный паттерн составлением матрицы системы.

Предположим, что мы представили одномерный стимулирующий контрастный паттерн как вектор-столбец ( $a$ ) и матрицу системы как  $1 \times N$ -матрицу ( $R$ ). Мы можем предсказать отклик нейрона простым умножением —  $\Delta r = R_a$ , что в табличной форме выглядит как:

$$\Delta r = (r_1 \dots r_i \dots r_N) \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_{N-1} \\ a_N \end{pmatrix}$$

Мы говорим, что элементы матрицы  $R$  характеризуют линейное, устойчивое рецептивное поле нейрона.

Определить элементы матрицы мы можем измерением контрастного отклика на множество стимулов, векторы контрастов которых равны нулю по всем элементам кроме одного ( $0, \dots, a_i, 0, \dots, 0$ ). Стимулы эти — суть простые линии в разных участках сетчатки. Наклон функции контрастного отклика на такую-то линию — это такой-то элемент матрицы. Кривая рис. 4.1.17 — это графическое представление элементов



**Рис. 4.1.17** Одномерный устойчивый контрастный отклик ганглиозной клетки приматов. Когда отклик нейрона на контрастный паттерн стимуляции линейен, мы можем построить матрицу, проецирующую контраст стимула на изменение частоты пульсации нейрона. Кривая на данном рисунке — есть графическое представление такой матрицы, представляющей, в свою очередь, одномерное устойчивое рецептивное поле нейрона.

матрицы системы типичного нейрона с рецептивным полем центр/окружение. То есть кривая эта представляет собой одномерное рецептивное поле нейрона.

**Двухмерное рецептивное поле.** Выполняя диагностику рецептивного поля с помощью точек (а не линий), мы сможем получить картину *двухмерного* устойчивого рецептивного поля. На рис. 4.1.18 показаны два варианта представления двухмерного рецептивного поля ганглиозной клетки сетчатки. Высота кривой на рис. 4.1.18 А демонстрирует изменение отклика нейрона на свет при стимуляции различных участков рецептивного поля точечным источником: высокие положительные значения в центре графика указывают на то, что данный нейрон возбуждается при стимуляции центра его рецеп-

тивного поля; отрицательные значения в окружении указывают на торможение при стимуляции кольцеобразного участка вокруг центра рецептивного поля.

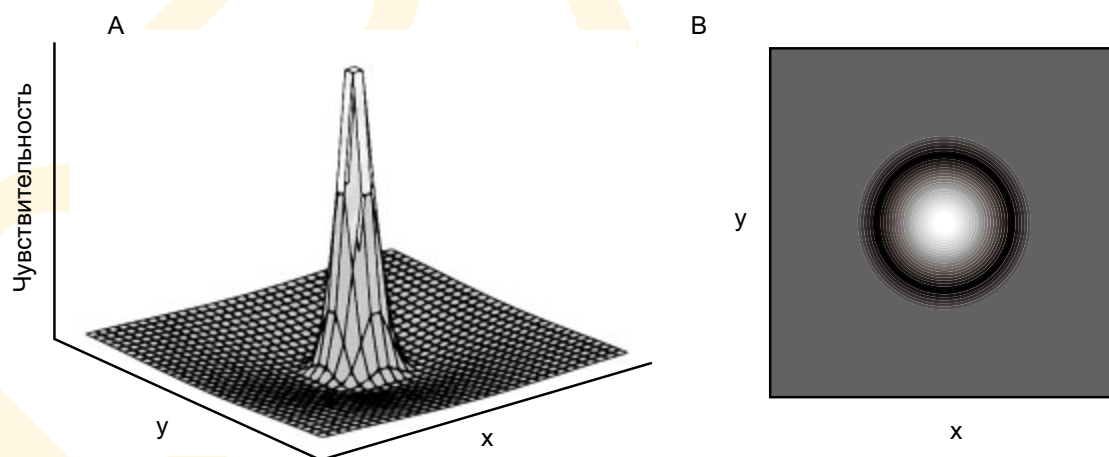
Отметим при этом, что эффект торможения в сравнении с возбуждением в центре весьма мал, но при этом область торможения занимает существенно бóльшую площадь, нежели область возбуждения. Отметим также, что предлагаемый график наглядно демонстрирует антагонизм центра и окружения, выявленный Кюфлером (1953) в рецептивных полях сетчатки кошек. В виде яркостной диаграммы результат кюфлерового опыта представлен на рис. 4.1.18 В: светлые участки — область возбуждения нейрона в ответ на световое воздействие; темные участки — область торможения также в ответ на световое воздействие.

#### 4.1.2.8 Биполярные клетки

Несмотря на то, что гистологически биполярные клетки располагаются перед ганглиозными, т.е. ближе к фоторецепторам, их функцию по ряду причин изучили позднее. Дело в том, что ганглиозные клетки — это первые клетки в зрительной системе, дающие спайковый разряд. Фоторецепторы и биполярные клетки откликаются *градуально* — непрерывным (т.е. недискретным) изменением электрического потенциала, которое движется весьма медленно и имеет место только у клеток с короткими аксонами, в частности у таких как клетки сетчатки.

Спайковый потенциал можно записать вне клетки, тогда как градуальный только путем т.н. *интраклеточной регистрации*, т.е. осложненной необходимостью внедрения электрода внутрь клетки и его длительного присутствия в ней.

Новейшие исследования, выполненные с применением интраклеточной методики, показали, что также как и у ганглиозных, у биполярных клеток структура рецептивных полей циркулярно симметрична с антагонистичными центром и окружением (Dacey, 1996; Werblin, 1969).



**Рис. 4.1.18** Двухмерное *стационарное* рецептивное поле ганглиозной клетки класса «он-центр/off-окружение», представленное двумя путями. А — высота объемного графика демонстрирует пространственную чувствительность. В сравнении с возбуждающим центром тормозящее окружение занимает ощутимо большую площадь, но в совокупный отклик нейрона вносит существенно меньший вклад, чем центр. В — изображение, за счет перепада яркостей демонстрирующее пространственную чувствительность рецептивного поля ганглиозной клетки. Светлый участок — зона, в которой стимуляция фоторецепторов возбуждет ганглиозную клетку; темное кольцо — зона торможения; серое поле — зона отсутствия эффекта.

Точная нейронная архитектура, организующая рецептивные поля биполярных клеток сегодня хорошо изучена: фоторецепторы сетчатки образуют синапсы с биполярными и горизонтальными клетками, то есть соединены с ними непосредственно; горизонтальные клетки, в свою очередь, образуют синапсы с биполярными клетками, образуя дополнительный не прямой путь от фоторецепторов к биполярам.

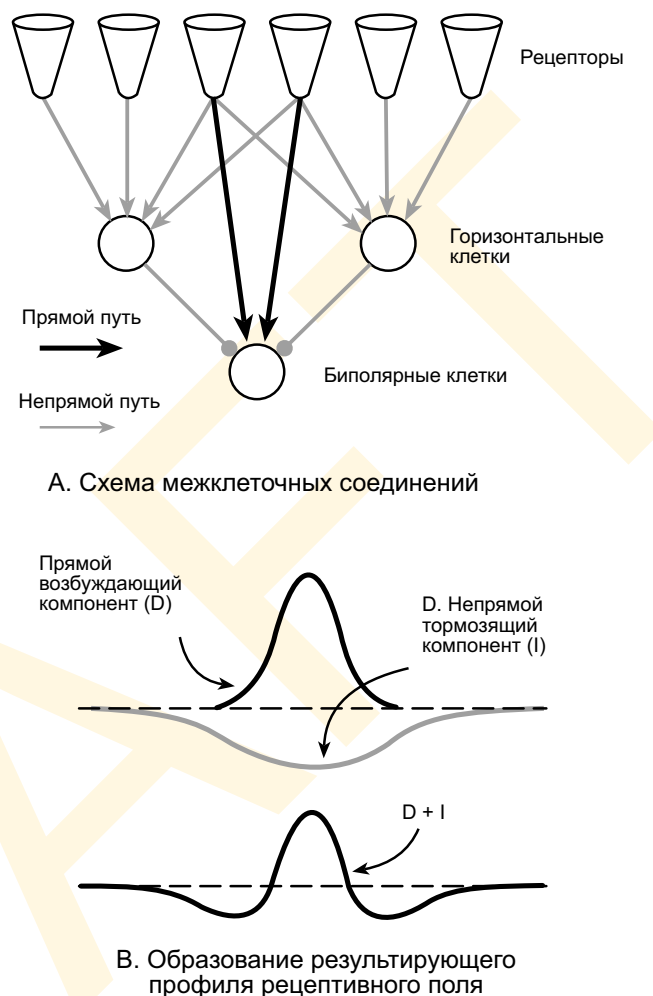
Прямой путь от рецептора к биполярной клетке бывает как возбуждающим, так и тормозящим, но при этом не прямой путь всегда окажется оппозитным. Пример сказанному дан на рис. 4.1.19 А: мы видим малый центральный возбуждающий участок и широкий (но слабый) тормозящий от непрямого пути. Оба участка (возбуждения и торможения) суммируются биполярными клетками, образуя «сомбреро» рецептивного поля (рис. 4.1.19 В).

#### 4.1.2.9 Связи с колбочками разных типов

В центральной ямке сетчатки приматов отклик по центру рецептивного поля миджет-ганглиозных клеток зависит от света, захваченного единичной колбочкой. То есть, мы говорим, что, центры рецептивных полей «наследуют» волновую чувствительность фотопигмента, заключенного во внешнем сегменте данной колбочки. У приматов центр-отклик нейронов парвоклеточного проводящего пути, с рецептивными полями возле fovea, возникает примерно в зоне оптической нерезкости, вызванной дифракцией (2.4 — 4.2 угловых минуты).

Существует минимум две основных точки зрения, толкующих связи колбочек разных типов с окружением рецептивных полей ганглиозных клеток.

В 1980 г. американский нейрофизиолог Петер Ленни (Рочестерский университет) показал, что окружение образуется случайным набором колбочек, включающих в себя колбочки L- и M-типов. Однако позднее — в 1992 г. — Клэй Рейд и Роберт Шепли пришли к выводу, что в отклик окружения рецептивного поля вносят вклад колбочки лишь одного типа. Когда отклик по центру рецептивного поля парвоклеточного пути



**Рис. 4.1.19** Схема поступления входного сигнала на биполярные клетки с рецептивными полями типа он-центр/офф-окружение. Прямые пути (черные стрелки) идут от рецепторов, расположенных в центре рецептивного поля. Непрямой путь оппозитной полярности (серые стрелки) идет по всему полю через горизонтальные клетки.



начинается с L-колбочки, отклик окружения строится на работе колбочек М. И наоборот: когда отклик центра рецептивного поля начинается с М-колбочек — окружение начинает откликаться с сигнала от колбочек L. Так или иначе, но на сей момент вопрос сегрегации колбочковых сигналов в окружении рецептивных полей нейронов парвоклеточного пути остается открытым.

Точка зрения исследователей на то, что центры и окружения в рецептивных полях нейронов магноклеточного пути получают сигнал от колбочек обоих типов — в целом, едина. Парасоль-клетка с рецептивным полем оп-центр получает возбуждающий сигнал от L- и М-колбочек, тогда как окружение — тормозящий сигнал от колбочек обоих типов, относительная мощность сигнала которых весьма изменчива (Derrington et al., 1984).

В 1984 г. американец Эндрю Марьяни (Национальный институт здоровья, США), а, спустя 10 лет, Дэннис Дэци и Барри Ли (1994) показали, что сигналы от S-колбочек кодируются внутри специальных зрительных потоков. Нейроны, получающие сигнал от колбочек S имеют очень крупные рецептивные поля (порядка 18 угловых минут). Из курса оптики мы знаем, что хроматические аберрации особо выражены в коротковолновой части спектра. Поскольку коротковолновый компонент ретинального изображения сильно размыт, крупный размер рецептивных полей S-колбочек оправдан.

Анатомические исследования Марьяни свидетельствуют о существовании биполярных нейронов, синаптирующих с некими весьма рыхло рассеянными по сетчатке колбочками, сильно напоминающими S-колбочковую мозаику. Наблюдения исследователя показали, что пространственные связи S-колбочковых сигналов также играют роль в создании рецептивных полей с большим центром (см. также Kouyama & Marschak, 1992).

Дэци и Ли подтвердили существование зрительных потоков, специализированных на передаче информации о S-колбочковом сигнале и выполнили ряд важных и весьма оригинальных наблюдений: в частности, Дэци показал, что существует морфологически самостоятельный тип ганглиозных клеток сетчатки, откликающихся на S-колбочковый сигнал и названных *бирасслоенными* (поскольку кроны их дендритов расслоены на два яруса возле внутреннего и внешнего краев внутреннего плексиформного слоя сетчатки). Опираясь на характер связей этих ганглионаров, Дэци показал, что они несут S-колбочковый возбуждающий сигнал.

Прибегнув к электрофизиологическим измерениям, Дэци и Ли сделали два дополнительных и совершенно неожиданных наблюдения — они обнаружили, что миджет-и парасоль-клетки не получают входного сигнала от S-колбочек и что все бирасслоенные ганглиозные клетки принимают возбуждающий входной S-сигнал (Dacey, 1993; Dacey & Lee, 1994).

Выходной сигнал бирасслоенные нейроны посылают в парвоклеточные слои ЛКТ (Rodieck & Watanabe, 1993). С помощью электрофизиологических методов мы можем измерить рецептивные поля в тех ядрах, чьи возбуждающие центры управляются сигналами S-колбочек. Сие возможно еще и потому, что нейроны эти дают проекции во вставки коленчатого тела. Поскольку пространственное разрешение S-колбочковой мозаики очень низкое, мы не ожидаем того, что бирасслоенные нейроны составляют весомую фракцию в нейроклеточной популяции сетчатки. Однако же их функциональное значение весьма велико, поскольку бирасслоенные клетки — это единственный морфологически идентифицированный класс ретинальных ганглионаров, несущих S-колбочковый сигнал в ЛКТ (Dacey & Lee, 1994).

### 4.1.3 Кортиковое представление

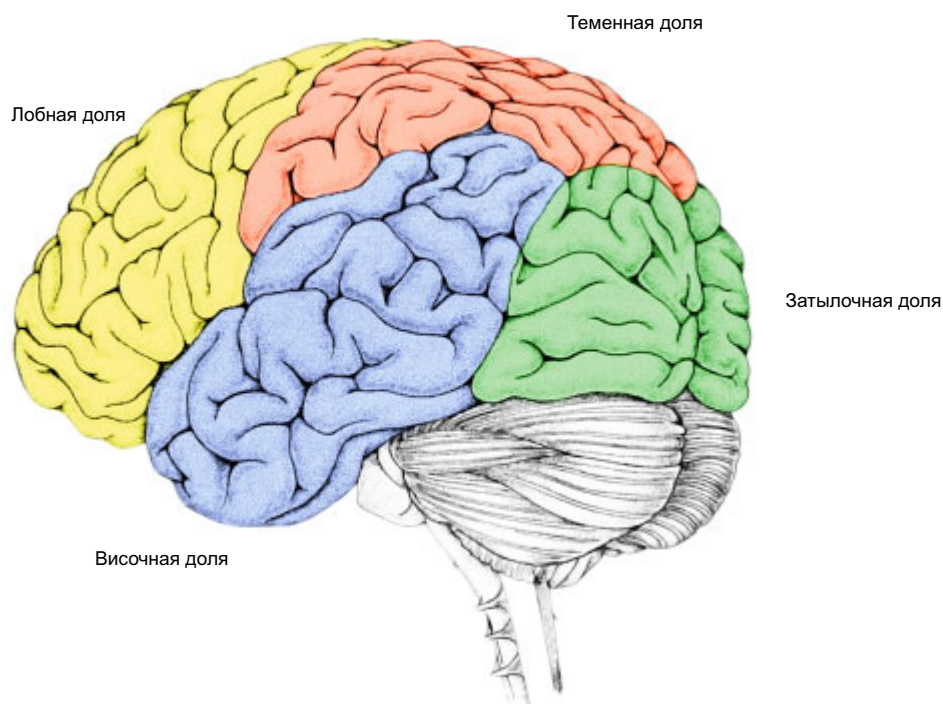
В данном разделе мы обсудим представление ретиальной информации в зрительной коре. За последние 50 лет наши знания о функции последней серьезно расширились и углубились, но даже сегодня наша точка зрения на функцию зрительной коры все время меняется — новые данные поступают порой чуть ли не каждую неделю, и поэтому, разумеется, говорить мы будем лишь о тех аспектах анатомии и физиологии зрительного отдела коры головного мозга, что безоговорочно приняты научным сообществом.

#### 4.1.3.1 Общие положения

Если посмотреть на головной мозг сбоку (рис. 4.1.20), то первое, что обращает на себя внимание — это его кора. Кора головного мозга человека — это пласт нейронов толщиной около 2 мм с общей поверхностью порядка  $1400\text{ см}^2$ . Кора, в отличие от сетчатки, гладко выстилающей внутреннюю поверхность глаза, подобна мятому листу, которым набили черепную коробку изнутри.

Места, где складчатая кора образует гребни, легко заметные невооруженным глазом, называются *извилинами*, а как щели между ними — *бороздами*.

У разных видов животных разная картина борозд и извилин, и понятно, что в мозге человека больше борозд, чем у других приматов. Мозг одной особи одного биологического вида может серьезно отличаться от мозга другой особи того же вида, но общий



**Рис. 4.1.20** Кортиковые доли головного мозга человека. Мозг принято делить на отделы, исходя из их формы и выделяя затылочную, височную, теменную и лобную доли. Исходя из внутренних связей, кору можно далее делить на множество функциональных отделов. Зрительный сигнал приходит в область V1, расположенную в затылочной доле.

рисунок борозд и извилин обычно остается неизменным. Отметим также, что борозды и извилины — это весьма удобные топографические ориентиры, при этом, скорее всего, не имеющие особого функционального значения.

Наиболее заметные борозды используются в качестве маркеров анатомического деления каждого полушария человеческого мозга на четыре доли: лобную (фронтальную), теменную (париетальную), височную (темпоральную) и затылочную (окципитальную). Каждая доля содержит множество мозговых отделов, то есть, смежных друг другу групп корковых нейронов, которые представляются работающими в единой функциональной связке.

Корковые области можно идентифицировать несколькими путями, хотя наиболее важный из них, возможно, это выявление анатомических связей с другими отделами мозга — каждый отдел мозга создает читаемую картину анатомических связей с другими отделами: входной сигнал, поступающий в один участок, прибывает обычно из нескольких мест мозга, а выходной сигнал, исходящий из этого участка, предназначен обычно для некоего специфического множества целевых участков.

У приматов большая часть сигнала сетчатки и коленчатого тела идет в определенную область коры затылочной доли, именуемой областью V1 или первичной зрительной корой. Это довольно крупный фрагмент коры головного мозга, несущий в себе порядка  $1.5 \times 10^8$  нейронов, что намного больше, чем в латеральном коленчатом теле (примерно  $10^6$ ). Область V1 можно идентифицировать по характерной полосчатости, которая есть следствие плотной паковки в одном из слоев зрительной коры миелинированных аксонов. Полосчатость эта точно соответствует размеру V1 и невооруженному глазу видится белой. Поскольку V1 топографически определена именно полосками, ее именуют *стриарной корой*.<sup>1</sup> Благодаря своей заметности, важному анатомическому расположению и большому размеру (24 см<sup>2</sup>) V1 стала предметом интенсивных исследований. Поэтому разговор в данном разделе мы начнем с описания анатомических и электрофизиологических особенностей именно стритарной коры.

Отметим, что помимо V1 были открыты еще около 20 дополнительных корковых областей, получающих сугубо зрительный сигнал. Их анатомия, электрофизиология и количественное назначение интенсивно изучаются и еще много лет будут оставаться предметом активного научного обсуждения. На протяжении нашей книги мы не раз еще вернемся к разговору о функциональной роли этих участков мозга (Zeki, 1978, 1990; Felleman & Van Essen, 1991).

Большая часть из того, что нам известно сегодня о зрительной коре пришло из экспериментов с кошками и обезьянами. Отметим, что существуют серьезные отличия в анатомии и физиологии коры у разных видов животных, и отличия эти можно продемонстрировать следующими примерами:

— в 1977 г. американец Джеймс Спрэгви (Пенсильванский университет) показал, что удаление первичной зрительной коры у кошек не приводит к их слепоте: животное прыгает, бегает и нормально воспринимает экспериментатора;

— в 1974 г. британец Николас Хампфри, изучавший поведение обезьян с удаленной V1, отмечает, что «вначале казалось, будто животное полностью ослепло, однако со временем часть зрительных функций у обезьяны восстанавливалась и зверь начал потихоньку ходить вокруг предметов, залезать на деревья и даже находить лакомство неподалеку от себя». У человека, однако, повреждение V1 ведет к полному краху всех зрительных функций. Поэтому мы будем говорить в основном об экспериментах с

<sup>1</sup> От лат. *stria* — полоса. — Прим. пер.

приматами и всегда, когда это возможно, описывать измерения, выполненные на мозге человека.

Итак, стриарная (зрительная) кора — это тонкий листок нейронов, толщиной примерно 2 мм и несколько десятков квадратных сантиметров площади. Это самая крупная область в коре приматов, содержащая порядка 200 млн. клеток, что примерно в 100 раз больше, чем количество клеток в коленчатом теле (порядка 1.5 млн.) или центральной ямке сетчатки (1 млн. клеток). Сей факт указывает на то, что в коре идут гораздо более сложные зрительные процессы, чем в нижележащих отделах зрительной системы.

Несмотря на то, что из опыта изучения травм головного мозга о первичной зрительной коре известно было давно, функциональные детали не были изучены до тех пор, пока исследователи в 50-х годах прошлого века не применили методы экстраклеточной регистрации.

#### 4.1.3.2 Исследования Хьюбела и Визеля

Дэвид Хьюбел и Торстен Визель были первыми, кто в изучении стриарной коры успешно применил, открытые Кюфлером и Барлоу, методики мэппинга рецептивных полей. Многие пытались и ранее, но не смогли подобрать соответствующие стимулы, вызывающие устойчивый отклик.

Хьюбел и Визель провели исследование весьма необычным образом (Hubel, 1995): вначале они попытались вызвать клеточный отклик, используя мелкую черную точку на предметном стекле, которое двигали перед проектором, и за счет чего на экране образовывалось темное плавающее пятно. Экран с подвижной точкой предъявлялся анестезированной кошке, у которой клетки коры давали вспышку активности.

Однако исследователям так и не удалось узнать, как с помощью такого рода стимуляции добиться устойчивого отклика: ученые полагали, что в стимуляции главную роль играет именно точка, но после серии опытов они пришли к неожиданному заключению, что клетки не дают отклика собственно на точку, но на тень края предметного стекла, движущуюся в определенном направлении по сетчатке. Дальнейшие опыты подтвердили, что ориентация и направление движения края являются определяющим фактором. Так с края предметного стекла началась эпоха изучения корковых рецептивных полей.

Дальнейшие исследования Хьюбела и Визеля (1959) быстро выявили то, что существуют несколько видов клеток коры с разным характером рецептивных полей, в частности: *простые клетки, сложные и гиперсложные*.

**Простые клетки.** Часть нейронов зрительной коры Хьюбел и Визель окрестили «простыми клетками», поскольку их отклики на сложные стимулы можно было довольно легко прогнозировать по их же откликам на стимулы простые, т.е. отдельные световые точки. Когда точка оказывается внутри возбуждающего участка, частота пульсации простых клеток увеличивается в сравнении с их спонтанной фоновой активностью. Когда световая точка оказывается в ингибиторном участке, частота пульсации падает от исходного уровня. Реакция простых клеток на большой и сложный паттерн стимуляции подчиняется принципу линейной суперпозиции и реакцию эту легко спрогнозировать обыкновенным суммированием откликов на каждую из мелких точек, образующих этот паттерн. Иными словами, относительная линейность отклика — одна из главных характеристик простых клеток.

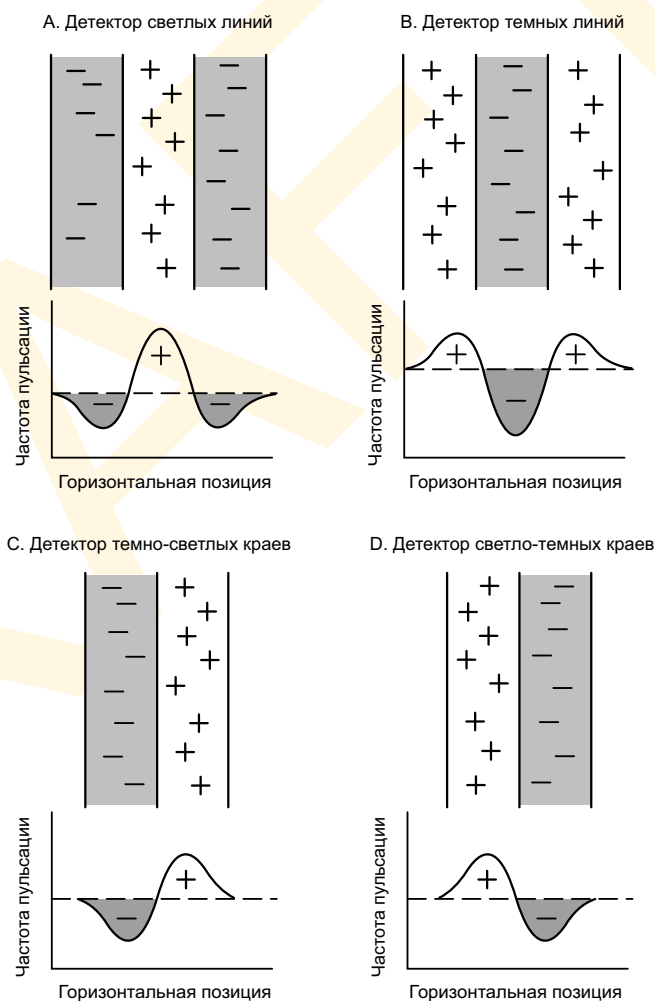
Используя такую, весьма незатейливую методику построения рецептивных полей,



Хьюбел и Визель идентифицировали несколько подтипов простых клеток. У подавляющего большинства из них форма рецептивных полей оказалась вытянутой, а отклик наиболее выраженным на стимуляцию линиями или краями в специфической ретинальной позиции и ориентации. Примеры находим на рис. 4.1.21: хорошо видно, что у одних полей область возбуждения расположена на одной стороне рецептивного поля, а область торможения — на другой. Поскольку клетки с такими рецептивными полями энергичнее откликаются на яркостные края определенных ориентации и полярности (яркий-темный или темный-яркий), их часто называют *краевыми детекторами*.

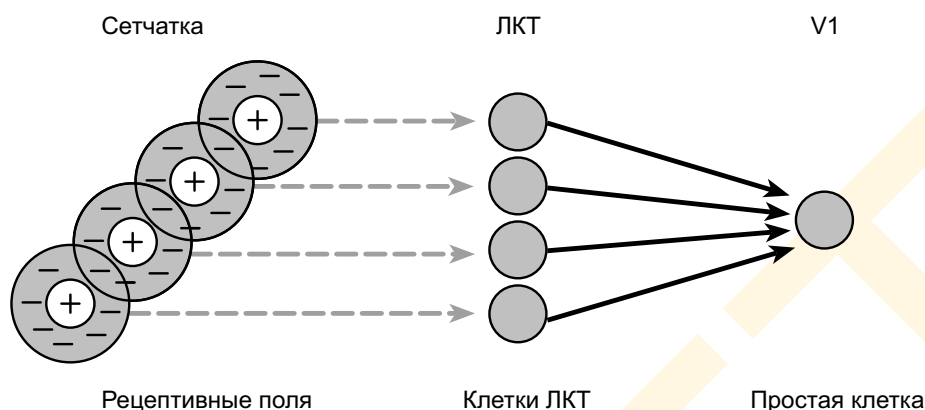
У рецептивных полей простых клеток второго подвида имеется вытянутый центральный участок, который может быть либо возбуждающим, либо тормозящим и, разумеется, с антагонистичным полем по обеим сторонам. Поскольку отклик этих клеток максимален на яркие или темные линии, эти клетки часто называют *детекторами линий* или *бардетекторами*. Хьюбел и Визель полагали, что такие краераспознающие и линиераспознающие образования в коре формируются из соответствующим образом расположенных образований колленчатого тела, организованных по принципу центр-окружение (рис. 4.1.22).

Точка зрения на обработку ретинального изображения, следующая из описанных открытий, такова, что начальный этап пространственной обработки ретинального изображения — это поиск линий и краев в этом изображении. В этой логике концепция детекции краев и линий весьма близка к атомистической точке зрения, обсуждавшейся в гл. 2, за исключением того, что «атомы», из которых строится пространственное зрение собраны из краев и линий определенных позиции, ориентации и контраста. Высокоуровневые свойства предметов, такие как их форма и ориентация, можно выявить сложив воедино множест-



**Рис. 4.1.21** Рецептивные поля и профили откликов вертикально ориентированных простых клеток V1. Показаны рецептивные поля (схемы) и профили откликов (графики) простых клеток четырех типов: детекторов светлых линий (А), детекторов темных линий (В), детекторов темно-светлых краев (С) и детекторов светло-темных краев (D). В V1 находятся клетки, настроенные также на другие пространственные ориентации (не показаны).





**Рис. 4.1.22** Возможные нейронные межсоединения клеток колленчатого тела с простыми клетками V1. Предполагается, что простые клетки получают входной сигнал от отдельных клеток ЛКТ с рецептивными полями, центры которых расположены в определенном направлении, что определяет пик чувствительности к стимулу такого же пространственного направления.

во откликов на локальные края и линии, идентифицированные клетками-детекторами в V1. Верно сие рассуждение или нет? — вопрос открытый, но уже в течение нескольких десятилетий оно доминирует в научном сообществе.

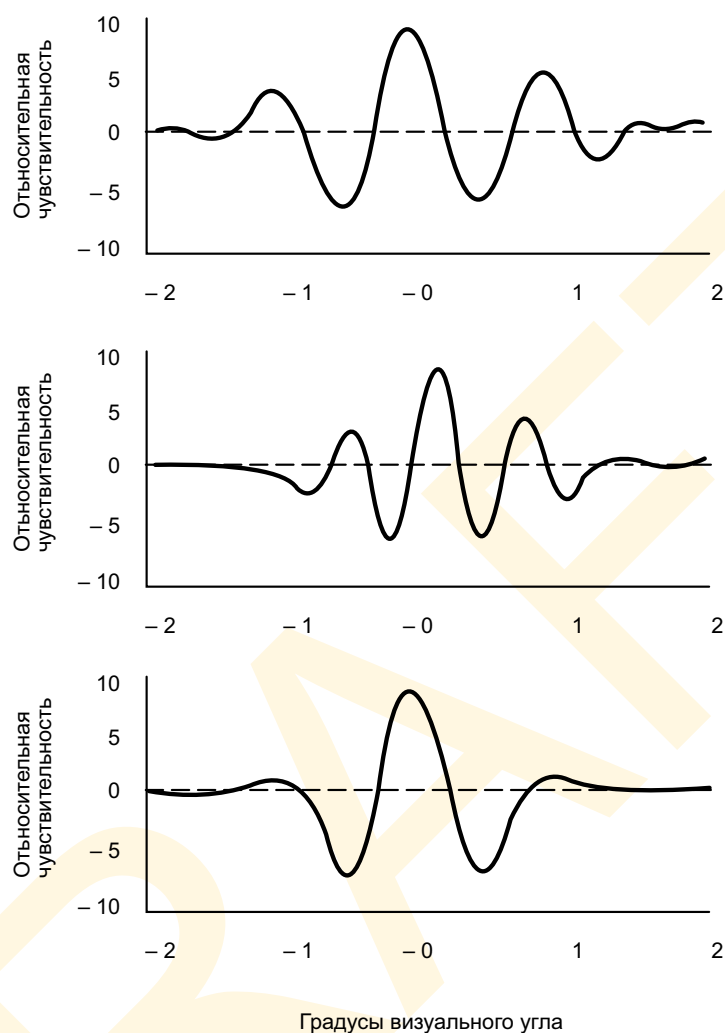
Новейшие исследования показали, что рецептивные поля простых клеток организованы чуть сложнее, чем представляли себе Хьюбел и Визель, к примеру: у многих нейронов на стороне основного горба имеются дополнительные небольшие горбы возбуждения и торможения (рис. 4.1.23) (De Valois & De Valois, 1988). Также обнаружено было множество рецептивных полей, разнящихся по размеру: большие поля отвечали на пространственно-крупные стимулы, а мелкие поля — на мелкие. Более того, вероятно, существует определенная корреляция между первым и вторым свойствами рецептивных полей: у простых клеток с малыми рецептивными полями больше горбов возбуждения-торможения, у клеток с большими полями таких горбов меньше (De Valois & De Valois, 1988).

Описанные поля можно толковать как подварианты рецептивных полей Хьюбела — Визеля, но отметим, что последние проявляют себя в ситуациях, когда клетки просто реагируют на обычную стимуляцию краями и линиями.

Профессоры университета Беркли, супруги Рассел Де Валоис и Карен Де Валоис в своей фундаментальной работе «Пространственное зрение» (1980, 1988) дали развернутое, но в целом иное толкование функциональной роли этих клеток. Однако мы отложим его разбор до тех пор, пока не начнем говорить о психофизических теориях обработки ретинального изображения, поскольку рассуждения ученых строятся именно на этих теориях.

**Сложные клетки.** Основной вид клеток стриарной коры — *сложные клетки*. Несмотря на то, что их рецептивные поля тоже вытянуты, сложные нейроны отличны от простых по ряду весьма важных свойств:

1. **Нелинейность.** Отклик сложных клеток сильно нелинеен и в целом они едва отвечают на стимуляцию мелкими неподвижными точками. В результате мэппинг их рецептивных полей неосуществим по технике измерения откликов на точечную стимуляцию по разным локалям сетчатки, поскольку не подчиняются правилу линейной суперпозиции.



**Рис. 4.1.23** Количественные профили откликов рецептивных полей трех простых клеток. На каждом из графиков дана частота пульсации коркового нейрона в ответ на тонкий штрих приоритетной для этого нейрона ориентации в разных участках рецептивного поля (ось абсцисс). Стоит обратить внимание на дополнительные горбы возбуждения и торможения в верхних двух графиках (De Valois & De Valois, 1988).

2. *Чувствительность к движению.* Сложные клетки легко реагируют на *движение* линий или краев в любой точке их рецептивных полей, но зачастую чувствительность проявляет себя лишь при определенном направлении движения стимула (рис. 4.1.24).

3. *Нечувствительность к позиции.* Сложные клетки не особо чувствительны к местоположению стимула — легкие отличия в локализации полос в целом не сильно влияют на величину отклика.

4. *Размер рецептивных полей.* Рецептивные поля сложных клеток в среднем несколько больше, чем таковые у клеток простых.

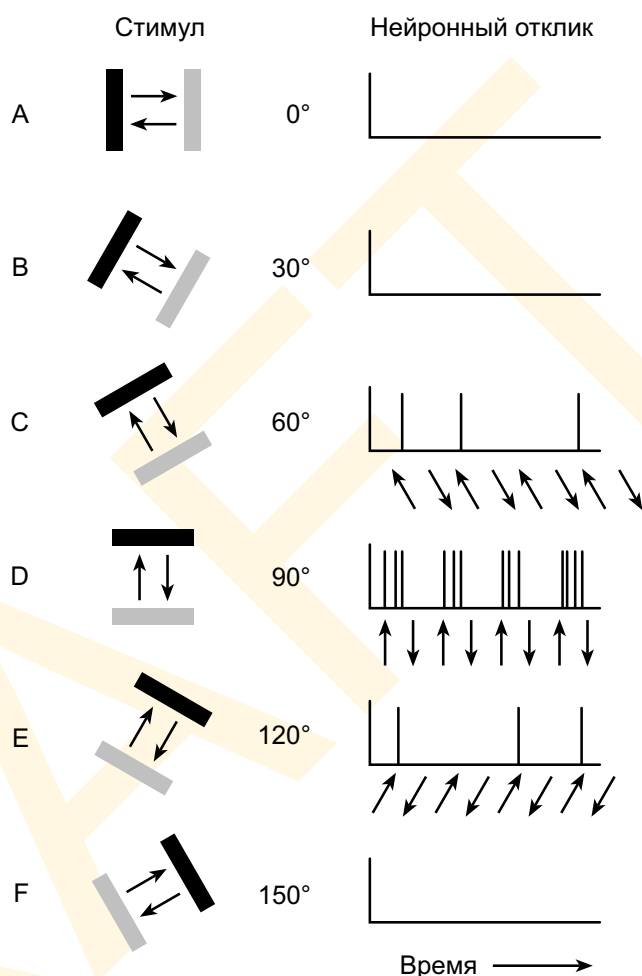
Первые кортикальные нейроны, с которыми Хьюбел и Визель стали «разговаривать» краем предметного стекла — это, фактически, были сложные клетки. Приняв то, что примерно 75% клеток стриарной коры — это клетки сложные, легко понять,

почему при предыдущих исследованиях кора не давала явного отклика на стимуляцию рассеянным светом, неподвижными световыми и темновыми точками и даже точками движущимися. Чтобы добиться выраженного отклика сложных клеток, необходимо предъявлять зрению вытянутые контуры, движущиеся в определенном направлении и воздействующие на определенный участок сетчатки. На что, в итоге, и наткнулись Хьюбел с Визелем: им повезло в том, что для создания стимула они взяли обычное предметное стекло — оказалось в их руках, как говорят, «навороченная» оптическая установка, Бог весть, сколько лет еще исследования кортикальных рецептивных полей не сдвинулись бы с мертвой точки!

Исследователи предположили, что сложные клетки собирают отклики от множества простых клеток (рис. 4.1.25). В целом так и оказалось, хотя некоторые сложные клетки получают сигнал непосредственно из колленчатого тела. Сия архитектура объясняет то, почему сложные клетки откликаются только на определенным образом ориентированные штрихи и края.

**Гиперсложные клетки.** Хьюбел и Визель идентифицировали третий тип стриарных клеток, которые окрестили *гиперсложными*, поскольку их рецептивные поля оказались еще более селективными, нежели поля сложных клеток. Главная характеристика гиперсложных клеток такова, что чем длиннее стимулирующая сетчатку линия или край, тем ниже активность этих клеток и наоборот: короткие линии и края вызывают максимум активности гиперсложных нейронов коры (рис. 4.1.26). По этой причине их часто называют *стопорными клетками*.

Сегодня исследователи полагают, что на самом деле — это простые или сложные клетки, играющие стопорную роль, но не отдельный качественно отличный тип клеток.



**Рис. 4.1.24** Отклик сложной клетки зрительной коры домашней кошки на движущиеся штрихи. Стимулы двигались вперед-назад перпендикулярно их ориентации (левая колонка). Отклик на стимуляцию показан в правой колонке. Максимальный отклик данных клеток приходится на стимуляцию горизонтальными штрихами, движущимися снизу вверх (D и в некоторой степени C и E), но не наоборот и не в каком-то другом направлении (Hubel & Wiesel, 1959).

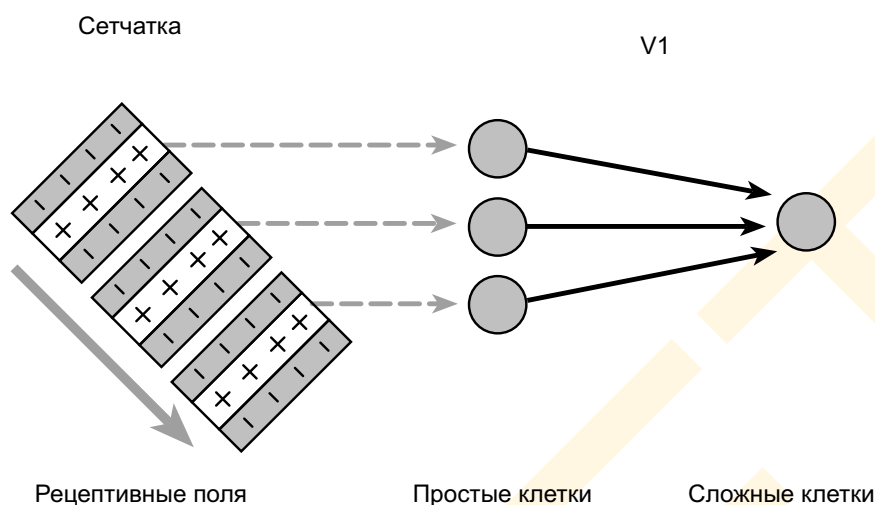


Рис. 4.1.25

Дальнейшие исследования показали, что мера «стопорности» непрерывна, т.е. не дискретна и не подпадает под определение «всё или ничего» (De Angelis, Freeman & Ohzawa, 1995). Действительно, сие может иметь место за счет латерального торможения от соседних корковых клеток. На рис. 4.1.27 дана схема, демонстрирующая свойства стопорных клеток за счет объединения выходного сигнала двух сложных клеток.

#### 4.1.3.3 Архитектура первичной зрительной коры

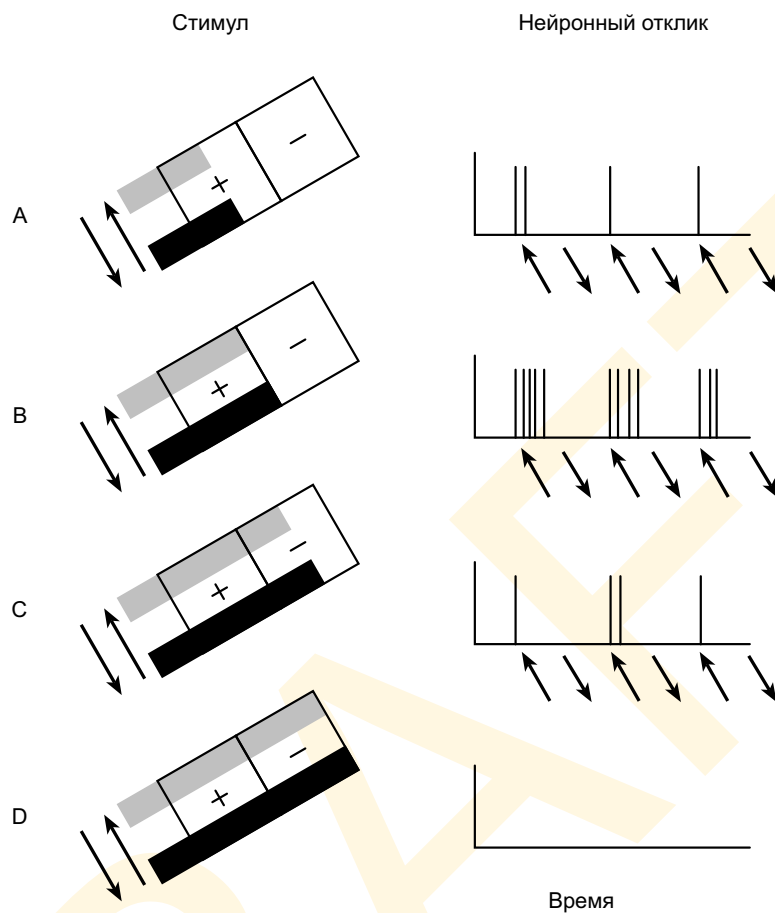
Понимать характер взаимосвязей между зрительными областями коры крайне важно. Специфическая картина сигналов, транслируемая в V1 от обеих сетчаток через коленчатое тело таламуса, строго закономерно согласуется с архитектурой первичной зрительной коры. Однако мы вначале рассмотрим анатомию V1, а затем обсудим то, как она вписывается в общую организацию зрительной информации, представленной в стриарной коре.

Как и вся кора головного мозга, зрительная кора состоит из слоев. На рис. 4.1.28 А показан ее срез, и несколько крупных слоев видны сразу. Однако, исходя из отличий в относительной плотности залегания нейронов, аксонов и синапсов, а также отличий в связях с прочими отделами мозга, область V1 делят на шесть слоев.

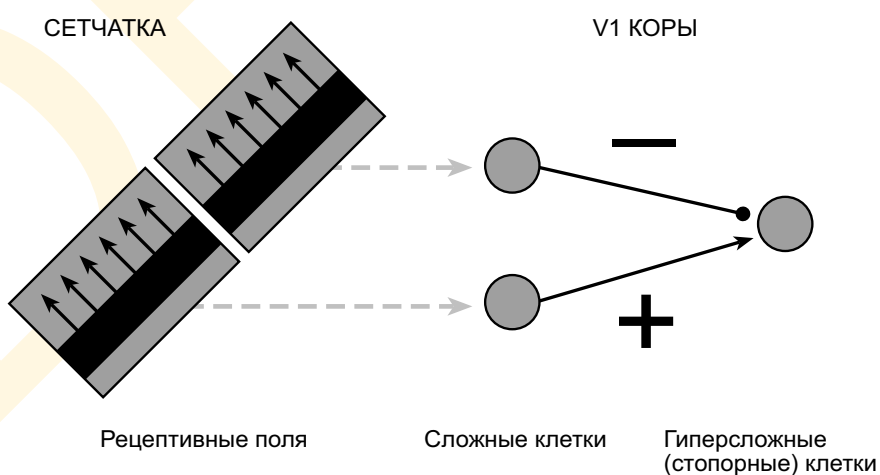
**1-й слой.** В первом, поверхностном слое зрительной коры нейронов совсем немного, но множество аксонов и синапсов, вместе именуемых как *нейропиль*.

**2-й и 3-й слои.** Слои 2 и 3 несут в себе плотный массив клеточных тел и множество локальных дендритовых межсоединений. Эти слои, судя по всему, получают прямой входной сигнал от вставок коленчатого тела (Fitzpatrick et al., 1983; Hendry & Yoshioka, 1994) и при этом посылают выходной сигнал в прочие корковые области. Слои 2 и 3 выявляются лишь с помощью специальных гистологических окрасок.

Чаще всего слои с 1-го по 3-й объединяют и именуют *поверхностными слоями* зрительной коры.

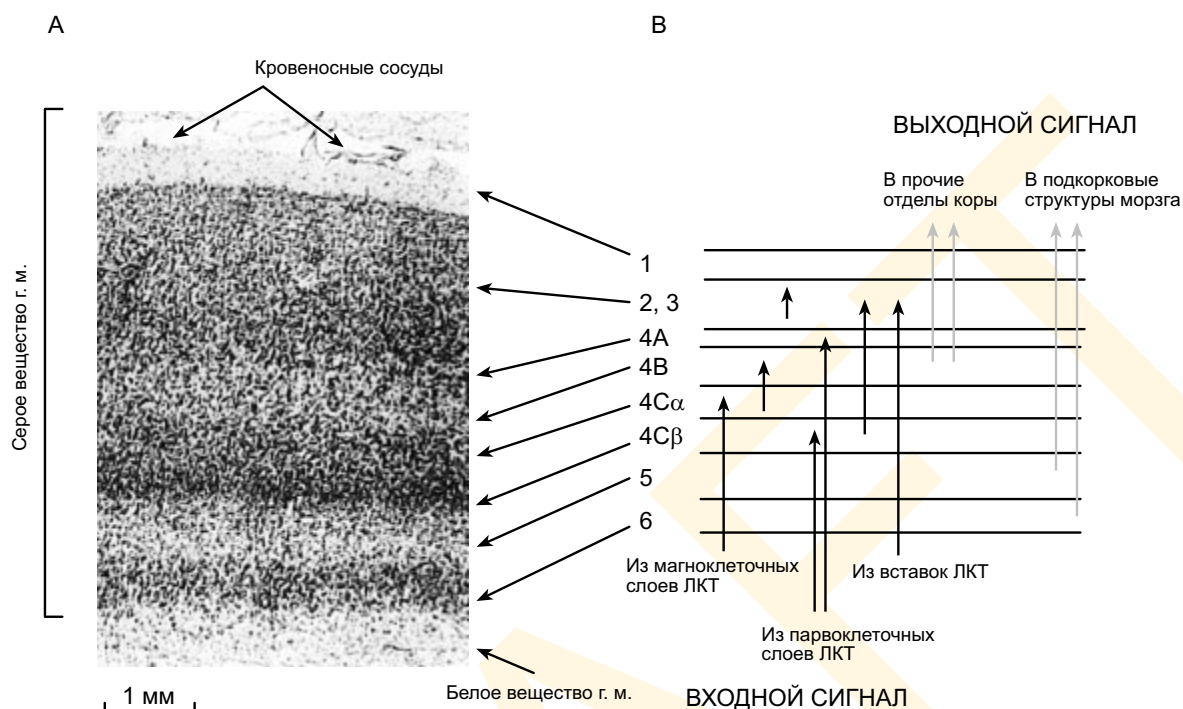


**Рис. 4.1.26** Отклик гиперсложных (стопорных) клеток зрительной коры кошки на движущиеся штрихи различной длины. Отклик этих клеток растет на движение вверх и влево, когда темная диагональная линия попадает в левую нижнюю часть рецептивного поля. Отклик падает, если линия дотягивается до правой верхней части поля.



**Рис. 4.1.27** Возможные связи сложных клеток с гиперсложными. Гиперсложные клетки ориентированы на получение входного сигнала от двух и более сложных клеток, причем в том числе и на возможное тормозящее влияние (Kuffler & Nichols, 1976).





**Рис. 4.1.28** Слои V1. А — специальным образом окрашенный срез зрительной коры головного мозга макаки с четко различимой слоистой структурой. Каждый слой содержит разные количества клеточных тел, дендритов и аксонов и поэтому идентифицируется по интенсивности окраски и другим признакам. В — организация входного и выходного сигналов: парвоклеточные и магноклеточные сигналы ЛКТ приходят в слой 4С; нейроны из вставок ЛКТ синаптируют в поверхностных слоях V1; выходной сигнал зрительной коры идет назад в колленчатое тело и прочие подкорковые ядра головного мозга.

С любезного разрешения Jennifer Lund.

**4-й слой.** Четвертый слой разделен на несколько подслоев согласно межсвязям его нейронов с другими отделами мозга и слоями. Слой 4С получает первичный входной сигнал из парвоклеточного и магноклеточного слоев колленчатого тела. Магнонейроны посылают свои выходные сигналы в верхнюю половину этого слоя (подслой 4Cα), тогда как парвонероны образуют связи с нижней его половиной (подслой 4Cβ).

Нейроны слоя 4В получают мощный входной сигнал из 4Cα и посылают выходной сигнал в прочие области коры. Слой 4В виден невооруженным глазом по наличию крупной полосчатости (*стрия Геннари*), образованной в основном аксонами корковых нейронов.

**5-й слой.** Пятый слой в сравнении с окружающими его слоями несет в себе относительно небольшое число клеток и посылает мощный выходной сигнал в верхние бугры четверохолмия — структуру среднего мозга.

**6-й слой.** В шестом слое клетки упакованы плотно и посылают мощный обратный сигнал в колленчатое тело (Lund et al., 1975).

В целом (хотя бывают и исключения) выходной сигнал, идущий в новые корковые отделы, чаще приходит из поверхностных слоев и заканчивается в четвертом слое. Об-

ратные проекции обычно идут из нижних слоев и заканчиваются в первом слое и шестом (Rockland & Pandya, 1979; Felleman & Van Essen, 1991).

Принципиальная схема рис. 4.1.28 В демонстрирует нам то, что сигналы, идущие к и от V1 — сложны и высокоспецифичны. Поэтому мы можем предположить, что нейронные межсоединения внутри стриарной коры тоже специфичны: действительно, примерно 25% нейронов во всех слоях — это ингибиторные промежуточные нейроны, и их межсвязи должны управляться наличием биохимических маркеров, идентифицирующих, какой нейрон с каким нейроном соединен (и как). Гистологическая классификация клеточных типов зрительной коры и идентификация локальной схемы дадут нам очень многое в понимании огромной функциональной роли V1.

#### 4.1.3.4 Пути в зрительную кору

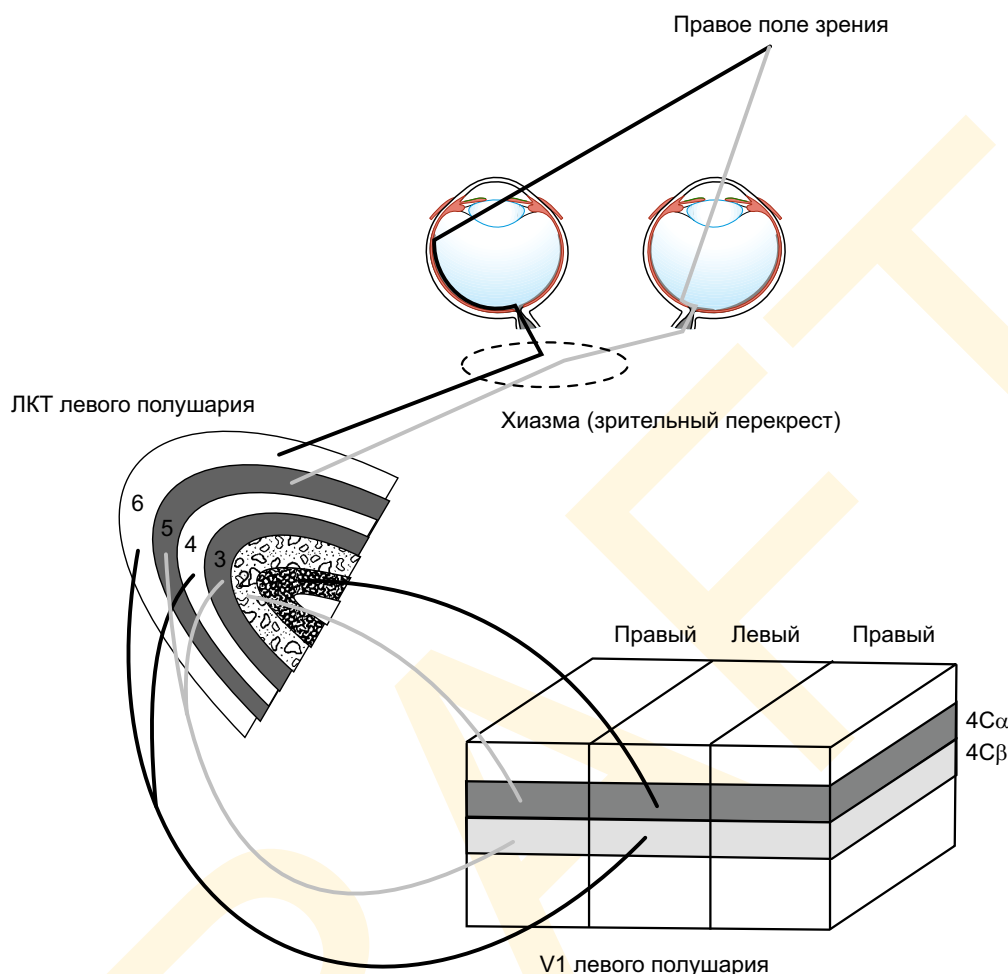
Анатомическая структура проводящих путей, идущих от обеих сетчаток в кору, определяет базовые свойства V1. В первую очередь, отметим, что V1 левого полушария получает сигнал лишь от правой половины поля зрения, а V1 правого полушария — от левой половины. Причины сего легко понять, наблюдая ход ретинальных сигналов через хиазму (рис. 4.1.29), в которой сигналы эти делятся на две новых группы (каждая из которых образует связи только в одном из полушарий). Аксоны ганглиозных клеток, чьи рецептивные поля стимулируются левой половиной поля зрения, посылают свой выходной сигнал в коленчатое тело правого полушария головного мозга, тогда как аксоны ганглиозных клеток, стимулируемые правой половиной поля зрения — в ЛКТ левого полушария. Следовательно коленчатое тело в обоих полушариях мозга получает ретинальный сигнал от обоих глаз, но при этом только от одной части поля зрения (рис. 4.1.29).

**Ведущий глаз.** Внутри коленчатого тела таламуса информация о глазе — источнике сигнала сохраняется, и волокна зрительного нерва каждого глаза образуют связи в разных слоях ЛКТ: парвоклеточные и магноклеточные слои (нумеруются от 1 до 6) 2, 3 и 5 получают входной сигнал от сетчатки «своего» глаза, а слои 1, 4 и 6 — от сетчатки «чужой» стороны. При этом по сей день остается загадкой то, почему картина окулярных связей именно такова. Глаз — источник сигнала для вставок ЛКТ все еще не установлен.

Сигнал от обоих глаз остается сепарированным вплоть до его прибытия в слой V1, и мы можем отследить это деление измерением электрофизиологических откликов нейронов в слое 4С: если электрод осторожно сдвинуть внутри 4С, то в какой-то момент смены ведущего глаза возникнет резкая смена сигнала. В слое 4С смена ведущего глаза происходит при смещении электрода на расстояние порядка 50 мкм.

Выше и ниже слоя 4С сигналы от обоих глаз сходятся на общих нейронах, хотя по-прежнему имеет место тенденция отдельных клеток получать входной сигнал преимущественно от одного глаза. Смена ведущего глаза в поверхностных слоях коры не столь резка — порядка 100 мкм.

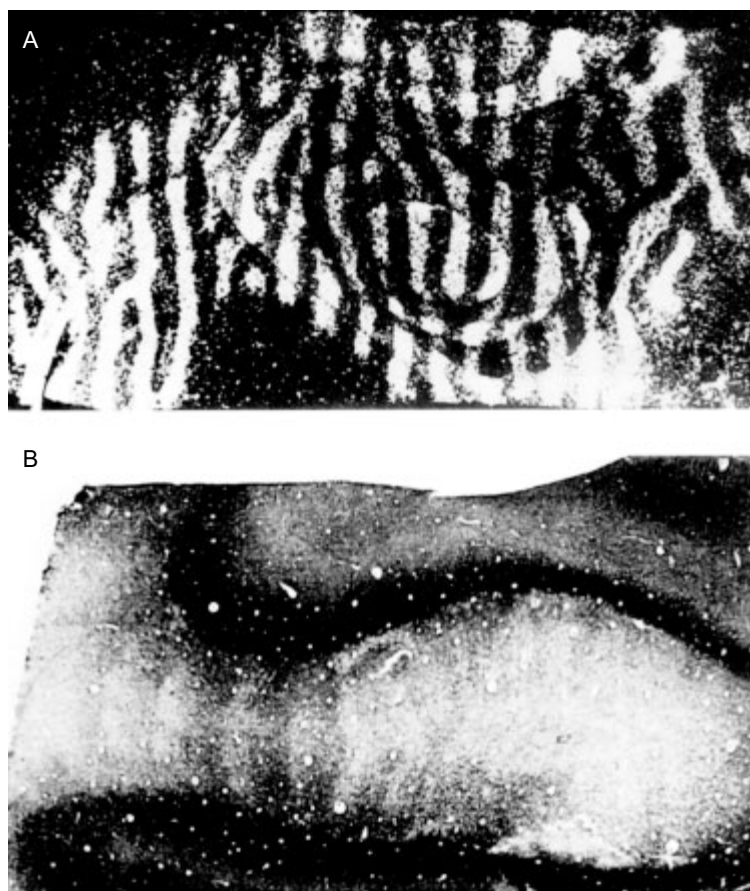
Корковое деление визуальной информации по ведущим глазам привело исследователей к тому, что стали рассуждать в логике т.н. *пластов окулярных доминант* (Hubel & Wiesel, 1977; Bishop, 1984). Дело в том, что в дополнение к фактам, полученным от электрофизиологических измерений, мы можем применять и традиционные анатомические методы визуализации зон коркового доминирования левого и правого глаз, в частности, использовать специальный маркер — помеченную тритием аминокислоту пролин. Пролин после инъекции в сетчатку переносится через синаптические связи в



**Рис. 4.1.29** Сигналы от обеих сетчаток поступают в V1 зрительной коры через ЛКТ. Точки из правого поля зрения проецируются на темпоральную (височную) сторону левого глаза и назальную (носовую) сторону правого. Аксоны ганглиозных клеток из этих ретинальных областей синаптируют в разных слоях ЛКТ левого полушария. Нейроны магноклеточных и парвоклеточных слоев ЛКТ посылают сигналы в корковые слои 4Cα и 4Cβ соответственно. Сигналы от каждого глаза разделены по полосам в V1. Сигналы от этих полос конвергируют на отдельных нейронах поверхностных слоев коры.

кору. Выполнив тангенциальный разрез коры по слою 4C и приложив срез к фотоэмульсии, мы получим на ней (после проявки, разумеется) картину темных и светлых полос, соответствующую картине абсорбции меченого пролина. На рис. 4.1.30 дан полосчатый паттерн, светлые полосы которого маркируют те участки коры, что получили входной сигнал от инъецированного глаза; промежуточные темные участки — это участки, получившие входной сигнал от второго (не инъецированного). В коре макаки шаг полос составляет примерно 400 мкм; у человека — примерно 800 мкм (Hubel et al., 1978; Horton & Hoyt, 1991).

**Бинокулярные нейроны.** В поверхностных слоях V1 приматов множество нейронов откликаются на сигнал от обоих глаз, к примеру: у обычной макаки 80% в поверхностных слоях V1 бинокулярно ведомы. Развитие межсоединений, необходимых для управления бинокулярными нейронами зависит от зрительного опыта, приобретаемо-



**Рис. 4.1.30** А — пласты окулярных доминант в стриарной коре макаки; В — то же в коре человека.

го в процессе созревания особи. Хьюбел и Визель (1965) показали, что веко, ушитое сразу после рождения животного, или подрезанная окуломоторная мышца радикально влияют на развитие нейронов в V1. В частности, не развиваются бинокулярные нейроны. Если один из глаз оставался закрытым в критический период развития, животное останется слепым на этот глаз в течение всей оставшейся жизни. Сие принципиально отлично от ситуации, когда глаз взрослой особи оказывается закрытым на несколько месяцев: отсутствие стимуляции коры от этого глаза не оказывает на кору сколько-нибудь значимого эффекта (Hubel et al., 1977a; Shatz & Stryker, 1978; Movson & Van Sluyters, 1981; Mitchell, 1988). У кошек нормальное развитие пластов окулярных доминант и особенно бинокулярных межсоединений также зависит от нейронной активности обеих сетчаток (Stryker & Harris, 1986).

Итак, коль скоро у нас два глаза, и оба дают проекции в полушария мозга, то автоматически встает вопрос о том, сколько же ретинотопических карт в коре — по одной на каждый глаз или одна общая? Ответ, как водится, где-то между: в коре каждого из полушарий одна глобальная карта, внутри которой клетки, принимающие проекции от одного глаза, доминируют над клетками, принимающими проекции от другого (Le Vay, Hubel & Wiesel, 1975).



**Гиперколонки.** Общая ретинотопическая карта собрана из множества мелких корковых единиц — т.н. *гиперколонок*, задающих мелкозернистую структуру этой карты. Гиперколонки — это длинные тонкие столбцы корковой ткани примерно 1×1 мм в сечении, идущие перпендикулярно всем шести корковым слоям. Одна из таких колонок показана на рис. 4.1.31 по левоглазо- и правоглазодоминантным клеткам. Внутри каждой колонки имеет место регулярная клеточная последовательность по одной из пространственных размерностей, что представляет настройку нейронов на определенную пространственную ориентацию стимула? (Hubel & Wiesel, 1968). Исследования, основанные как на экстраклеточной записи, так и на автордиографии, единодушно свидетельствуют именно о такой корковой архитектуре.<sup>1</sup>

**Ретинотопическая организация.** Пространственное взаиморасположение ретинальных ганглионаров копируется пространственной организацией нейронов колленчатого тела: в задней части ядер ЛКТ находятся нейроны, рецептивные поля которых близки к центральной ямке сетчатки, но по мере продвижения кпереди рецептивные поля постепенно смещаются к периферии сетчатки. Такой принцип пространственного расположения именуется *ретинотопической организацией*.

Сигналы V1 тоже ретинотопичны. Электрофизиологически, с помощью электрода тангенциально идущего через слой 4С и перпендикулярно пластам окулярных доминант, мы можем определить расположение рецептивных полей V1 макаки. Рецептивные поля, центрированные по нейронам вдоль траектории электрода, устойчиво согласуются с аналогичными траекториями от центральной ямки к периферии сетчатки. Сия тенденция, разумеется, прерывается короткими скачками на краях пластов. Внутри первого пласта позиции центров рецептивных полей меняются плавно, но по мере того как мы продвигаемся далее к следующему пласту, возникает резкий сдвиг положения рецептивного поля равный примерно половине пространства занимаемого рецептивными полями нейронов первого пласта. Хьюбел и Визель (1977) окрестили такую пространственную организацию как «два шага вперед, шаг назад».

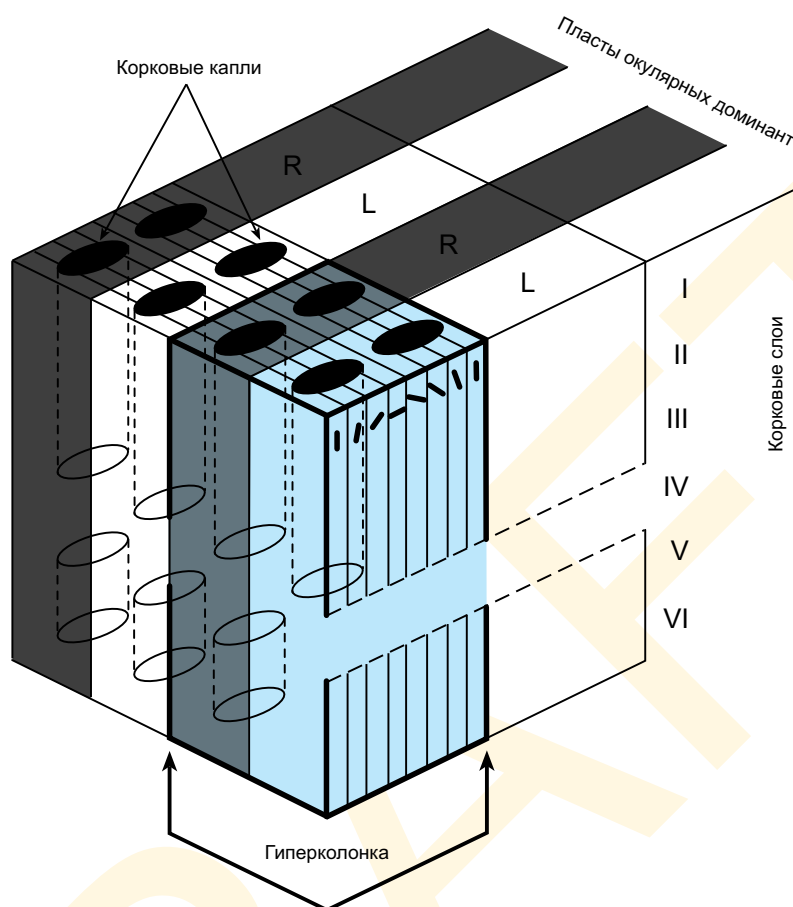
У человека область V1 расположена в т.н. *sulcus calcarinus* — *шпорной борозде* затылочной доли мозга. Шпорная борозда головного мозга одного из трех соавторов нашей книги — Брайана Ванделла, а также ее ретинотопическая организация, показаны на рис. 4.1.32.

Нейроны с рецептивными полями в центральной области поля зрения расположены в задней части шпорной борозды, а нейроны с рецептивными полями на его периферии — в передней части. На определенном отрезке вдоль борозды рецептивные поля располагаются по полуокружности поля зрения. Нейроны с рецептивными полями в верхней, средней и нижней частях этой полуокружности обнаруживаются соответственно в нижней, средней и верхней частях шпорной борозды (Inouye, 1909; Holmes, 1918, 1945; Horton & Hoyt, 1991).

В 1994 г. Стивен Эйнгель (Калифорнийский университет) диагностировал ретинотопическую организацию зрительной системы человека, пройдя от центральной ямки до периферии сетчатки. Стимул, им применявшийся, показан на рис. 4.1.33. Стимул состоял из серии медленно расширяющихся колец, каждое из которых представляло собой совокупность мигающих квадратиков. Кольцо начиналось с небольшой точки в определенной позиции поля зрения и затем увеличивалось до тех пор, пока не исчезало за его пределами. Как только кольцо пропадало, новое появлялось в центре.

<sup>1</sup> Альтернативную точку зрения, предложенную Расселом и Карен Де Валоис (1988) мы рассмотрим в текущей главе, но несколько позднее.

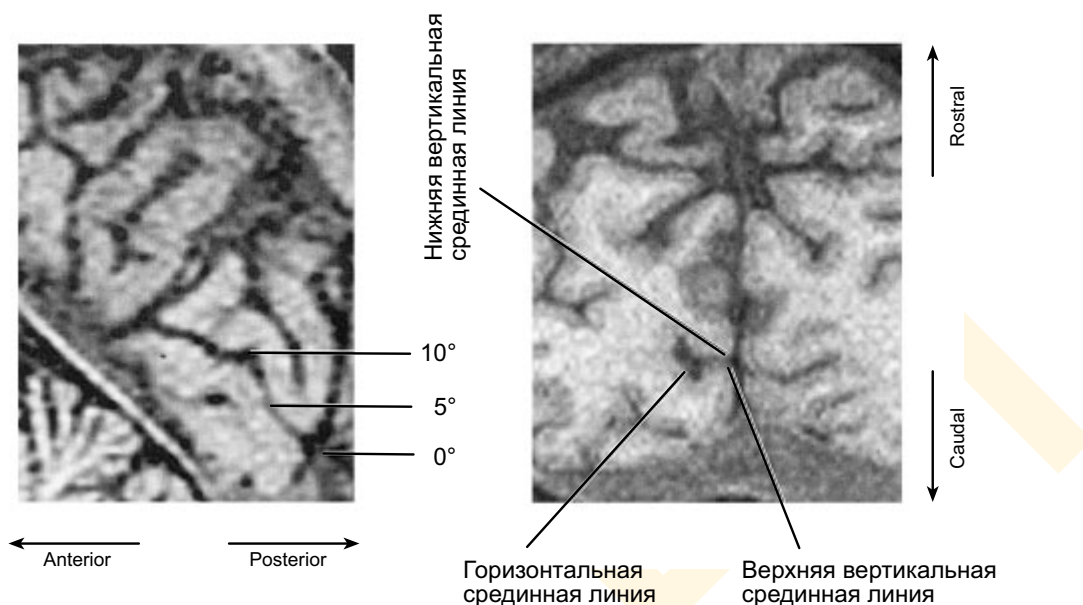




**Рис. 4.1.31** Организация гиперколонок в стриарной коре. По одному направлению приоритетная ориентация систематически варьирует внутри одной гиперколонки. По другому направлению чередуются пласти окулярных доминант. Пласти окулярных доминант по правому глазу выделены серым. Отметим, что колонки ориентации не проходят сквозь слой IV.

Поскольку нейронная организация шпорной борозды ретинотопична, то, каждое кольцо вызывает движущуюся волну нейронной активности, начинающуюся в задней части sulcus calcarinus и катящуюся к ее переду. Ход этой волны мы можем отследить, измерив ФЯМР-сигнал (см. раздел 2.2.3.4) в разных участках шпорной борозды: на рис. 4.1.33 В дано изображение мозга в плоскости шпорной борозды; позиции внутри борозды отмечены черным; ФЯМР-сигнал в каждой точке борозды как функция от времени дан на сетчатом графике рис. 4.1.33 С. Отметим, что амплитуда ФЯМР-сигнала меняется параллельно изменениям фотометрических яркостей в стимуле: сигнал растет и падает четырежды в течение четырех периодов расширения кольца. Временная фаза ФЯМР-сигнала систематически меняется от задней части борозды к передней, при этом задняя часть активна дольше, чем передняя. Волновым движение активности является из-за того, что стимул вначале активизирует нейроны задней части шпорной борозды, а затем передней.

К сказанному стоит добавить, что существует ряд иных оценок мэппинга ретинального изображения в шпорную борозду — рис. 4.1.34: данные ФЯМР-измерений по двум наблюдателям отмечены черными символами; данные по прямой электрической

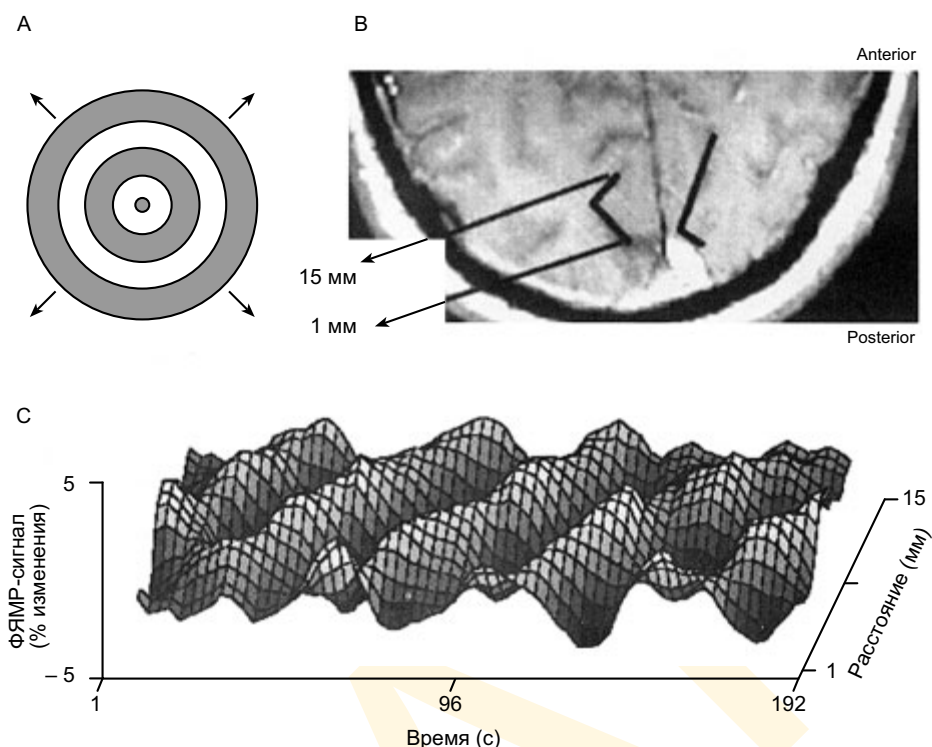


**Рис. 4.1.32** V1 головного мозга человека расположена большей частью в шпорной борозде и в некоторых случаях может выходить за ее пределы. А — в боковой проекции шпорная борозда имеет длину порядка 4 см. В разных положениях продемонстрирована эксцентричность рецептивных полей. В — в коронарной проекции шпорная борозда выглядит как впадина в срединной стенке мозга. На определенном протяжении вдоль борозды рецептивные поля нейронов располагаются по полуокружности поля зрения (каждое полушарие представляет одну его половину). Нейроны с рецептивными полями постоянной эксцентричности в верхней, срединной и нижней частях поля зрения обнаруживаются соответственно в нижнем, среднем и верхнем участках шпорной борозды.

стимуляции коры — серыми квадратиками (Dobelle et al., 1979). В обсуждаемых экспериментах мозг наблюдателей — волонтеров вначале стимулировался, после чего волонтер говорил, в какой точке поля зрения был предъявлен стимул (см. также Brindley & Lewin, 1968): три серых ромбика демонстрируют данные, полученные с помощью ПЭТ, и представляющие усредненные величины по пяти разным наблюдателям и нормированные к общему размеру мозга; серая линия демонстрирует результат, полученный в 1991 г. американцами Джонатаном Хортоном и Вильямом Хойтом (Калифорнийский университет, Сан-Франциско) при изучении позиции слепого пятна у наблюдателей с локальными поражениями мозга, а также рассчитанный экстраполяцией данных, полученных при исследовании обезьян. Мы видим, что данные эти четко согласуются между собой и демонстрируют то, что ощутимо большая корковая область отведена фовеальному представлению, нежели периферическому.

То, что информации, поступающей из центральной ямки, в коре отведено ощутимо больше места, чем, поступающей с периферии сетчатки, выглядит вполне естественным следствием того, что большее число фоторецепторов и ганглиозных клеток представлено в центральной ямке, нежели на периферии сетчатки.

В 1990 г., упоминавшийся уже ранее Хайнц Вессле (см. также Schein, 1988), показал, что расширенное фовеальное представление можно объяснить, предположив, что каждой ганглиозной клетке отведен примерно один и тот же объем представления в коре.



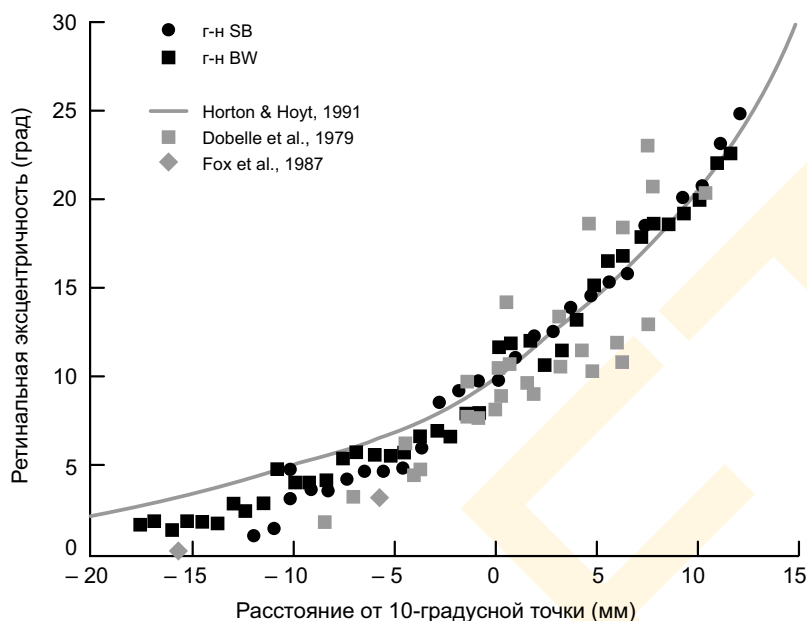
**Рис. 4.1.33** Расположение рецептивных полей нейронов шпорной борозды, диагностированное методом ФЯМР. А — наблюдателю на сером фоне предъявляется серия concentрических расширяющихся колец, в каждом из которых мигает высококонтрастный шапечный паттерн. Когда одно кольцо исчезает за границами экрана, в центре появляется новое, задавая цикл. Всего в одной сессии измерений использовалось четыре цикла. В — изображение в плоскости шпорной борозды. Темные линии — это совокупность точек, расположенных внутри шпорной борозды. С — временной ФЯМР-сигнал в разных точках левой шпорной борозды, следующий строго по направлению стимуляции.

По материалам Engel, 1994.

Позднее — в 1993 г. — британские исследователи Поль Аззопарди и Алан Кови показали, что имеет место дальнейшая экспансия фовеального представления и что ганглиозным клеткам, принимающим сигнал, от рецепторов центральной ямки, отведено в коре в 3-6 раз больше места, чем периферическим ганглионарам.

Итак, нейронные слои, образующие стриарную кору каждого из полушарий мозга, организованы согласно ретинотопическому принципу — принципу двух ретинотопических карт, каждая из которых отвечает в основном за одну половину поля зрения. При этом мы говорим, что ретинотопическая карта хранит *ретинальную топографию*, то есть что соседние области сетчатки также окажутся соседними и в коре. Метрические свойства такой карты сильно искажены и подобны резине, растянутой в одном из направлений. При этом проекции соседних областей сетчатки все равно остаются соседними.

Дисторсия же (первичная) возникает из-за т.н. *кортикального увеличения* центральных областей относительно периферийных. Сие увеличение (см. рис. 1.3.30) наглядно демонстрируют результаты некоего весьма изящного эксперимента, выпол-



**Рис. 4.1.34** Методы диагностики локализации рецептивного поля нейрона в шпорной борозде. Черные символы — данные по двум наблюдателям, полученные с помощью ФЯМР (Engel et al., 1994). Серые квадратики — данные, полученные с помощью микроstimуляции слепого волонтера (Dobelle et al., 1979). Серые ромбики — усредненные данные, полученные с помощью позитронно-эмиссионной томографии (ПЭТ) по пяти наблюдателям (Fox et al., 1987). Серая кривая построена на основе данных, полученных при исследовании зон слепоты у больных с разного травмами и данных, полученных при экстраклеточной регистрации активности единичных нейронов у приматов разных видов (Horton & Hoyt, 1991).

По материалам Engel et al, 1994.

ненного в 80-е годы прошлого века Тутеллем, Сильверманом, Свиткесом и Де Валоис (1982).

На рис. 4.1.35 показана автораддиограмма ретинотопической карты в V1 и дисторсия, вызванная корковым увеличением. Получено изображение было следующим образом: обезьянке показывали пестрый колесообразный паттерн, состоящий из беспорядочных черных и белых прямоугольничков (рис. 4.1.35 А). «Колесо» ритмично вспыхивало и в эти моменты радиоактивный сахар потреблялся активированными клетками. Затем животное молниеносно умервляли, кору его головного мозга расправляли, выполняли с нее тончайшие срезы параллельные ее наружной поверхности и помещали их на пленку, чувствительную к рентгеновским лучам. На рентгенограммах, полученных с коры одного полушария — которое несло информацию только об одной половине «колеса» — четко просматривались все пять «спиц» и все три дуги, сохранившие пространственные (топографические) отношения друг с другом (рис. 4.1.35 В).

Отметим однако, что логарифмически распределенные дуги в стимуле оказались равноотстоящими друг от друга в своих корковых проекциях. Сие указывает на то, что, как уже говорилось выше, весьма небольшой области сетчатки возле ее центра принадлежит непропорционально большая область в коре (и что периферия представлена много меньше). Частично сие действительно есть следствие плотной паковки фоторецепторов в центральной ямке сетчатки (см. рис. 1.3.12), частично — результат

наиболее тщательного коркового анализа этого небольшого, но очень важного участка поля зрения.

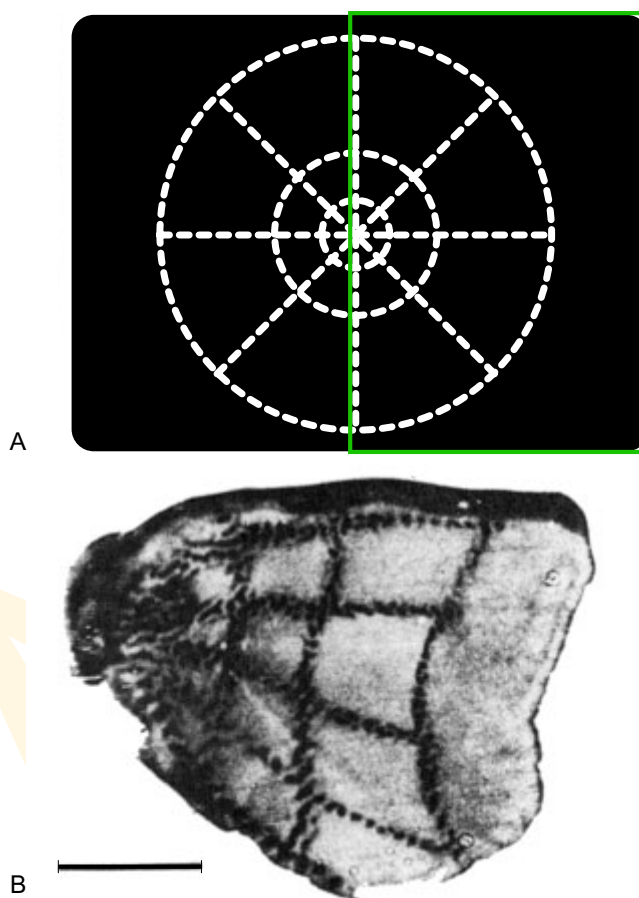
#### 4.1.3.5 Электростимуляция V1 человека

Прямая электростимуляция зрительной коры вызывает зрительное ощущение. Образ, спровоцированный не оптической стимуляцией, а, скажем, надавливанием на глазное яблоко или стимуляцией электрической, называют *зрительным фосфеном*.

При разработке протезов для больных, страдающих неизлечимыми заболеваниями сетчатки, несколько научных групп активно изучали свойства зрительных фосфенов, вызванных электростимуляцией зрительной коры (Brindley & Lewin, 1968; Dobelle et al., 1979; Back et al., 1990).

В 1968 г. британцы Джилс Бриндли и Вальполь Левин (1968) описали эксперименты с волонтером — диабетиком, страдавшей т.н. билатеральной глаукомой с правосторонним отслоением сетчатки и, соответственно, почти полной слепотой по большей части поля зрения и полной слепотой в его центральных 15-ти градусах. С согласия больной Бриндли и Левин имплантировали ей специальный стимулятор, пускавший электрический ток по поверхности мозга неподалеку от первичной V1. Исследователи просили больную описать ее ощущения при стимуляции разными электродами в разных участках зрительной коры. Больная показала, что стимуляция вызывает у нее фосфен в виде яркой точки или пятна, висящего в пространстве. Ее описание зрительного впечатления, вызванного большим числом электродов звучало «как рисовое зерно с вытянутой руки». Иногда один из электродов провоцировал чуть более крупный образ: «как полспички с вытянутой руки».

Как и следовало ожидать от ретинотопической организации зрительной коры, «пространственная позиция» фосфенов в поле зрения менялась, когда менялось пространственное положение электрода. На рис. 4.1.36 показаны позиции электродов и



**Рис. 4.1.35** Авторадиограмма, демонстрирующая ретинотопичность V1 макаки. Обезьяне предъявлялся стимул (А), центрованный по фовеа, в кровоток вводилась радиоактивная 2-дезоксиглюкоза. Картина абсорбции радиоактивного препарата, полученная в V1, показана внизу (В) и воспроизводит правую половину колесообразного стимула.

По материалам Tootel, Silverman, Switkes & De Valois, 1982).



соответствующих фосфенов: как видим, картина результатов вполне ожидаемая — стимуляция электродами в задней части мозга ведет к появлению фосфенов в центре поля зрения (и в рамках 5 градусов зрительного угла), но по мере продвижения электрода к переднему отделу зрительной борозды, локализация фосфенов становится все более эксцентричной. То есть, мы видим, что большая часть зрительной коры отведена под центральные области ретинального сигнала и много меньшая — под периферические.

Бриндли и Левин также выполнили тест на суперпозицию, стимулируя кору отдельными электродами, а затем обоими как одним. Когда электроды далеко отстоят друг от друга, зрительный фосфен, возникший под их одновременным действием, можно спрогнозировать при раздельной стимуляции. Отметим, что принцип суперпозиции работает и в отношении близко стоящих электродов (но не всех).

Легко догадаться, что для получения сложных зрительных паттернов от стимуляции непосредственно V1, необходимо использовать комплекты электродов. При этом, ежели система остается линейной, то мы можем, оценив ощущение от стимуляций единичными электродами, предсказать результат комплексной стимуляции.

Итак, то, что принцип суперпозиции более или менее работает, следует из прогнозируемости визуального результата комплексной электростимуляции на базе результатов стимуляции единичной. Однако без суперпозиции у нас нет логической базы создания изображения из интенсивностей стимула по каждой отдельно взятой точке.

Сравнительно недавно прошли еще несколько докладов о непосредственной стимуляции зрительной коры человека. К примеру, в экспериментах с больными, у которых были удалены эпилептические очаги, американский нейрофизиолог Мартин Бэк с сотрудниками (1990) использовал сверхтонкий электрод (37.5 мкм). Пациенты находились под местной анестезией и поэтому могли говорить о своих зрительных ощущениях. Бэк обнаружил при этом, что когда электроды погружены в толщу зрительной коры, зрительных ощущений можно добиться при очень слабых токах.

Бриндли и Левин (1968) использовали ток силой порядка 2 мА, но Бэк зафиксировал пороги примерно в сто раз меньшие, то есть при токах 0.02 мА. Зрительные фосфены у испытуемых Бэка были вполне устойчивы, и порой даже окрашены. Однако срок экспериментов был крайне ограничен и удалось провести совсем небольшое число манипуляций. Но ученый все же доложил о том, что когда микроэлектроды разнесены более чем на 0.7 мм, фосфены воспринимаются раздельными; при расстоянии между электродами порядка 0.3 мм ощущения сливаются в единую точку. У одного из пациентов почти все фосфены оказались ярко окрашенными, в отличие от тех фосфенов, о которых докладывали пациенты Бриндли и Левина. В момент стимуляции больные могли разумеется, воспринимать и обычные световые стимулы. Фосфены появляются при этом на фоне естественного поля зрения.

#### *4.1.3.6 Рецептивные поля нейронов первичной зрительной коры*

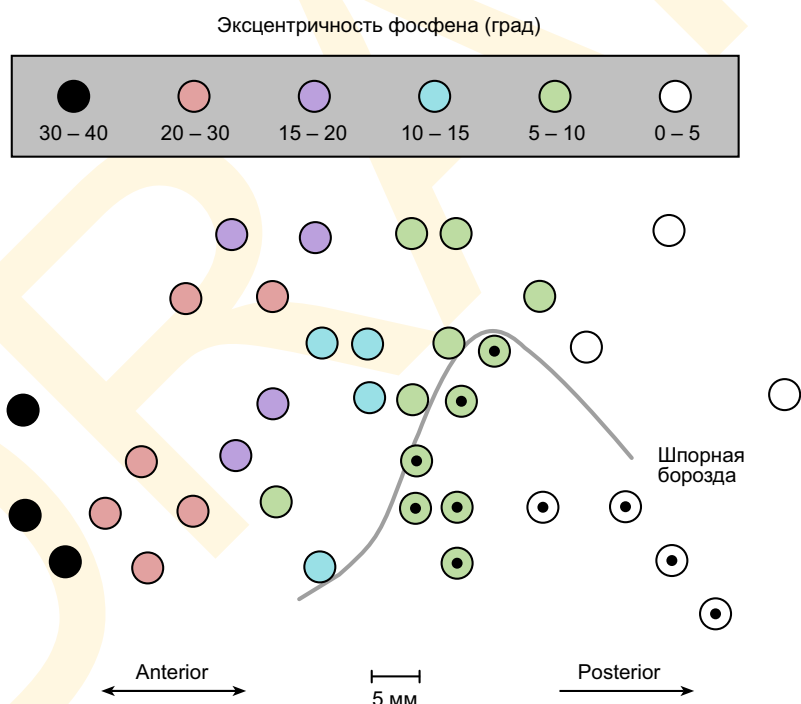
Рецептивные поля нейронов V1 качественно отличны от таковых у коленчатого тела таламуса. К примеру, в отличие от полей коры, форма рецептивных полей ЛКТ циркулярно симметрична. При этом в отличие от нейронов коленчатого тела некоторые нейроны V1 откликаются на стимулы, движущиеся в одном направлении, но «молчат» в ответ на стимулы противоположного направления движения. Некоторые области V1 бинокулярны, то есть откликаются на стимуляцию обоих глаз. Сии новые свойства рецептивных полей должны относиться к зрительным количественным преобразованиям («вычислениям»), выполняемым в недрах коры, в частности таким пре-

образованиям как анализ формы и текстуры, восприятие движения и оценка объема предметов. И у нас есть все основания предполагать, что сие так.

Большая часть из того, что нам известно о корковых рецептивных полях, появилась трудами нобеленосцев Дэвида Хьюбела и Торстена Визеля, проработавших вместе более 25 лет. Эксперименты многих других исследователей натолкнулись на непреодолимые сложности, связанные с записью сигналов от единичных корковых нейронов. Плюс к тому, в самых первых опытах в качестве источника стимуляции использовалось диффузное комнатное освещение, но, как было показано ранее, — контраст стимулирующего паттерна — это критично важная переменная в ретинальном нейронном представлении, и, следовательно, корковые клетки отвечали очень слабо на рассеянное освещение (von Baumgarten & Jung, 1952).

Хьюбел и Визель, применяя высокорелевантные зрению стимулы и дав полученным результатам исключительно проницательную оценку, добились колоссального прогресса в понимании откликов корковых нейронов. Их исследования заложили основу нашего понимания отличий между рецептивными полями коры и коленчатого тела таламуса, и отметим, что методы Хьюбела — Визеля по сей день доминируют в сфере экспериментальной физиологии зрения (Hubel & Wiesel, 1959, 1962, 1968, 1977; Hubel, 1982).

Вспомним, что Хьюбел и Визель фиксировали активность корковых нейронов,



**Рис. 4.1.36** Электрическая стимуляция V1 зрительной системы человека посредством имплантированных микроэлектродов выявляет ретинотопическую организацию зрительной коры. Кружки на диаграмме символизируют позицию микроэлектродов на медиальной (срединной) стенке шпорной борозды. Цвет заливки кружков указывает на эксцентricность фосфена, вызванного электродной стимуляцией. Точка внутри кружков означает, что фосфен возник в верхней части поля зрения. Серая кривая символизирует пространственное расположение шпорной борозды.

По материалам Brindley & Lewin, 1968.

предъявляя сходные стимулы (главным образом отрезки линий и точки) на экране, дававшим проекцию через роговицу и хрусталик на сетчатку животных. По мере продвижения электрода вглубь коры, исследователи предъявляли отрезки переменных длины и ширины. Но, в первую очередь, ученые играли пространственной позицией стимула, отыскивая границы рецептивного поля нейрона. Как только поле устойчиво диагностировалось, исследователи измеряли отклик коркового нейрона на линии, штрихи и точки, предъявляемые индивидуально.

Одной из главных задач их работы являлась систематизация видов корковых нейронов по их откликам на малые наборы стимулов. Исследователи предложили классификацию, представляющую свойства рецептивных полей, а также радикально прояснили функцию нейронов зрительной системы. Напомним при этом, что классификация рецептивных полей не менее важна и в контексте разговора об откликах ганглиозных клеток сетчатки. Плюс к тому, необычайно интересно попытаться понять, возможно ли объединить данные по ретинальным ганглиозным клеткам и клеткам коры, дабы получить ясное представление о функции проводящих путей.

Вторым весьма важным аспектом характеристики корковых нейронов является измерение преобразования контрастного паттерна стимуляции в нейронную активность. Для постановки экспериментов и построения количественных моделей такого преобразования ганглиозными клетками мы применяли метод линейных систем. Линейность — весьма важный принцип, как оказалось, вполне применимый и к корковым клеткам. Действительно, главная область применения линейного принципа в физиологии — это хьюбелово-визелевое деление корковых нейронов на две категории: простые и сложные клетки. Сия классификация основана по большей части на т.н. *ненормированном тесте линейности* (Skottun et al., 1991). Как писал Хьюбел:

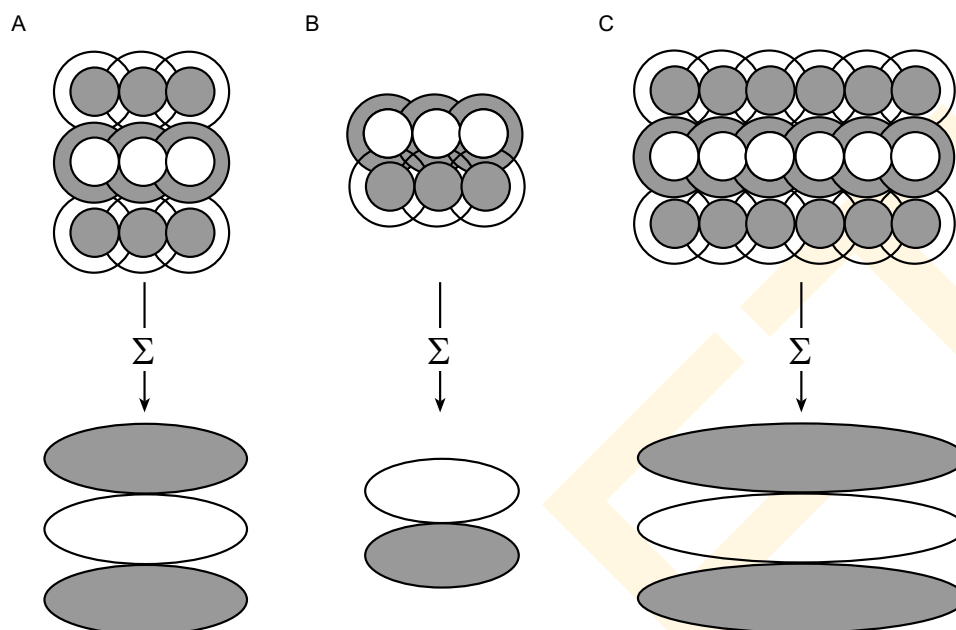
«В целом мы можем прогнозировать отклики простых клеток на сложные стимулы по их откликам на стимулы мелкоточечные» (Hubel, 1988).

С другой стороны, сложные нейроны не отвечают принципу суперпозиции: отклики, полученные прорисью штриха через рецептивное поле нейрона, нельзя достоверно предсказать с помощью откликов на отдельные световые вспышки, образующие этот штрих.

#### 4.1.3.7 Чувствительность к пространственной ориентации стимула

Поскольку простые клетки отвечают на световую стимуляцию зрительной системы более или менее удовлетворяя условиям гомогенности и суперпозиции, их рецептивные поля можно диагностировать, используя описанные ранее линейные методы. Рецептивные поля простых клеток отличны от таковых у ганглиозных клеток сетчатки и клеток коленчатого тела и состоят из соседствующих участков возбуждения и торможения (рис. 4.1.37). Рецептивные поля простых клеток, как говорят, *пространственно ориентированы*, то есть, в большей степени откликаются на стимулы определенной ориентации (при прочих равных оптических характеристиках). Такие рецептивные поля мы назовем *ориентационно-селективными*. Ориентация стимула, вызывающая наибольший отклик, именуется *клеточно приоритетной ориентацией*, или просто — *приоритетной*.

Селективность по ориентации — это новое свойство рецептивных полей, присущее лишь корковым нейронам: вспомним, что рецептивные поля ретинальных нейронов и нейронов коленчатого тела, как говорят, *циркулярно-симметричны* и равновелико откликаются на стимулы любой ориентации, тогда как ориентационно-селективные



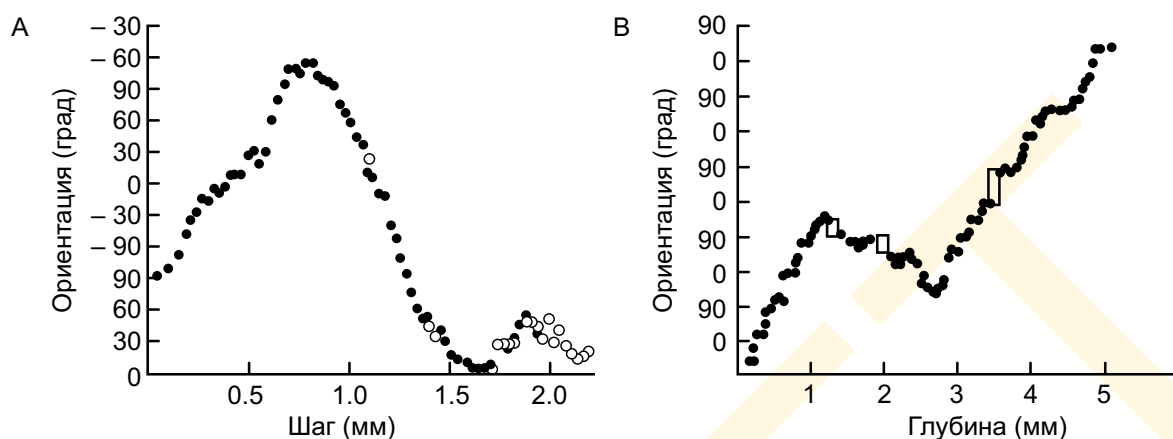
**Рис. 4.1.37** Рецептивные поля селективные по ориентации стимула можно диагностировать, суммируя отклики нейронов с неориентированными, циркулярно-симметричными рецептивными полями. Показаны гипотетические рецептивные поля неких трех нейронов. Каждое из гипотетических рецептивных полей контактирует с соседними участками возбуждения и торможения. Сравнение А и С иллюстрирует, что уровень селективности по ориентации может меняться в зависимости от числа нейронов, объединенных общим направлением.

нейроны коры обнаруживаются на всем протяжении слоев 2 и 3 (но относительно редко на входе слоя 4С).

На рис. 4.1.37 показаны несколько селективных по ориентации линейных рецептивных полей и их организация из выходных сигналов нейронов колленчатого тела<sup>1</sup>. Рецептивные поля простых клеток состоят из соседствующих возбуждающих и тормозящих участков, которые в одном направлении длиннее, чем в другом. Главная ось рецептивного поля определяет приоритетную ориентацию его стимуляции: стимулы, направленные вдоль главной оси, наиболее эффективно возбуждают, либо подавляют нейрон, нежели стимулы прочих ориентаций.

Сравнив три фрагмента рис. 4.1.37, мы увидим, что рецептивные поля, обладающие общей ориентацией стимула, могут отличаться по ряду других параметров: на рис. 4.1.37 А и рис. 4.1.37 В показаны два рецептивных поля одной общей приоритетной ориентацией, но отличных друг от друга по пространственной организации участков возбуждения и торможения; на рис. 4.1.37 А и 4.1.37 С показаны два рецептивных поля одной общей приоритетной ориентации, одинаковой пространственной организации участков возбуждения и торможения, но разной длины самих полей. Понятно,

<sup>1</sup> В принципе, мы можем построить пространственно-ориентированное рецептивное поле из выходных сигналов единичной линейки нейронов колленчатого тела. Но при этом вспомним, что у рецептивных полей нейронов ЛКТ весьма слабое оппозитное окружение. У корковых нейронов участки возбуждения и торможения намного более сбалансированы по данному эффекту.



**Рис. 4.1.38** Приоритетную ориентацию нейронов V1 измерили с помощью т.н. *тангенциальной пенетрации*. По осям абсцисс отложена глубина пенетрации, а по осям ординат — ориентация рецептивного поля. А — данные Hubel & Wiesel (1977). В — данные аналогичного эксперимента, проведенного Hubel & Livingstone (1987). Полыми элементами отмечены локалы, в которых отклик по всем ориентациям равновелик.

что нейроны более длинного рецептивного поля сильнее откликаются на вытянутые стимулы меньшего углового диапазона, нежели нейроны того поля, что короче.

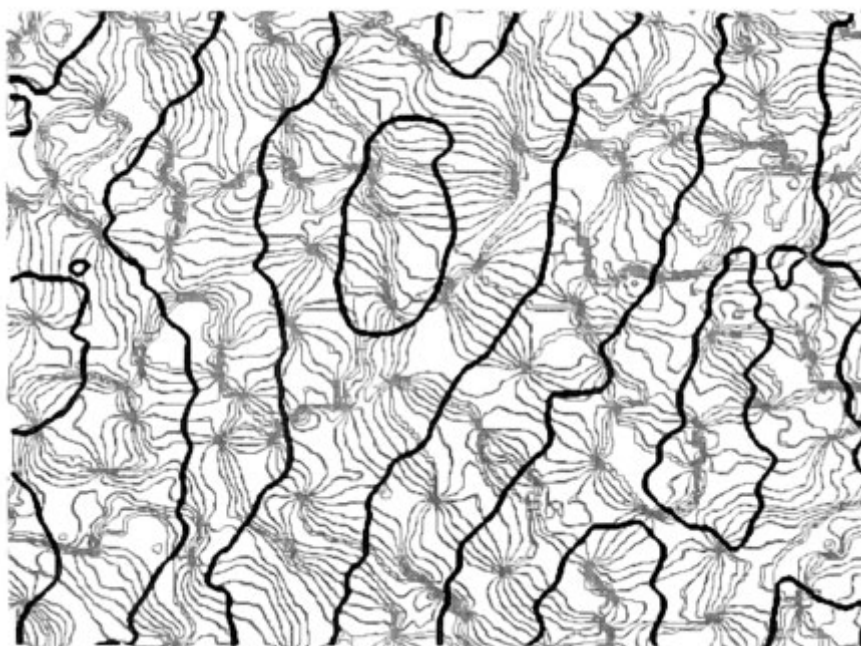
Приоритетная ориентация рецептивных полей варьирует в зависимости от положения нейрона в корковом листке: на рис. 4.1.38 А дана приоритетная ориентация группы нейронов, активность которых измерялась методом т.н. *единичной тангенциальной пенетрации коры*.<sup>1</sup> Мы видим, что в нескольких малых участках слоев 2 и 3 приоритетная ориентация почти одна и та же; по мере движения сквозь кору ориентация меняется систематически, проходя через все углы (Hubel & Wiesel, 1977). Позднее Дэвид Хьюбел и Маргарет Ливингстон (1987) отмечали, что при этих измерениях были выявлены некие постоянные интервалы, в течение которых ориентация рецептивного поля была неоднозначной.

На рис. 4.1.38 В показано второе пенетрирующее измерение, при котором выявились рецептивные поля, не имеющие пространственной ориентации. Хьюбел и Ливингстон обнаружили также, что участки, теряющие ориентационную избирательность совпадают с участками корковых слоев 2 и 3, в которых накапливается фермент цитохромоксидаза. Однако по-прежнему продолжают споры о том, представляют ли эти измерения подлинные отличия в рецептивных полях отдельных нейронов, или же лишь отличия в распределении активности отдельных нейронных групп (Leventhal et al., 1993; O'Keefe et al., 1993).

В 1993 г. американцы Клаус Обермайер и Гарри Бласдел диагностировали участки общей приоритетной ориентации, применив метод оптической визуализации высокого разрешения. Метод строился на применении электрочувствительного красителя, наносимого на кору головного мозга: локальная активность нейронов вызывала изменение спектрального коэффициента отражения красителя, что видно в обычном свете; предъявляя стимулы разных пространственных ориентаций и замеряя изменения в

<sup>1</sup> От лат. «penetratio» — проникновение. — Прим. пер.





**Рис. 4.1.39** Участки общих приоритетов по ориентации стимуляции отмечены серыми линиями. Черные линии — пласты окулярных доминант. На краях пластов окулярных доминант участки общей ориентации объединены в параллельные линии, которые почти перпендикулярны этим пластам. Линии эти сходятся в сингулярных точках, расположенных возле центров пластов окулярных доминант.

По материалам Obermayer & Blasdel, 1993.

коэффициенте отражения, исследователям удалось визуализировать участки общей приоритетной ориентации. В частности, стимулируя левый и правый глаза отдельно, Обермайер и Бласдел визуализировали пласты окулярных доминант (см. также Hubel & Wiesel, 1977).

На рис. 4.1.39 дано контурное представление данных Обермайера — Бласдела (1993): участки общей приоритетной ориентации отмечены серыми изолиниями, а границы пластов окулярных доминант — черными. Иллюстрация демонстрирует нам то, что вариации приоритетной ориентации согласуются с вариациями окулярных доминант. Полный диапазон приоритетных ориентаций имеет место внутри пространства коры, размером около 0.5 мм, что примерно равно ширине одного пласта. Возле краев окулярных пластов изолинии собираются в линейные параллельные стрипы, распрямляющиеся примерно до 1 мм. Стрипы эти ориентированы примерно перпендикулярно краям пластов. В середине пластов изолинии сходятся в одной точке, именуемой *сингулярной*: в этих участках нейроны с рецептивными полями различных приоритетных ориентаций располагаются очень близко друг к другу и, следовательно, в этих местах достаточно высокой может оказаться концентрация цитохромоксидазы (Blasdel, 1992). Метаболическая активность (активность обменных процессов) в этих зонах тоже весьма высока, поскольку активны нейроны по всем возможным ориентациям стимуляции. Это служит объяснением того, почему в участках коры с высокой концентрацией цитохромоксидазы снижается селективность нейронного отклика к ориентации стимулов.

Множество вопросов о характере чувствительности нейронов зрительной коры к ориентации стимула остаются по сей день открытыми. Например, неясно, как характер входного сигнала влияет на свойства рецептивных полей корковых нейронов. На рис. 4.1.37 показано, что мы можем объяснить селективность к ориентации теоретически, поскольку объединение сигналов от нейронов с соседними рецептивными полями класса центр/окружение образуют ориентированное рецептивное поле — но при этом не существует эмпирического подтверждения данному объяснению. Во-вторых, регулярная структура изолиний демонстрирует то, что приоритеты по ориентации у нейронов собраны в высокорегулярный и организованный паттерн.

Каковы же правила образования межсоединений, приводящих к такой удивительной избирательности по ориентации? Какую функциональную роль эти правила играют в перцепционной обработке? Является ли такая пространственная организация исключительной прерогативой нейронных «вычислений» или это типичная схема работы тех областей, чей выходной сигнал предназначен для других модулей пространственной обработки?

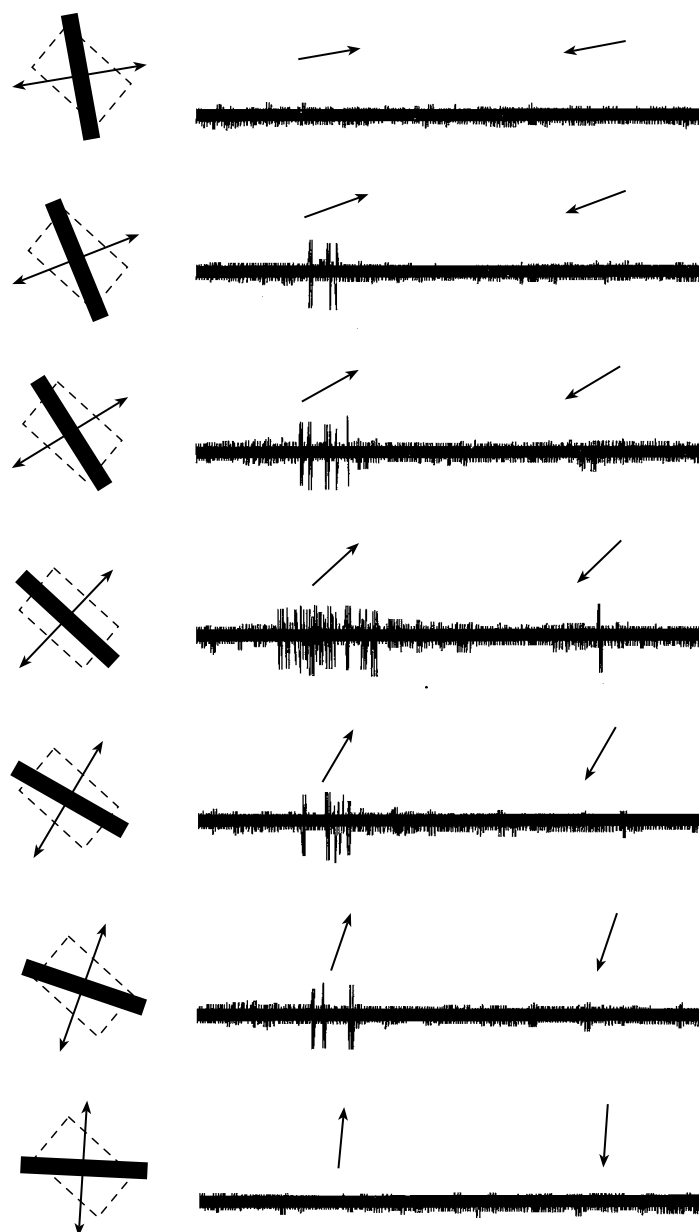
#### 4.1.3.8 Селективность по направлению движения

Хьюбел и Визель (1968) обнаружили еще один вид специализации рецептивных полей нейронов V1: оказалось, что отдельные корковые нейроны у обезьян откликаются на стимулы, движущиеся в определенном направлении и «молчат», когда такой же стимул движется в обратную сторону. Сие есть *селективность по направлению*.

На рис. 1.4.40 дан отклик нейрона V1 макаки на стимуляцию линией, двигавшейся вначале в одном направлении, а затем в противоположном. Клетка, как видим, устойчиво отвечает на стимул определенной пространственной ориентации, но, плюс к тому, является еще и селективной по направлению: когда линия движется вверх и вправо — нейрон дает выраженный отклик; но когда эта же линия движется вниз и влево — дает едва заметный сигнал. Поскольку у данного нейрона частота спонтанной пульсации довольно низка (что типично для большинства корковых нейронов), на основании этих измерений мы не можем четко сказать, что происходит: то ли просто сам нейрон не отвечает, то ли активно заторможен стимулом, движущимся в противоположном направлении.

Нейроны, селективные по направлению стимуляции, обнаруживаются как правило лишь в некоторых определенных слоях коры (и крайне редко в остальных). Основные слои, несущие нейроны данного класса — это 4А, 4В, 4С $\alpha$  и 6 (Hawken et al., 1988). Слои эти получают входной сигнал в основном из магноклеточного пути и посылают свой выходной сигнал в разные отделы мозга. Мы предполагаем при этом, что нейроны в этих слоях могут являться частью зрительного потока, специализированного по передаче информации о движении объектов.

Отклик рецептивных полей, избирательных по направлению движения, может возникнуть от нейронных связей, аналогичных связям, лежащим в основе селективности по ориентации. Функция корковой клетки с рецептивным полем селективным по направлению, может строиться за счет выходных сигналов нейронов ЛКТ с пространственно сдвинутыми рецептивными полями и введением некоторой временной задержки во входной сигнал. Как мы покажем в дальнейшем, результатом объединения пространственных и временных смещений является то, что (когда задержка *суммирует* сигналы) целевой корковый нейрон начинает отвечать на стимулы, движущиеся в одном определенном направлении; мы говорим при этом, что задержка *инвертирует* сигналы, когда стимул начинает двигаться в направлении обратном. Отметим,



**Рис. 4.1.40** Селективность отклика корковых нейронов по направлению движения стимула. В каждом ряду слева показана локализация рецептивного поля, ориентация стимулирующего штриха и два направления движения. Справа даны потенциалы действия, которыми нейроны откликаются на движение стимула по каждому из направлений. Четко видно, что отклик нейрона зависит как от направления движения, так и от пространственной ориентации штриха.

По материалам Hubel & Wiesel, 1993.

что данное объяснение хоть и весьма удобно, но также как и в отношении механизмов селективности по ориентации, точная нейронная схема достижения селективности по направлению корковыми нейронами приматов по сей день остается неизвестной.

### 4.1.3.9 Контрастная чувствительность корковых клеток

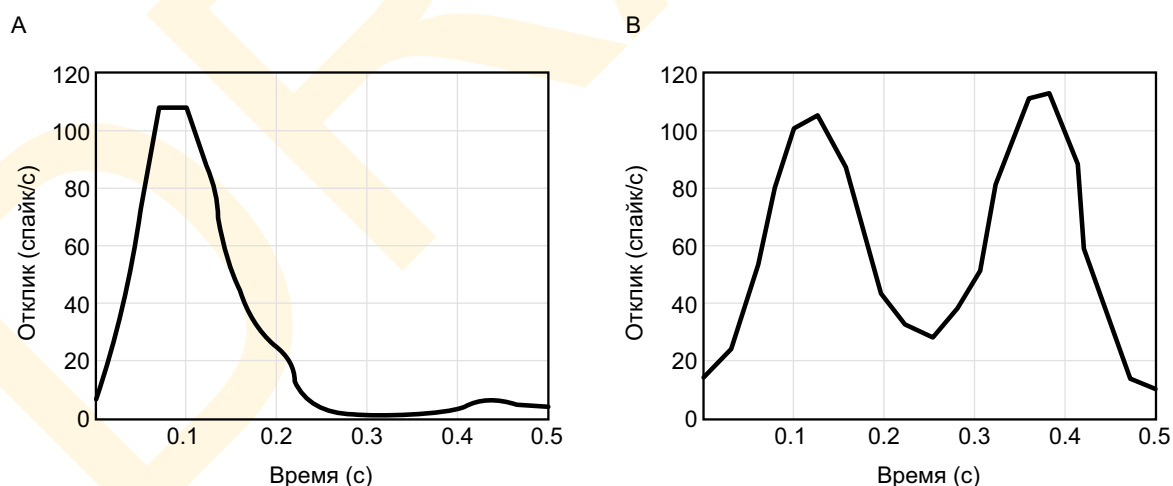
Возможно, что самый прямой и понятный путь классификации простых и сложных корковых клеток основан на их отклике к контрастно-реверсивным синусоидальным паттернам (примеры см. на рис. 4.1.41).

Вспомним, что контрастно-реверсивные паттерны периодичны в пространстве и времени. На рис. 4.1.41 А дана частота пульсации простой клетки, усредненная по множеству повторов в ответ на контрастно-реверсивную стимуляцию с периодом 0.5 с. Там, где простая клетка идеально линейна, вариации частоты пульсации будут синусоидальны и один период отклика окажется равным одному периоду стимуляции. Однако же синусоидальный ритм отклика невозможен, поскольку спонтанная частота разряда нейрона близка к нулю, и, следовательно — частота пульсации не может упасть ниже спонтанной частоты.

Отклик, показанный на рис. 4.1.41, типичен для простых корковых клеток Хьюбела — Визеля, поскольку у многих из них низкий уровень спонтанной активности. Когда сигнал следует только на положительную часть стимулирующей синусоиды и имеет нулевой отклик на ее отрицательную часть, такой отклик называют *выпрямленным по полупериоду* (что свойственно очень многим простым клеткам зрительной коры).

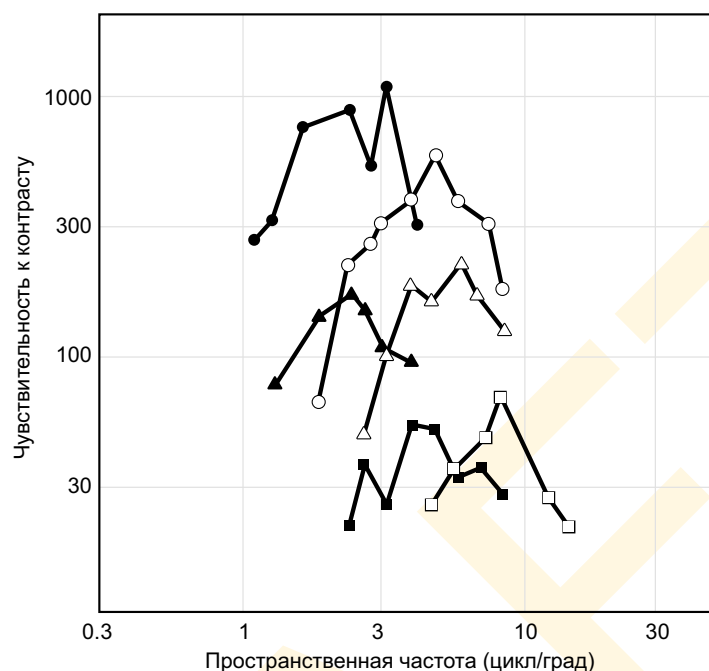
На рис. 4.1.41 В показан средний отклик сложных клеток V1 в течение одного периода стимуляции. В отличие от простых, отклик сложных клеток не синхронизирован с частотой стимуляции и растет в течение обеих фаз мигания контрастного паттерна. Картина отклика этих клеток именуется *выпрямленной по двухполупериоду* и темпоральный компонент такого отклика вдвое превышает временную частоту стимуляции. Сие нелинейное *удвоение частоты* типично для сложных клеток. Отметим при этом, что сложные клетки составляют большинство нейронов зрительной коры.

В 1982 г. Рассел и Карен де Валоис (1982) диагностировали т.н. *пространственную контрастно-сенситивную функцию* клеток Хьюбела — Визеля (рис. 4.1.42): оказалось, что контрастные чувствительности этих нейронов острее, чем таковые у гангли-



**Рис. 4.1.41** Период отклика корковых клеток на контрастно-реверсивный пространственно-частотный паттерн в течение полусекунды. А — отклик простой клетки — это выпрямленная по полупериоду синусоида. В — отклик сложной клетки — это синусоида, выпрямленная по двухполупериоду. Следовательно, темпоральный отклик равен удвоенной частоте стимуляции.

По материалам DeValois, 1982.



**Рис. 4.1.42** Пространственно-частотная селективность шести нейронов в V1 макаки. Отклики были записаны в соседних областях коры, где нейроны еще обладают пространственно-частотной селективностью.

По материалам Hubel & Wiesel, 1993.

озных клеток сетчатки. Более того, даже если бы эти измерения выполнялись по нейронам, расположенным вплотную друг к другу, имеет место ощутимая неоднородность в большинстве эффективных пространственных частот стимула. Сей вариабельности в пространственной настройке не обнаруживается у ретинальных нейронов одного класса. Вариабельность эту могла вызвать новая специализация в коре или же попросту тот факт, что мы до сих пор еще не идентифицировали все типы корковых нейронов V1. В любом случае отличие в положениях пиков контрастно-сенситивных функций ставит вопрос о том, как сигналы от ретинальных нейронов внутри малого участка рекомбинируют свои сигналы, формируя отклик корковых нейронов с такими разными пространственными свойствами рецептивных полей?

В 1978 г. американец Энтони Мовшон с сотрудниками (Movshon, J. A., 1978 a, b), а, спустя почти 9 лет, британцы Дэвид Толхарст и Эндрю Дин (1987) протестировали линейность отклика простых клеток зрительной коры домашней кошки. Приняв во внимание низкий уровень спонтанного сигнала и выполнив выпрямление по полупериоду, они обнаружили, что могут количественно прогнозировать диапазон откликов простых клеток из показателей контрастно-сенситивных функций. Прогнозы работали четко в отношении стимулов с приглушенным или исходно слабым контрастом, то есть по стимулам, вызывавшим отклик меньший половины максимального отклика нейрона. Масштабных тестов на линейность рецептивных полей нейронов коры обезьян не проводилось, но контрастно-сенситивные кривые, вероятно, вполне адекватны задаче прогнозирования откликов простых клеток и в коре обезьян.

На рис. рис. 4.1.42 также даны контрастно-сенситивные функции нелинейных сложных нейронов. Вспомним, что когда система нелинейна, ее отклик на синусои-



дальнюю решетку не является базовым критерием нейронного исполнения, то есть мы не можем использовать его для предсказания отклика на другие стимулы. У нелинейных нейронов контрастно-сенситивная функция определяет отклик клетки на интересующий нас набор стимулов. Такие измерения, в частности, могут помочь понять природу нелинейности. Однако мы говорим, что контрастно-сенситивные функции нелинейной системы — это *неполное* количественное описание рецептивного поля нейрона.

#### 4.1.3.10 Нормировка контраста

Ежели мы станем учитывать низкий уровень спонтанной пульсации нейронов, то выяснится, что отклик простых клеток более или менее линеен, когда речь идет о стимулах невысокого контраста. Однако по мере роста диапазона стимуляции, некоторые важные параметры отклика простых корковых клеток становятся нелинейными. Одно из отклонений, именуемое *нормировкой контраста*, можно продемонстрировать функцией контрастного отклика (сравним с рис. 4.1.10).

На рис. 4.1.43 показана функция контрастного отклика нейрона V1 на четыре различных синусоидальных решетки. Контраст стимула и отклики нейрона даны в логарифмических осях. Сдвиг кривых вправо указывает на то, что нейрон по-разному чувствителен к разным пространственным стимулирующим паттернам. Сей сдвиг таков, какого мы и ожидаем от простой линейной системы, и за которым начинается т.н. *статическая нелинейность*.

**Статическая нелинейность.** Чувствительность фотопигмента можно определить по фототоку, поскольку взаимоотношения между тестовым стимулом и фототоком очень просты (хотя и нелинейны). Процесс поглощения фотонов линеен, но за линейным кодированием следует нелинейный процесс, конвертирующий частоту поглощений фотонов в мембранный фототок. От линейного этапа кодирования свойства нелинейного процесса не зависят и поэтому мы можем назвать последний *статической нелинейностью*. Когда в системе за линейным этапом следует нелинейный, мы можем полностью охарактеризовать структуру такой системы.

Итак, рассмотрим линейный процесс поглощения света фотопигментами. Существует системная матрица, скажем  $A$ , проецирующая тестовый стимул на частоту поглощения фотонов —  $At$ . Во-вторых, статическая нелинейность конвертирует частоту поглощения фотонов в пик фототокового отклика. Вместе оба процесса определяют отклик нелинейной системы —  $F(At)$ .

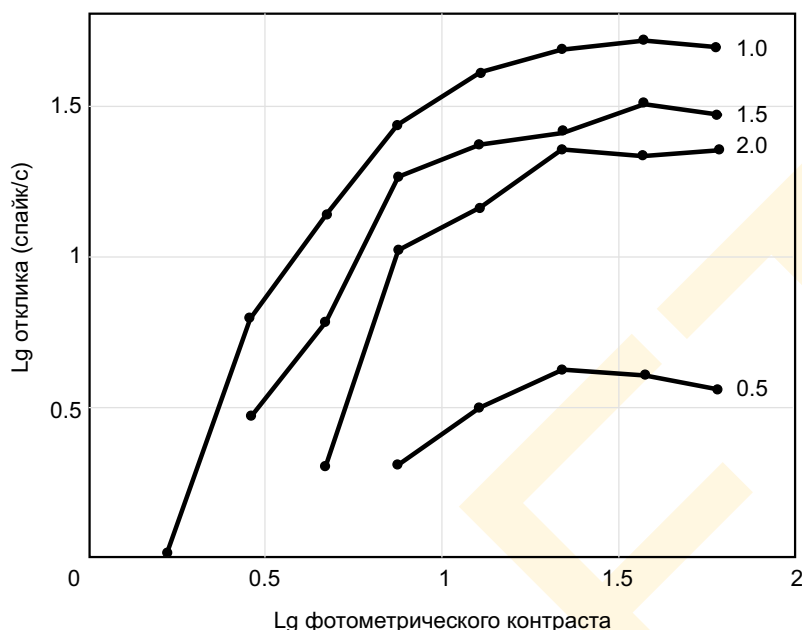
Когда мы установим соответствие между пиком фототока от тестового стимула и стимула уравнивающего, то получим выражение:

$$F(At) = F(aAp),$$

где  $a$  — это интенсивность уравнивающего светового потока, необходимая для достижения визуального равенства с тестовым стимулом. Поскольку нелинейная функция  $F$  монотонна, мы можем сократить ее в обеих частях уравнения и записать:

$$At = aAp$$

Из данного уравнения мы видим, что отношения между интенсивностями тестового и уравнивающего потоков — линейны, поскольку, если  $t$  соответствует  $ap$ , то  $kt$  бу-



**Рис. 4.1.43** Функции контрастного отклика нейронов V1. Каждая кривая демонстрирует результаты измерений, выполненных с применением решеток разной пространственной частоты (справа). Чувствительность нейронов и максимум отклика зависит от пространственной частоты стимуляции.

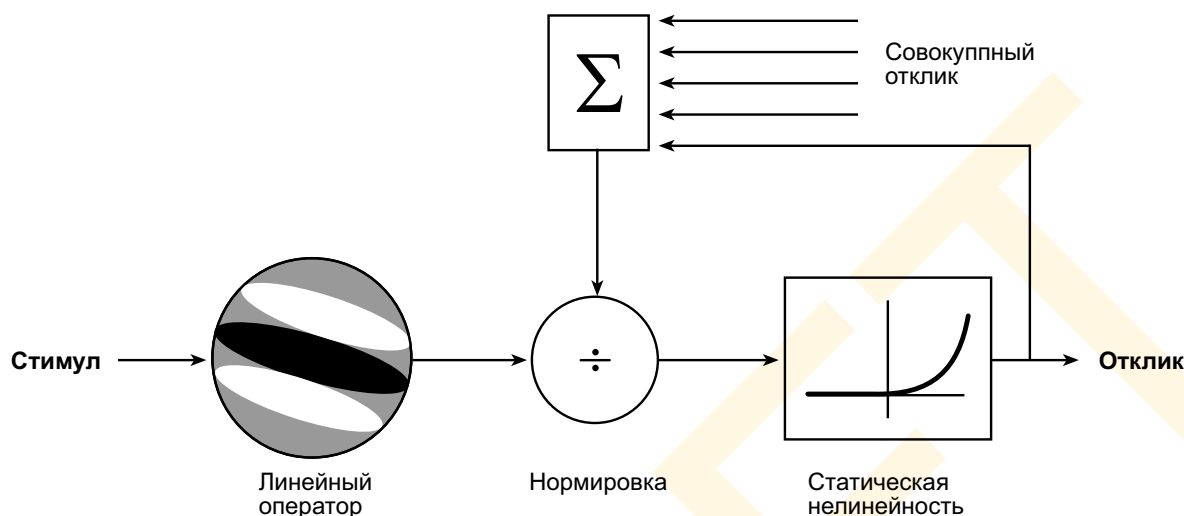
По материалам Albrecht & Hamilton, 1982.

дет соответствовать *кар*. Таким образом, даже если системе свойственна статическая нелинейность, то система эта в эксперименте по уравниванию удовлетворит условию гомогенности. Мы можем показать также, что система эта удовлетворит принципу суперпозиции.

Однако вернемся к графику функций отклика на контрастный синусоидальный стимул: мы видим, что полный набор данных не согласуется с предлагаемой моделью, поскольку уровень насыщения отклика зависит в том числе и от пространственной частоты стимула. Если нелинейность статическая, то уровень насыщения отклика будет одним и тем же вне зависимости от того, какой стимул мы использовали. Поскольку уровень насыщения стимулозависим, он не может основываться на внутренних свойствах нейрона и скорее опосредован каким-то активным процессом (Albrecht & Geisler, 1991; Heeger, 1992).

**Хигерова нелинейная модель.** В 1992 г. профессор Нью-Йоркского университета Дэвид Хигер (Heeger D. J., 1992) предложил модель такого процесса (см. рис. 4.1.44), предполагающую, что отклик одного отдельно взятого нейрона линеен. К линейному сигналу при этом добавляется второй — нелинейный, значение которого зависит от суммарной активности *популяции* нейронов. Однако же отметим, что речь идет не о статической нелинейности, поскольку делитель зависит от контраста стимула.

Хигерова модель толкует данные рис. 4.1.43 следующим образом: во-первых, чувствительность нейрона варьирует вместе с пространственной частотой стимула, поскольку исходно линейное рецептивное поле охотнее откликается на одни стимулы, и слабее на другие (что вызывает смещение графика вдоль логарифмической оси абсцисс); во-вторых, уровень насыщения отклика зависит от соотношения внутренней



**Рис. 4.1.44** Модель нормировки контраста, согласно которой отклик каждого нейрона определен исходным линейным кодированием стимула. Линейный отклик разделен на коэффициент, зависящий от активности популяции нейронов. Совокупный сигнал видоизменен статической нелинейностью.

По материалам Heeger, 1992.

чувствительности нейрона к суммарной чувствительности нейронов в популяции. Уровень насыщения устанавливается с помощью процедуры нормирования: если нейрон относительно нечувствителен к стимулу в сравнении с общей чувствительностью популяции нейронов, то пик отклика нейрона будет подавлен делящим сигналом. Наконец, общая форма функции отклика определена характером следующей затем статической нелинейности.

В чем же биологический смысл нелинейности контрастного отклика? Из данных рис. 4.1.43 мы видим, что пропорция отклика на различные паттерны остается примерно постоянной на разных уровнях контраста. Без нормировки контраста нейронный отклик будет насыщаться на одном и том же уровне, вне зависимости от стимула. В этом случае пропорции отклика на стимулы с различными уровнями контраста будут различными, например: на высоких уровнях контраста все нейроны V1 насытятся и их сигналы окажутся неизменными в ответ на один и тот же входной сигнал. Нормировка же управляет уровнем насыщения, ставя его в зависимость от чувствительности нейрона. При таком раскладе пропорция нейронных откликов остается постоянной в широком диапазоне уровней контраста.

#### 4.1.4 Зрительные потоки в коре

Итак, мы разобрали два основных принципа, характеризующих поток зрительной информации из сетчатки в кору: суть первого в том, что информация эта разделена на потоки, берущие начало в сетчатке и текущие по нейронным проводящим путям в мозг; второго — в том, что рецептивные поля нейронов по ходу в высшие отделы мозга прогрессивно усложняются — рецептивные поля корковых нейронов реагируют на

много более сложно организованные стимулы, нежели поля нейронов сетчатки. Функциональный прогресс рецептивных полей — это ключ к пониманию специализации «вычислительных» процессов, проходящих в зрительной коре.

Однако наши представления о корковой обработке все время расширяются и углубляются, в частности, ожидается, что будут найдены новые зрительные потоки, играющие определенную роль в корковых «вычислениях», и, вероятно также, что обнаружатся новые потоки в самой зрительной коре (и, наверняка выяснится, что некоторые из них, подобно палочковому пути в сетчатке, решают строго специфические задачи и отделены от других потоков). Мы ожидаем также открытия нейронов с новыми свойствами рецептивных полей (и нам предстоит охарактеризовать эти поля, дабы понять их «вычислительную» роль в зрительном акте).

Увы, сегодня наши представления об архитектуре и функции зрительной коры находятся буквально в зачаточном состоянии и потому мы рассмотрим лишь основные принципы организации корковых областей. В частности, мы поговорим о том, как информация из V1 расходится по остальным отделам коры, а также рассмотрим экспериментальные и логические методы нейронауки, имеющие активное отношение к происходящему в коре в контексте зрения.

#### *4.1.4.1 Исход парво- и магноклеточного путей*

Деление зрительной информации на потоки — это очень важный организационный принцип нейронного представления этой информации. Два более или менее хорошо изученных потока — это потоки, текущие по магноклеточному и парвоклеточному проводящим путям, аксоны которых заканчиваются в слоях 4Cα и 4Cβ стриарной коры. Что же происходит с сигналами, поступившими из этих путей, в зрительной коре?

Часть сигналов, поступивших из магноклеточного пути, продолжают движение из V1 в целевые отделы коры, например: магноклеточный путь в слое 4Cα образует синапсы с нейронами слоя 4B, рецептивные поля которых чувствительны к направлению движения стимула; нейроны эти затем посылают непосредственную проекцию в корковую область MT (middle temporal). То есть, можно предположить, что информация, переносимая магноклеточным потоком, особо значима для зрительной обработки в MT.

Как было показано ранее, по магноклеточному проводящему пути течет информация о компонентах ретинального изображения высокой временной частоты. Еще чуть ранее мы говорили о том, что нейроны слоя 4B обладают строгой селективностью по направлению, что свойственно нейронам MT (Zeki, 1974, 1990 a,b). Объединив эти наблюдения, приходим к гипотезе о том, что область MT играет определенную роль в восприятии движения.

Итак, один из рукавов магноклеточного потока течет самостоятельно, но второй конвергирует с парвоклеточным потоком в поверхностных слоях V1.

В 1981 г. американец Джозеф Мальпели (Malpeli J. G., 1981), а спустя время Феррера Тара Нили и Джон Манселл провели исследования, наглядно продемонстрировавшие то, что сигналы из парвоклеточного и магноклеточного потоков, конвергируют на отдельных нейронах. В этих исследованиях парвоклеточные и магноклеточные сигналы блокировались анестезией небольших участков коленчатого тела лидокаином (Malpeli et al., 1981), либо γ-аминобутировой кислотой (Nealy & Maunsell, 1994). По завершении обоих экспериментов исследователи доложили об обнаружении нейронов, на чьи отклики можно было влиять блокадой парво- и магноклеточных путей. Также

были идентифицированы и анатомические пути, по которым текут соответствующие сигналы.

В 1992 г. американец Эдвард Лачика доложил о том, что вводил т.н. *ретроградные анатомические трэйсеры* (вещества — индикаторы, движущиеся в направлении обратном движению импульса) в поверхностные слои зрительной коры. Исследователь пришел к выводу, что нейроны магноклеточного и парвоклеточного путей дают входной сигнал, накладывающимся друг на друга участкам зрительной коры, то есть подтвердилось то, что анатомические пути могут маршрутизировать физиологические сигналы.

Подобно тому, как самостоятельные палочковые пути сливаются в итоге с колбочковыми, сигналы одной из ветвей магноклеточного потока сливаются с парвоклеточными сигналами. Вполне возможно, что целью периферической сегрегации парво- и магноклеточных сигналов является скорейшее объединение определенных видов ретинальной информации в МТ коры. После такого объединения та же информация может использоваться в других корковых областях в комбинации с информацией из парвоклеточных путей.

#### 4.1.4.2 Окраска на цитохромоксидазу

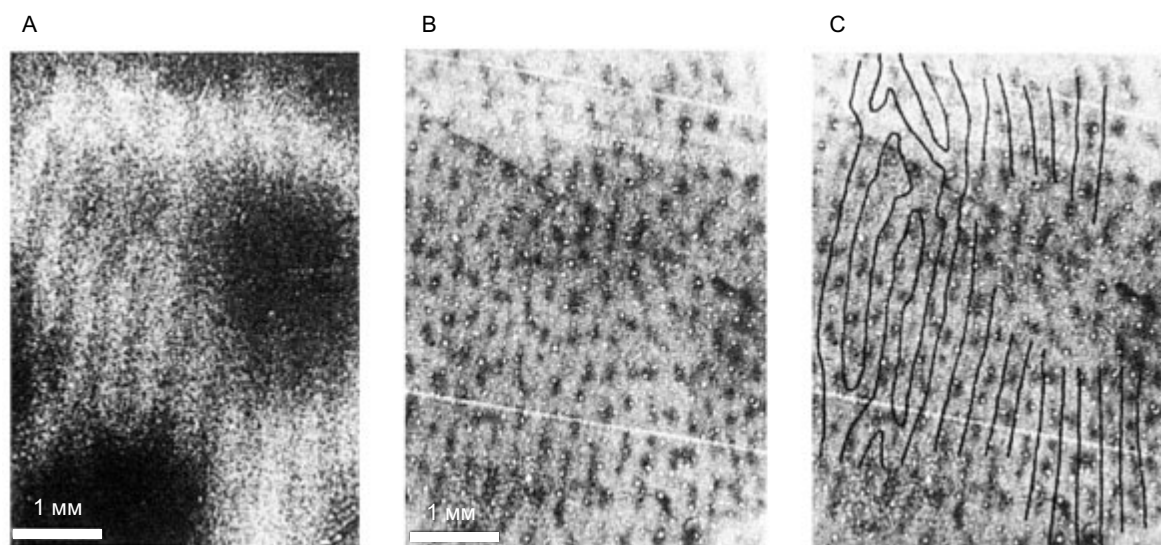
Маргарет Ливингстон и Дэвид Хьюбел (1984 а, b, 1987 а, b, 1988) убедительно показали, что V1 дает исток новым зрительным потокам, начав при этом с апелляции к исследованиям американки Маргарет Вонг-Рили (1979), использовавшей гистохимические маркеры для детекции некоего специфического фермента — *цитохромоксидазы*. Вонг-Рили обнаружила, что фермент этот представлен главным образом в непрерывных и «тяжелых» паттернах слоя 4, и лишь фрагментарно в подлежащих и в поверхностных слоях. Скопления цитохромоксидазы обнаружились микроскопией тангенциальных срезов поверхностных слоев коры и покрашенных на присутствие цитохромоксидазы специальным красителем. Участки с высокой концентрацией цитохромоксидазы выглядели как россыпи черных точек (см. рис. 4.1.45), названных разными исследователями по-разному — «пузыри», «пятна», «угольки» и пр. Пятна эти обычно хорошо видны у приматов (в том числе у человека).

Замечено, что концентрация цитохромоксидазы четко коррелирует с нейронной активностью, несмотря на то, что нет никаких данных за то, что сам по себе фермент играет хоть сколько-нибудь значимую роль в зрительной функции (Wong-Riley, 1979; Humphrey & Hendrickson, 1980; Livingstone & Hubel, 1982; Hendrickson, 1985).

Есть также данные за то, что присутствие цитохромоксидазы коррелирует с наличием отдельных зрительных потоков в коре:

— во-первых, в 1981 г. Джонатан Хортон и Дэвид Хьюбел обнаружили, что картина прокраски на цитохромоксидазу относится как минимум к одному аспекту зрительной функции — окулярным доминантам: на рис. 4.1.45 А показаны пласты окулярных доминант, диагностированные с помощью инъекции радиоактивного трэйсера — меченого тритием пролина; разрез идет через слой 4С (и небольшой участок слоя 5), где пласты окулярных доминант обозначены четче всего. Данные рис. 4.1.45 А — это фактически копия того, что было показано на рис. 4.1.30. На рис. 4.1.45 В при этом показаны точки, четко видимые на тангенциальном срезе через слои 2 и 3, расположенные сразу над участком, показанным на рис. 4.1.45 А. На рис. 4.1.45 С показано наложение пластов окулярных доминант и пятен цитохромоксидазы (изображения были совмещены по общим кровеносным сосудам) и четко видно при этом,





**Рис. 4.1.45** Капли цитохромоксидазы, залегающие в середине пластов окулярных доминант. Три изображения демонстрируют нам тангенциальные срезы V1. А — пласты окулярных доминант в слое 4С, идентифицированные монокулярной инъекцией радиоактивного трэйсера. В — темные пятна — участки с высокой концентрацией цитохромоксидазы в поверхностных слоях коры. С — наложение А и В, демонстрирующее то, что капли залегают в середине пластов окулярных доминант (пространственная привodka изображений выполнена по кровеносным сосудам).

По материалам Horton & Hubel, 1981.

что пятна прокрашенной цитохромоксидазы попадают в центры пластов окулярных доминант;

— во-вторых, доказательство того, что присутствие цитохромоксидазы идентифицирует отличия в зрительных потоках, идет из данных по изучению связей между нейронами V1 — В 1989 г. профессор Вашингтонского университета Андреас Буркхальтер (1989) показал, что нейроны внутри т.н. *цитохромных капель* V1 у человека соединены с другими нейронами в аналогичных каплях, тогда как нейроны между каплями соединены с другими нейронами также между каплями (Rockland & Lund, 1983);

— в-третьих, доказательством являются взаимоотношения между каплями цитохромоксидазы и связями с другими корковыми областями. Также как и в V1 цитохромоксидаза шахматным порядком залегают в соседней корковой зоне — V2. В V2 прокраска на цитохромоксидазу дает регулярный трехполосный паттерн, определяя тем самым район разной концентрации фермента. Три типа полосок поименованы как *толстые, тонкие и вставки*. В толстых и тонких полосках концентрация цитохромоксидазы выше, чем во вставках. Трехполосный паттерн четко различим в V2 мозга человека, а также некоторых видов обезьян (Tootell et al., 1983).

Ливингстон и Хьюбел (1982, 1987 а) продемонстрировали специфичность этих соединений: участки с высокой концентрацией цитохромоксидазы в V1, то есть, зоны зрительной коры с высокой метаболической активностью, образуют связи с аналогичными участками в V2, а именно — нейроны слоя 4В области V1 посылают выходные сигналы в толстые полосы V2; нейроны в каплях посылают сигнал в тонкие полосы; нейроны между каплями посылают импульсы во вставки.

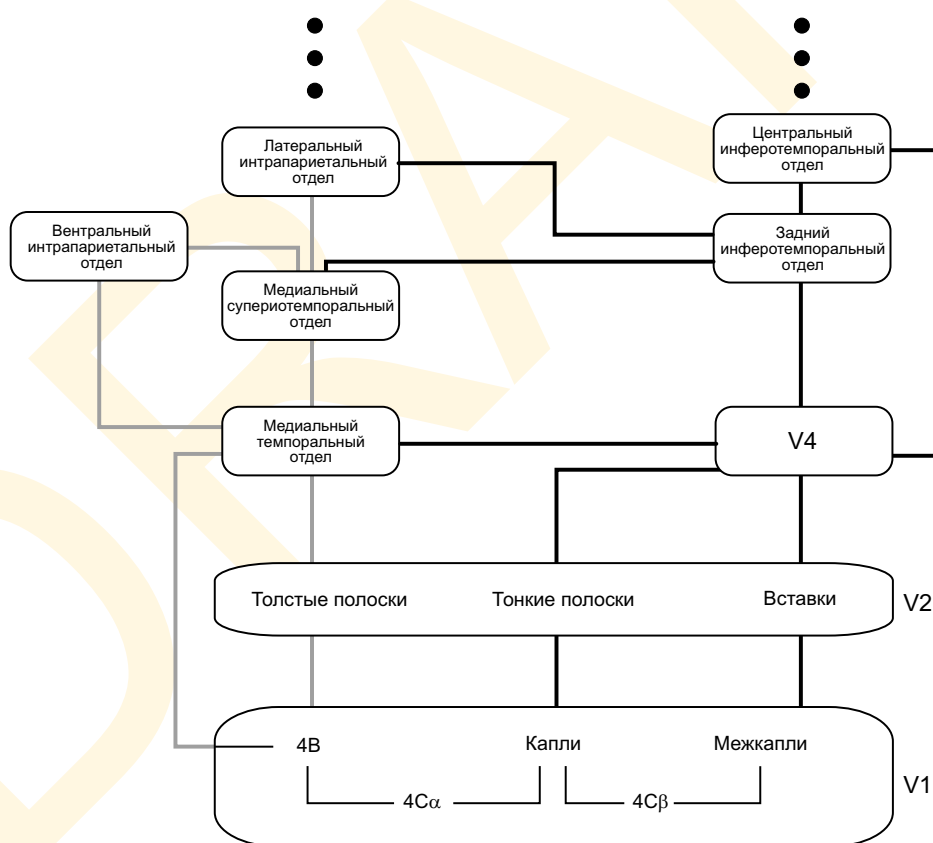
Наконец картина прокраски на цитохромоксидазу в V2 коррелирует с картиной

связей нейронов V2 с другими зрительными областями: у обезьян зрительная область MT получает мощную проекцию из толстых полос V2; область V4 получает сигналы главным образом из тонких полос и вставок (Deyoe & Van Essen, 1988; Merigan and Maunsell, 1993).

#### 4.1.4.3 Зоны первичной зрительной коры

На рис. 4.1.46 показаны некоторые из множества зон зрительной коры. Сразу отметим, что зачастую связи между отделами зрительной коры реципрокальны, к примеру: восходящие аксоны образуют связи с первичным входным слоем — 4С, а аксоны нисходящие образуют соединения со слоями 1 и 6. Единичная зрительная область может иметь связи с несколькими корковыми отделами (Rockland & Pandya, 1979; Felleman & Van Essen, 1991).

Отметим, что на рис. 4.1.46 особо отмечен один из аспектов корковой сегрегации зрительной информации, а именно то, что информация из зрительных областей в затылочной доле делится на два зрительных потока: один посылает выходные сигналы главным образом в заднюю часть теменной доли, а второй образует связи в основном в передней части доли височной.



**Рис. 4.1.46** Общая схема организации зрительных областей. Сигналы из затылочной доли направляются по двум основным потокам: первый уходит в заднюю теменную долю; второй — в нижнюю теменную. Судя по всему сигнал в эти доли идет от разных нейронов затылочной коры.

По материалам Merigan & Maunsell, 1993.

Сегодня известны несколько соединений между этими двумя потоками, но в целом, еще в 1991 г. американец Джоан Бэйзер убедительно показал, что сегрегация ярко выражена: в мозг обезьян исследователь вводил большое количество двух ретроградных трэйсеров: один трэйсер — в заднюю часть теменной доли, а второй — в кору переднего отдела височной. Оба трэйсера со временем обнаружили в затылочной доле, включая V1, V2, V4 и MT. При этом не нашлось ни одного нейрона, содержащего оба вещества. Сие значит, что сигналы от отдельных нейронов объединяются либо в париетальной, либо в темпоральной коре.

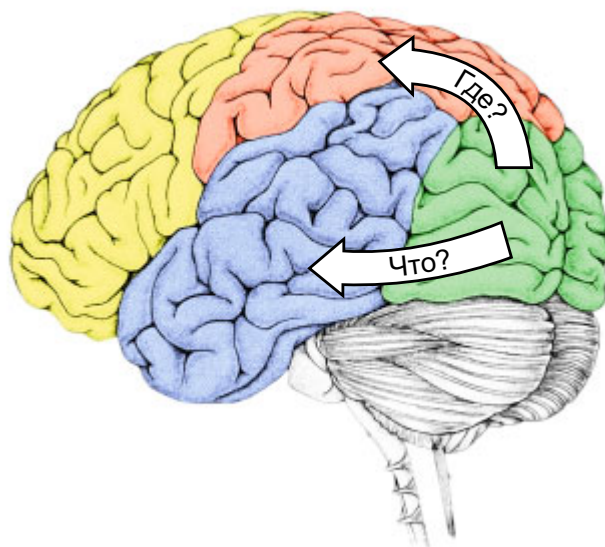
Бэйзер доложил также и о том, что нейроны теменной коры получают информацию в основном от нейронов, рецептивные поля которых лежат возле центральной ямки.

Наблюдения Бэйзера поддержали Лэсли Унгерляйдер и Мортимер Мишкин (1982), предположившие, что париетальные (теменные) и темпоральные (височные) потоки исполняют разные зрительные функции. Исследователи наблюдали, как хирургическое повреждение париетального потока в одном из полушарий вызывает у послеоперационных больных затруднения в зрительной и моторной ориентации, а также обуславливает т.н. *геминеглект* — состояние, при котором больной не реагирует на стимуляцию в той половине поля зрения, что дает проекцию в париетальную кору соответствующего полушария. Такие пациенты также плохо воспринимают местоположение предметов на периферии поля зрения и с трудом дотягиваются до них.

Клинические симптомы, характерные для повреждений темпорального потока весьма разнообразны: в одном из случаев у пациента наблюдалась недостаточность различения и распознавания, а также пострадала зрительная память. Проявления реактивной недостаточности при такой патологии весьма разнообразны, поэтому нейрофизиологи предполагают, что зрительные функции височной и теменной долей головного

мозга также весьма различны, к примеру: темпоральная (височная) система отвечает особи на вопрос «Что?», тогда как париетальная (теменная) на вопрос «Где?» — рис. 4.1.47 (Ungerleider & Mishkin, 1982; Merigan & Maunsell, 1993).

Анатомическое деление нейронных сигналов, идущих от затылочной доли по этим двум потокам, демонстрирует то, что темпоральные и париетальные области получают разнохарактерную информацию о ретинальном изображении. Следовательно, весьма вероятно, что количественные преобразования («вычисления»), проходящие в этих двух отделах мозга, нацелены на решение разных задач (хотя не лишним оказался бы дополнительный и более тонкий анализ этой разницы).



**Рис. 4.1.47** Вентральный (часто именуемый темпоральным или «Что?») и дорсальный (часто именуемый париетальным или «Где?») потоки, исходящие из стриарной коры.

По материалам Mishkin, Ungerleider & Mako, 1983.

### 4.1.5 Корковое представление и восприятие

Выдающийся британский исследователь зрения Вильям Раштон любил эпатировать научное сообщество: как-то раз он потряс нейрофизиологов заявлением, что, мол де «лишь надежда когда-нибудь постичь суть сознания придает хоть какую-то ценность нейронауке». Действительно, почти все работы, упомянутые нами, вдохновлялись надеждой и желанием понять механизмы осознанного восприятия. В этом смысле некоторые нейрофизиологи думают о себе как об исследователях, изучающих проблему «ум — тело». Они формируют вокруг себя сферу знания, которую мы называем *экспериментальной философией*.

Много позже Крик (1993), подхвативший и развивший раштоновскую идею, тоже подкинул каверзный вопрос: «А какие, собственно, компоненты коркового отклика соответствуют осознанным ощущениям?». Поскольку почти все нейрофизиологи склоняются к мысли, что сознание четко коррелирует с корковой активностью, Крик настаивал на том, что идентификация взаимоотношений между сознанием и нейронной активностью — это сугубо экспериментальная тема.

Какие опыты мы можем поставить для ответа на вопрос о взаимосвязи между осознанным ощущением и активностью мозга? Одним из путей поиска ответа на этот вопрос является сравнение информации, обрабатываемой той или иной областью коры с нашими зрительными ощущениями.

Весьма интересен один из экспериментов, рассматривающий осознанное ощущение в сравнении с информацией, подвергаемой обработке в V1. Из анатомии и физиологии V1 нам известно, что в зрительной коре достаточно информации для выявления того, какой глаз стал источником зрительного сигнала. Известно, что внутри слоя 4С полностью отделенные друг от друга группы нейронов, заключенные в пласты окулярных доминант, откликаются на сигналы только одного глаза, но доступна ли нам эта информация?

Ответить на этот вопрос мы можем экспериментально, попросив испытуемых разделить зрительные сигналы, исходящие от правого и левого глаз. Ежели мы точно определим ведущий глаз, то мы можем сделать вывод о том, что информация в V1 — это часть нашего осознанного ощущения. Если не можем, то мы должны признать, что информация из слоя 4С может потеряться до возникновения осознанного ощущения.

Отметим, что сигнал от ведущего глаза всегда доступен бессознательно: еще Гельмгольц (1896) говорил о том, что информация от обоих глаз используется для расчета т.н. *стереопсиса* (бинокулярных данных). Следовательно, вопрос не в том, представлена ли информация, но в том, открыта ли она для осознанного ощущения? Основываясь на собственной интроспекции, Гельмгольц ответил на вопрос отрицательно.

Спустя почти сто лет Оно и Барбайто (1985) провели эксперимент, показавший, что не существует убедительных данных в пользу нашей способности понимать, какой глаз по данному ощущению является ведущим. Следовательно, массив информации, поступающей во входные слои V1 (и ранее) о ведущем глазе не является элементом нашего осознанного ощущения.

Через сей отрицательный результат мы делаем крошечный шагочок в нашем поиске локализации сознания.

#### 4.1.5.1 Функция зрительных областей

Даже несмотря на то, что количественные процессы, прошедшие в зрительной коре, не являются фрагментом нашего осознанного ощущения, мы все равно хотим уз-



нать то, что это за процессы. За последние 25 лет на свет явилось множество гипотез, толкующих перцепционную значимость корковых областей, и главным образом мы наблюдаем поток предположений, касающихся того, что отдельные зрительные области отвечают за «обсчет» специфических перцепционных характеристик, таких как окраска предмета, объем, форма и пр.

Каков же логический и экспериментальный базис рассуждений о перцепционной значимости зрительных областей? Хорэйс Бэрлоу (1972) предложил весьма специфический тезис, соотносящий гистологию с физиологией — нейроны с восприятием — это т.н. *нейронная доктрина*.

Доктрина сия утверждает, что:

⇒ Рецептивное поле нейрона описывает перцепт, вызванный возбуждением этого нейрона.

Мы увидим, что идея сия была не раз выражена как на страницах фундаментальных трудов по физиологии восприятия, так и в отдельных исследованиях, интерпретирующих перцепционную значимость нейронных откликов.

Однако наше сегодняшнее понимание периферического представления не больно-то согласуется с нейронной доктриной. К примеру, принцип Бэрлоу беспомощен при анализе цветовых перцептов: мы хорошо знаем, что выраженный отклик L-колбочек еще не означает, что у наблюдателя возникнет ощущение красного по данному участку поля зрения. Действительно, цветовой перцепт зависит от стимуляции многих соседних участков сетчатки. Условия появления красного цветового перцепта включают в себя паттерн периферических нейронных откликов, в которых преобладают L- нежели M-колбочковые возбуждения. Данные с периферии в целом лучше согласуются с представлением, в котором ощущение зависит от отклика группы нейронов.

Забавно, что неприменимость нейронной доктрины к периферическим сигналам часто используют для поддержки именно самой нейронной доктрины: в конце концов, периферия не относится к сфере нашего осознанного опыта, поэтому неприменимость нейронной концепции вполне ожидаема. Пафос нейронной концепции основан на той идее, что в мозге (возможно в коре) должно существовать специальное место, в котором рецептивное поле нейрона формирует осознанное ощущение, когда нейрон активен. Сие местоположение в мозге должно существовать лишь в той точке зрительного пути, что располагается после участков, выполняющих количественные преобразования, необходимые для восприятия нами тех или иных свойств предметов (окраски, формы, объема и пр.).

Раньше очень любили цитировать работу Хьюбела и Визеля, посвященную V1, как дающую пример локали, к которой целиком относима нейронная концепция: рецептивные поля V1 позиционировались как основное перцепционное свойство этой области, тогда как их ориентация, способность к детекции движения, бинокулярность, а также присутствие сложных клеток первое время рассматривались лишь как дополнение. То есть, в течение определенного (и довольно продолжительного) времени идея рецептивных полей V1, определяющих явные перцепционные признаки, рассматривалась как базовая теория зрительного восприятия.

Однако в 1979 г. была продемонстрирована высокая значимость и других корковых областей (Zeki, 1974, 1978; Felleman & Van Essen, 1991; Van Essen et al., 1992). В своем обзоре зрительных путей Хьюбел и Визель (1979) писали:

«Клетки латерального колленчатого тела посылают аксоны непосредственно в пер-



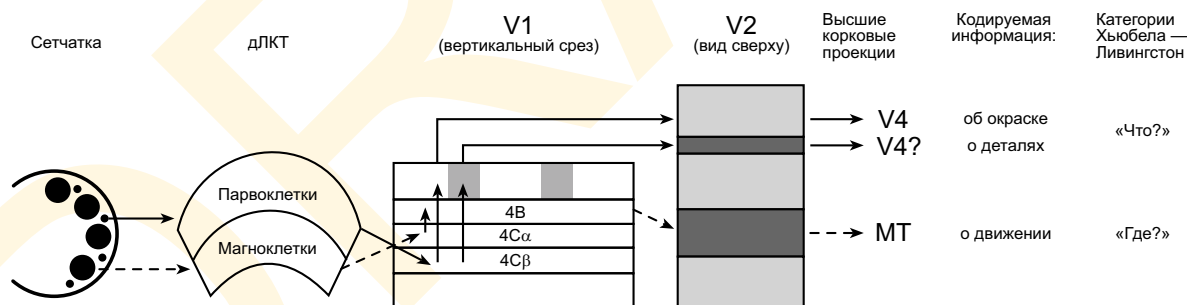
вичную зрительную кору. Оттуда, после каскада синапсов, сигнал идет по нескольким направлениям: в соседние корковые области, а также в глубину мозга. Некоторые клетки даже отправляют проекции назад в ЛКТ, и назначение этой отправки пока неизвестно. Главное во всем этом то, что первичная зрительная кора не является финальной инстанцией в зрительном тракте. Наоборот — происходящее в ней — это лишь один из этапов зрительного акта, причем возможно один из ранних, если рассуждать на определенном уровне абстракции в отношении обрабатываемой информации».

И все-таки, какова функция прочих корковых областей?

Ответ, нужно полагать, базирован на нейронной доктрине. К примеру, когда Семир Зеки (1980, 1983, 1993) обнаружил, что хроматический контраст является специфическим стимулом для V4, он тем самым показал, что данная зона коры участвует в цветовом восприятии. Поскольку движение — это специфический стимул для нейронов MT, мы говорим, что MT — это зона коры, отвечающая за восприятие движения (Dubner & Zeki, 1971). То есть, логика нейронной доктрины позволяет нам интерпретировать свойства рецептивных полей в логике перцепционной функции.

О весьма оригинальном применении нейронной доктрины повествуется в статьях Маргарет Ливингстон и Дэвида Хьюбела (1984 a,b, 1987 a,b, 1988): авторы в целом поддерживают точку зрения Зеки, но добавляют ряд новых гипотез, суть которых представлена на тщательно продуманной анатомо-перцепционной схеме (рис. 4.1.48).

Гипотеза, предложенная Зеки и Ливингстон с Хьюбелом, определила новую точку зрения на функцию коры. С этой точки зрения отношения между корковыми нейронами и восприятием должны строиться на уровне перцепционных свойств. Исследователи не ставили перед собой задачи изучения количественных, «вычислительных» преобразований, происходящих в нейронных потоках, но скорее обобщили свойства ре-



**Рис. 4.1.48** Анатомо-перцепционная схема организации парво- и магноклеточных путей из сетчатки в кору, идентифицирующая зрительные потоки в коре с определенными перцепционными свойствами. Анатомические потоки идентифицированы с применением анатомических маркеров; перцепционные свойства ассоциированы с этими потоками на базе нейронной доктрины.

Аксоны магноклеток, находящихся в дорсальной (задней части) ЛКТ, синаптируют в слое 4Cα стриарной коры. В свою очередь клетки этого слоя посылают аксоны в слой 4B, дающего проекции в широкие темные полосы V2, откуда поток идет в MT-коры. Парвоклетки ДЛКТ дают проекции в слой 4Cβ, откуда сигнал идет в капельные и межкапельные участки коры. Капельная система посылает аксоны в темные тонкие полосы V2, откуда сигнал отправляется в V4. Межкапельная система посылает сигнал в светлые полосы V2, откуда также идет в V4.

По материалам Livingstone & Hubel (1988).

цептивных полей, базировав интерпретацию этих свойств на нейронной доктрине (Hubel & Wiesel, 1977).

Забавно, что применение нейронной доктрины в интерпретации функции мозга весьма распространено, но при этом у нас на руках совсем мало прямых доказательств этой доктрины (Martin, 1992). Собственно главным и единственным «доказательством» является отсутствие внятных альтернатив. Однако же наиболее часто цитируемой альтернативой является предположение о том, что перцепт — это результат активности множества нейронов и что не существует отдельных нейронов, согласованных с тем или иным осознанным перцепционным событием. Такие модели, именуемые *моделями распределенной обработки*, в среде нейрофизиологов не нашли широкого распространения, поскольку не дают надежного инструмента интерпретации измерения активности *единичных* нейронов, тогда как нейронная доктрина дает быстрый и четкий ответ.

Однако сегодня мы склоняемся к толкованию функции мозга с количественных позиций, нежели с позиций мэппинга тех или иных перцептов на зрительные потоки, да еще и диагностированные с весьма сомнительной точностью. При этом мы полагаем истинным то утверждение, что магноклеточный путь несет в себе наилучшее представление высоких временных частот, но в меньшей степени отвечает требованиям пути, проводящем информацию о движении объектов (поскольку информацию, закодированную высокими временными частотами можно использовать при решении многих других зрительных задач).

Фундаментальными мы полагаем два количественных вопроса: «Как?» и «Где?». Каким образом по корковым цепям движутся сигналы высшей обработки, такие как умножение, сложение и синхронизация?; каким образом сохраняются временные данные и каким образом представлены финальные результаты количественных преобразований?; каким образом маршрутизируется информация при ее движении из одной локали мозга другую?

Несмотря на высокую популярность нейронной доктрины (бо являет собой простое средство согласования эффектов восприятия с мозговыми функциями), мы полагаем, что доктрина эта уводит нас от ответа на важнейший вопрос: «Какова количественная подоплека таких перцепционных категорий как, например, окраска, форма, объем?». Даже несмотря на то, что нейронный концепт свидетельствует о прогнозируемости ощущения по характеру рецептивного поля нейрона, моментально появляется второй вопрос: «А откуда тогда берется этот характер?».

Ответы на оба вопроса помогут нам во множестве практических ситуаций: от сенсорного протезирования до робототехники.

Специфические структуры внутри проводящих путей мы рассмотрим как способ воплощения этих принципов, но не как что-то обладающее самостоятельной внутренней ценностью.

Хьюбел и Визель (1979) однажды высказали нечто подобное сей точке зрения:

«Что происходит за пределами зрительной коры и как информация об ориентации используется на дальнейших этапах зрительного акта? Можем ли мы допустить мысль о существовании нейронов, отвечающих на определенные специфические признаки стимула? Наш ответ таков, что мы сильно сомневаемся в существовании таких нейронов, но при этом у нас нет иных идей: чтобы глубже понимать работу мозга, приходится идти окольными путями, в которых больше интуитивного, чем явного.

Было время — и не так давно — когда мы лишь хлопали глазами, глядя на миллио-

ны нейронов в разных слоях коры и страшно радовались, когда кто-то подавал хоть какую-нибудь идею на предмет их функции. Работают ли эти нейроны параллельно (подобно клеткам печени или почек), достигая своих целей просто числом, или же выполняют высокоспециализированные функции? По зрительной коре ответ примерно таков: специфические стимулы либо возбуждают нейрон, либо тормозят его, а вот группы нейронов действительно выполняют специфические преобразования. В целом, есть все основания полагать, что ежели секреты одного из отделов мозга будут раскрыты, то и остальные со временем приподнимут завесу тайны.»

Итак, рецептивное поле клетки определяет ее отклик на паттерн ретиальной стимуляции. Масштабные исследования, в которых применялись самые разнообразные стимулы, показали, что клетки зрительной коры селективно откликаются на весьма ограниченный набор пространственных вариаций: ориентацию, позицию, размер, хроматику и движение стимула. Отклик также зависит от того, какому глазу (правому или левому) был предъявлен этот стимул. Коль скоро все перечисленные факторы оказывают влияние на характер отклика кортикальных нейронов, встает вопрос о том, какова же организация этих клеток? Сие вопрос *архитектуры* стриарной коры: Каково пространственное взаиморасположение клеток и каков паттерн связей между ними, что они нацелены на столь различные количества и характеры стимуляции? Все ли они калейдоскопно перемешаны или напротив — упакованы в регулярные образования по некой системе?

Раньше для ответа на эти вопросы пришлось бы пуститься в долгий марафон экстраклеточных записей: исследователям понадобилось бы вначале отыскать нужную клетку, подвести электрод и записывать сигнал при различных видах стимуляции до тех пор, пока не были бы выявлены локаль, ориентация, размер и доминирующий глаз у рецептивного поля этой клетки: после чего продвинуть электрод дальше — до следующей клетки и повторить исследование. И т. д. По завершении сессии животное пришлось бы умертвить, а его мозг подвергнуть микроскопии, дабы отследить траекторию хода электрода. Критическим моментом такого исследования являлся бы поиск закономерностей в последовательности параметров, характеризующих изученные клетки. Однако кое-какие закономерности все же были выведены экстраклеточным путем, о чем мы поговорим ниже.

Вспомним, что исследования архитектуры коры головного мозга колоссально продвинулись с появлением автордиографических методов. Как было показано в разделе 2.2.3, методы эти основывались на эффекте накопления радиоактивных сахаров в активных клетках и получении благодаря этому визуальной картины пространственного распределения радиоактивности в больших участках стриарной коры. Данные радиогрaфии в целом подтвердили данные экстраклеточных экспериментов, но результат оказался гораздо убедительнее, получен гуманнее и по гораздо большим областям коры.

#### 4.1.5.2 Пространственно-временной анализ: линии и точки

До сих пор в наших попытках диагностики рецептивного поля нейрона мы рассматривали лишь устойчивый отклик, для простоты исключив из анализа временную разность. Теперь настало время вернуть ее на место.

На рис. 4.1.14 показан средний темпоральный отклик ганглиозной клетки сетчатки домашней кошки на световую вспышку. Для получения кривой, показанной на этом рисунке, экспериментаторы предъявляли животному тестовый стимул и записы-

вали результирующую нейронную активность. Если тест вспыхивал регулярно, то результирующая картина спайков с каждым разом оказывалась несколько отличной от предыдущей. Однако отличие было небольшим, что позволило просуммировать все отклики, полученные, скажем, по 50 вспышкам, вычислить среднее число потенциалов действия в течение стимуляции и вычертить кривую. Сия кривая получила название *перистимульной временной гистограммы* (ПСВГ).

Данные на рис. 4.1.14 и прочие экспериментальные измерения показали, что для определенных клеток сетчатки кошек отклик в целом линеен (Enroth-Cugell & Robson, 1984). Проверить линейность в сетчатке приматов экстенсивным путем не удалось, но с определенной долей точности ее подтвердили в измерениях в парвоклеточном проводящем пути внутри ЛКТ приматов (Derrington & Lennie, 1984). Следовательно, опираясь на линейный подход, мы можем измерить отклики на некий комплект базовых стимулов и использовать эти отклики для прогноза откликов на множество прочих стимулов.

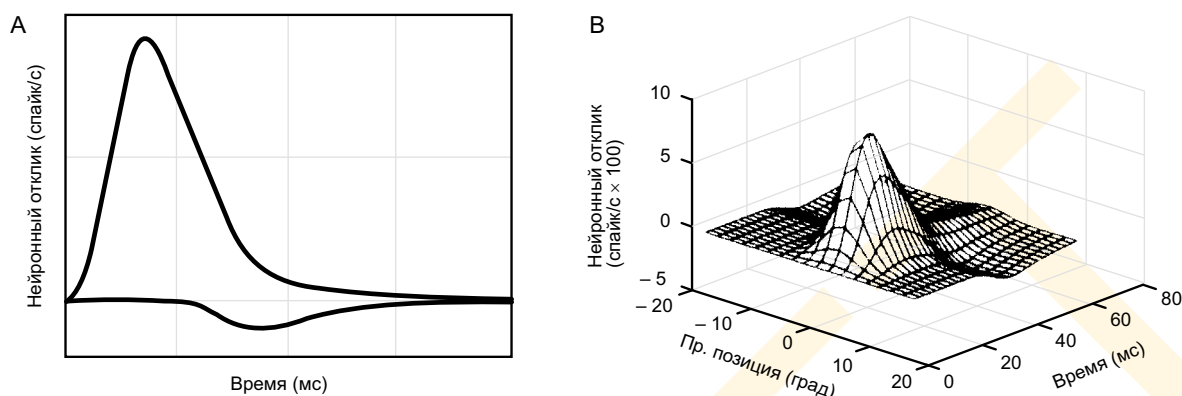
На рис. 4.1.49 показаны две условных перистимульных гистограммы отклика на мигающие линии клеток с рецептивными полями класса оп-центр. Одна из гистограмм демонстрирует условный отклик на мигающую линию, проходящую через центр рецептивного поля, а вторая — на линию, проходящую через его оппозитное окружение. Измеряя отклики на мигающие линии во многих участках рецептивного поля, мы можем точно определить *пространственно-временное* рецептивное поле нейрона (Stevens & Gerstein, 1976) — см. рис. 4.1.49 В: одна из осей — временная размерность; вторая — пространственная позиция тестовой линии. Каждая кривая, идущая вдоль временной оси — это условная перистимульная гистограмма по единичной позиции стимулирующей линии. Когда линия попадает в центр рецептивного поля, условный отклик велик и возникает сразу после стимуляции. Когда стимул ложится на окружение рецептивного поля — отклик слабее, отсрочен и противоположного знака.

Поскольку любой одномерный, но при этом переменный по времени стимул — это сумма множества мерцающих линий, полный набор измерений (подобно стимуляции на рис. 4.1.49 В) позволяет нам прогнозировать отклик на любой аналогичный стимул. Такие измерения определяют пространственно-временное рецептивное поле данного условного нейрона по одной пространственной размерности. Сей метод определения пространственно-временных рецептивных полей был применен главным образом в исследованиях зрительных путей домашней кошки (Stevens & Gerstein, 1976; Palmer & Davis, 1981; Emerson et al., 1992).

#### 4.1.5.3 Пространственно-временные измерения: гармонические функции

Свойства рецептивного поля можно измерить также с помощью т.н. *гармонических функций*, хотя в данном случае мы должны использовать гармонические функции в пространстве и времени. Построить пространственно-временные гармонические функции мы можем перемножением пространственного и временного косинусоидальных контрастных паттернов, создав при этом стимул, именуемый *контрастно-реверсивной решеткой*. Такой паттерн — это произведение пространственной гармонической функции с частотой  $f_x$  цикл/град и временной гармонической с частотой  $f_t$  Гц. Уравнение контрастно-реверсивной решетки, построенной из пространственного и временного косинусоидов выглядит следующим образом:

$$i_{x,t} = [1.0 + a \cos(2\pi f_t t) \cos(2\pi f_x x)] m,$$



**Рис. 4.1.49** Пространственно-временное рецептивное поле линейного нейрона можно диагностировать, измерив его темпоральные отклики на вспыхивающие линии. А — имитация темпорального отклика на мигающие линии, позиционированные либо по центру, либо по периферии рецептивного поля. В — трехмерный график, демонстрирующий имитат темпоральных откликов к отдельным линиям в разных пространственных позициях внутри рецептивного поля. Набор таких измерений можно использовать для прогноза линейного отклика нейронов на любой одномерный переменный по времени стимул. Следовательно, сии измерения дают один из возможных вариантов определения пространственно-временного поля нейрона.

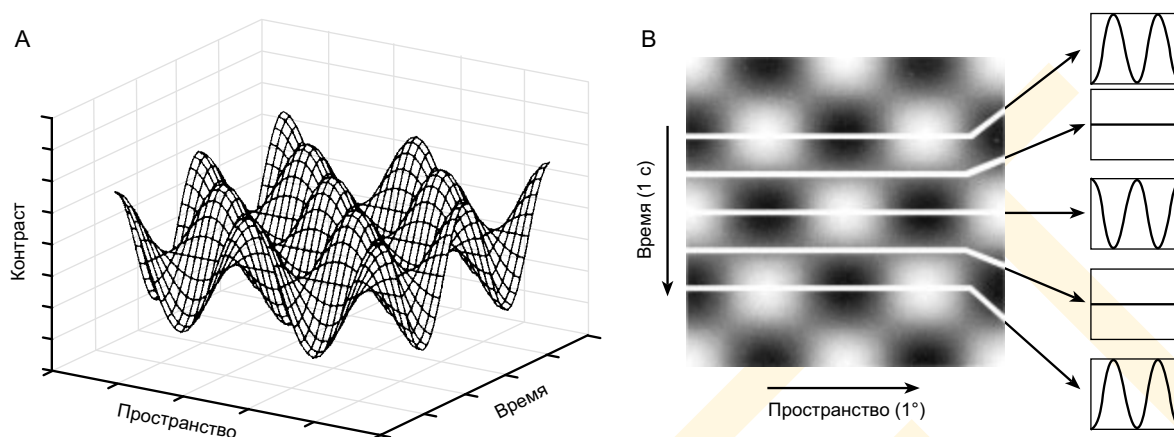
где  $m$  — это средняя фотометрическая яркость стимула и  $a$  — амплитуда модуляции.

На рис. 4.1.50 показана контрастно-реверсивная решетка в двух различных формах. В каждый момент времени решетка эта — это одномерная пространственно-частотная решетка с пространственной частотой  $f_x$ . В каждой точке пространства переменный по времени контраст — это косинусоидальная функция времени с частотой  $f_t$ . На рис. 4.1.50 А функция дана в виде трехмерного графика: по одной из осей отложено время, по другой — пространственный размер. Отдельные линии этого графика — синусоиды. На рис. 4.1.50 В та же функция представлена в виде модуляций яркости: опять одна размерность здесь представляет время, а вторая — пространственный размер — яркие участки соответствуют точкам, где функция имеет положительный контраст, а темные — отрицательный.

Нейроны ЛКТ линейно откликаются на стимулы до 30% контраста, то есть, в диагностике их пространственных рецептивных полей допустимо применять линейные методы. Более того, поскольку рецептивные поля ретинальных нейронов ЛКТ в парвоклеточном пути четносимметричны, мы можем измерить пространственно-временную контрастно-сенситивную функцию, используя лишь косинусоидальные функции по центру рецептивного поля. Поскольку нейроны линейны и по временному параметру шифт-инвариантны, перистимульная гистограмма — это гармоническая функция той же временной частоты, что и временная частота стимула —  $f_t$ . Амплитуда и фаза модуляций перистимульной гистограммы зависят от временной и пространственной частот контрастно-реверсивного паттерна, использованного в качестве стимула.

Отклик на контрастно-реверсивные паттерны парвонейрона ЛКТ макаки показаны на рис. 4.1.51. Этот нейрон откликается на пространственные и временные модуляции стимула в пределах 15 Гц и 15 цикл/град. Форма графика кое-что расскажет нам о реактивности нейрона: во-первых, пониженная чувствительность на низких про-





**Рис. 4.1.50** Пространственно-временное представление контрастно-реверсивной решетки. В каждый момент времени это изображение являет собой косинусоидальный пространственный паттерн — это стимул, амплитуда которого косинусоидально меняется во времени. Контрастно-реверсивный пространственно-временной паттерн представлен в виде трехмерного графика (А) и в виде яркостного изображения (В).

пространственных частотах указывает на то, что рецептивного поля нейрона выражено оппонентное окружение; во-вторых, что нейрон устойчиво откликается на все тестовые временные частоты при низкой пространственной частоте стимуляции, но намного слабее на высоких временных частотах при высоких частотах пространственных. Поскольку отклик на стимулы высокой пространственной частоты — это плод воздействия стимула на оп-центр, данные демонстрируют то, что временная чувствительность у оп-центра и у оппозитного ему окружения могут отличаться.

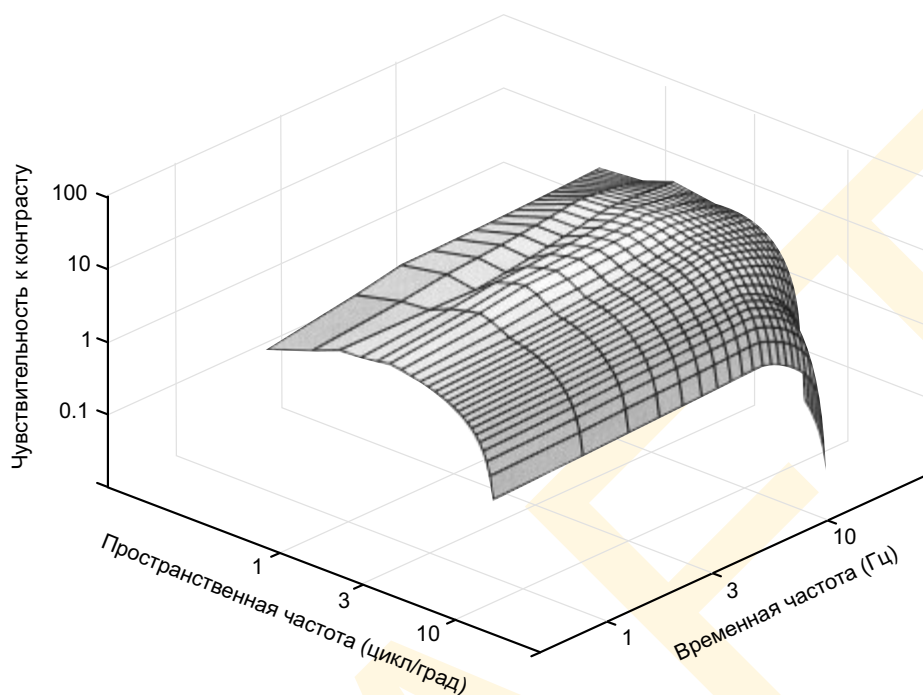
Итак, настало время рассмотреть общий вопрос взаимосвязи и взаимозависимости пространственной и временной чувствительностей.

#### 4.1.5.4 Пространственно-временная разделяемость

Мы диагностировали рецептивное поле нейрона, используя два комплементарных представления: в первом случае — мерцающие линии; во втором — пространственно-временные гармонические функции.

Понятно, что независимо от того, каким описанием мы воспользуемся — чувствительность нейрона остается зависимой как от пространственного размера стимуляции, так и от ее протяженности во времени. Когда наше описание рецептивного поля двухкомпонентно и включает в себя пространственную и временную составляющие, то как мы сумеем выделить пространственное поле? Или временное?

Предположим, что мы говорим о пространственно-временном рецептивном поле некоего нейрона, показанном на рис. 4.1.49 В, используя функцию  $R(x, t)$ . Функция эта определяет отклик на мигающую линию в позиции  $x$  в течение времени  $t$ . Нас интересует, действительно ли можно диагностировать уникальное пространственное рецептивное поле этого нейрона путем исследования этой функции по нескольким различным временным моментам? Мы могли бы зафиксировать время на значениях  $t_1$  и



**Рис. 4.1.51** Отклик на контрастно-реверсивный гармонический стимул нейрона парвовклеточных слоев ЛКТ макаки. По горизонтальным осям отложены пространственная и временная частоты контрастно-реверсивного паттерна. Высота поверхности графика характеризует амплитуду отклика нейронов. Отметим, что снижение временной чувствительности зависит от пространственной частоты стимула: временная чувствительность падает быстрее на высоких пространственных частотах. График построен путем интерполяции данных, предоставленных Деррингтоном и Ленни (1984).

рассмотреть функцию в позиции  $\mathbf{R}(x, t_1)$ , что задает пространственное рецептивное поле в момент  $t_1$ . Мы можем сравнить результат с измерениями во второй момент, скажем —  $\mathbf{R}(x, t_2)$ . Очевидно, что когда у нейрона имеется уникальное пространственное рецептивное поле, две функции оказываются совершенно одинаковыми, например:

$$\mathbf{R}(x, t_1) = a\mathbf{R}(x, t_2),$$

где  $a$  — это скалярная константа.

Если пространственное рецептивное поле в каждый момент времени одно и то же за исключением постоянного шкалирующего коэффициента, то мы говорим, что нейрон обладает «вполне определенным» пространственным рецептивным полем.

Нейрон будет иметь вполне определенное, уникальное пространственное рецептивное поле, если пространственно-временное рецептивное поле является собой *разделяемой* во времени и пространстве функцией.

Разделяемость означает, что функцию рецептивного поля можно записать как произведение двух функций, область определения одной из которых пространственная, а другой — временная, то есть:

$$\mathbf{R}(x, t) = S(x)T(t).$$

Если пространственно-временное рецептивное поле нейрона разделяемо, то, как следует из определения, два пространственных рецептивных поля по двум временам будут всегда соотноситься через шкалирующий фактор  $T(t_1)/T(t_2)$ . В данном случае функция  $S(x)$  даст полноценное определение пространственного рецептивного поля нейрона.

Функция  $T(t)$  — это полноценное определение временного импульсного отклика нейрона. Если пространственно-временное рецептивное поле разделяемо, возможно показать, что пространственно-временная контрастно-сенситивная функция также будет разделяемой.

Как видно по графикам рис. 4.1.49 и рис. 4.1.51, пространственно-временные рецептивные поля нейронов парвоклеточного пути не являются разделяемыми на пространственную и временную составляющие. Следовательно, нейроны эти не имеют самостоятельных пространственных или временных реактивных свойств.<sup>1</sup> Частично по причине этой сложности важно выйти за пределы описаний нейронного отклика, полученного путем линейных измерений и построить модель, прогнозирующую отклик на пространственно-временной паттерн стимуляции.

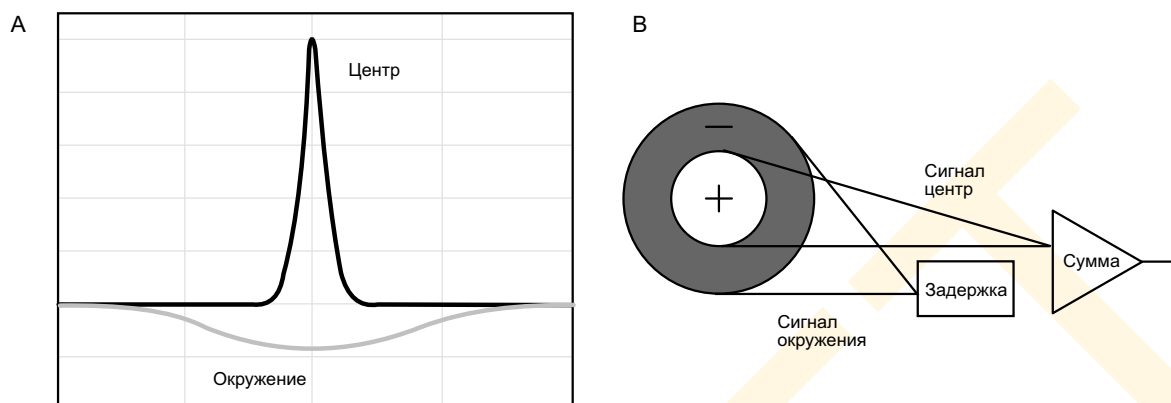
#### 4.1.5.5 Модель гауссовых отличий

В середине 60-х годов прошлого века американцы Роберт Родик и Кристина Энрот-Кугель, а также британец Джон Робсон предложили модель линейного рецептивного поля, послужившую базисом последующих моделей рецептивных полей ганглиозных клеток. В науке о зрении базовая модель Родика — Энрот-Кугель — Робсона имеет большое значение, поскольку идея, в нее заложенная, используется во многих областях, включая работы по психофизике зрительного восприятия пространственных стимулов, в сфере исследования компьютерного зрения применительно к краевой детекции и сегментации изображений. Модель Родика — Энрот-Кугель — Робсона получила название *модель гауссовых отличий* (Rodieck, 1965; Enroth-Cugell & Robson, 1966; Enroth-Cugell et al., 1983).

Модель гауссовых отличий предполагает, что нейронный отклик есть результат объединения сигналов от двух отдельных механизмов, именуемых «центром» и «окружением». Механизм «центр» все свои входные сигналы получает от небольшого центрального участка, а механизм «окружение» — как от центра, так и от области, окружающей его. Предполагается, что и «центр», и «окружение» отвечают на контрастный стимул как система, разделяемая по пространственно-временному критерию. Поскольку каждый из механизмов работает самостоятельно, мы можем описать их как имеющих полноценную пространственную и темпоральную чувствительности. Кривые, характеризующие пространственные чувствительности механизмов «центр» и «окружение» показаны на рис. 4.1.52 А: мы видим, что обе кривые повторяют форму *кривой гауссова распределения*, известную также как *кривая нормального распределения* или *нормальная кривая*.

Согласно модели гауссовых отличий выходной сигнал нейрона можно предсказать суммированием темпорального отклика механизма «центр» с темпоральным откликом механизма «окружение». Модель также предполагает, что темпоральный отклик

<sup>1</sup> Несмотря на то, что мы не углубляемся в детали, рассмотрев связи с колбочками разных типов мы можем убедиться в том, что рецептивные поля не разделяемы по критерию «пространство — длина волны».



**Рис. 4.1.52** Линейная модель пространственно-временного рецептивного поля. Нейронный отклик зависит от линейной суммы двух независимых механизмов: центра и оппозитного ему окружения. А — пространственные чувствительности нейрона по центру рецептивного поля и по его окружению в сумме аппроксимируют гауссову кривую. В — сигналы по центру и по окружению суммируются отдельно. Сигнал окружения имеет временную задержку и добавлен к сигналу по центру.

По материалам Enroth-Cugell et al., 1983.

окружения отличен от такового у центра и в целом медленнее проявляет себя во времени.

Из-за задержки отклика окружения поведение клетки в целом неразделимо по пространственно-временному критерию, даже несмотря на то, что отклики компонент этого механизма вполне разделимы. Следовательно нейронная модель не имеет ни уникального пространственного рецептивного поля, ни уникальной темпоральной функции отклика.

Модель гауссовых отличий весьма практична в науке о зрении минимум по двум причинам: во-первых, она упрощает наше понимание рецептивных полей (и вычисления с этим связанные) и прогнозирует отклик нейронов из расчетов, которые требуют от нас учета лишь нескольких неизвестных (таких, в частности, как ширина центрального и периферийного механизмов рецептивного поля); во-вторых, модель дает метод весьма эффективного описания нейронных пространственно-временных рецептивных полей и поэтому являет собой весьма удобный путь сравнения рецептивных полей нейронов.

Столь же важным представляется и то, что модель Родика — Энрот-Кугель — Робсона создает базу для новых и весьма интересных теоретических построений. В самом деле, идея оппозитных центра и окружения полностью теоретическая и не существует какого бы то ни было прямого доказательства их существования. Построив модель, эффективно описывающую результаты, мы выработаем новую гипотезу структуры нейронного отклика. Пытаясь убедиться в работоспособности теоретических инструментов, а также в факте биологического существования проводящих путей, построенных по принципу антагонизма центра и окружения, — приходим к новым и очень интересным аспектам кодирования световой информации в сетчатке.

Наконец, модель гауссовых отличий детально описана, и, как мы увидим далее, —

грает ключевую роль в описании ряда ограничений чувствительности зрения человека к контрастным паттернам.

Модель также можно использовать для описания отдельных количественных аспектов оптического представления ретинального изображения.

#### 4.1.6 Онтогенез рецептивных полей

Факт наличия в зрительной коре взрослых обезьян нейронов различных типов автоматически поднимает вопрос об их наличии и дифференцировке у новорожденных в противоположность тезису об их развитии в процессе взаимодействия особи с окружающим миром. Сие замечательный пример противостояния нэйтивизма эмпиризму (см. раздел 2.1), но на клеточном (а не органном) уровне. Поскольку биологический базис данного аспекта зрения неплохо изучен, то вполне возможен развернутый ответ на поставленный вопрос.

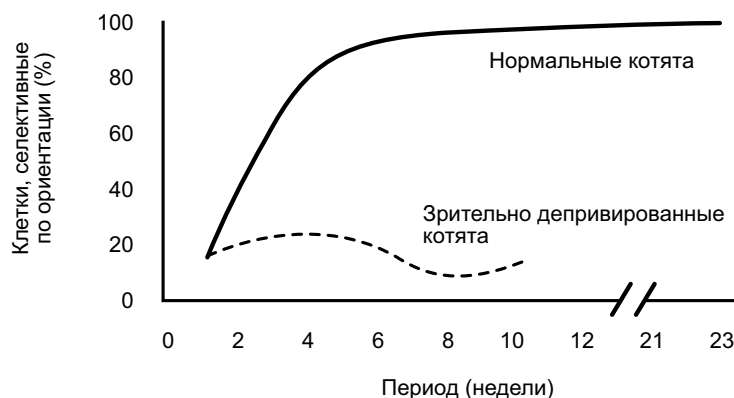
Дэвид Хьюбел и Торстен Визель (1963) исследовали новорожденных котят и обнаружили, что клетки, селективные по ориентации, имеются даже у самых юных животных, что явно указывает на врожденный характер признака. Но также очевидно и то, что после рождения имеет место дальнейшее развитие, поскольку у рецептивных полей новорожденных котят меньшая чем у взрослых животных специализация по ориентации стимула и меньший отклик на стационарные штрихи и края. Главный вопрос заключен в том, благодаря чему происходит дальнейшее развитие: благодаря запрограммированному созреванию или же в результате обучения особи по мере регулярной зрительной стимуляции после рождения? Поиски ответа шли по нескольким направлениям.

В первых экспериментах, котят, выросших без какой-либо зрительной стимуляции, сравнивали с обычными животными. Т.н. *визуальная депривация* (т.е. лишение зрительной информации) достигалась либо путем ушивания века при рождении, либо содержанием животных в полной темноте. Хьюбел и Визель обнаружили, что корковые клетки у ослепленных котят по-прежнему остаются селективными по ориентации стимула, но при этом, в отличие от нормально развивавшихся животных, эти котята оказались почти не способными улавливать ориентацию краев или направление движения.

Более поздние исследования (Blakemore & Van Sluyters, 1975) показали, что по критерию общего количества ориентационно-селективных клеток отличие между светло- и темновыращенными котятами растет ураганно в первые пять недель жизни (рис. 4.1.53). Авторадиографические измерения, построенные на поглощении 2-дезоксиглюкозы, также подтвердили то, что поглощение маркера клетками, селективными по ориентации, растет в первые пять недель жизни (Thomson, Kossut & Blakemore, 1983).

Эти и другие исследования указывают на то, что у зрительной системы имеется некий *критический период*: интервал времени, в течение которого под действием внешней стимуляции идет ее бурное развитие. До и после критического периода обнаруживается существенно меньшее число эмпирических эффектов. Сие значит, что эффекты зрительной депривации в критический период в частности, напрочь подрывают здоровье особи. Нормальное напряжение зрительной функции невозможно из-за недостатка стимуляции, и, следовательно, результирующая зрительная недостаточность некомпенсируема в дальнейшем, поскольку запоздалая стимуляция малоэффективна или неэффективна вовсе.





**Рис. 4.1.53** Динамика развития клеток, селективных по ориентации стимула, у нормальных и зрительно депривированных котят. У нормальных котят (сплошная линия) большинство корковых клеток становится высокоселективными по ориентации между первой и пятой неделями внеутробной жизни. У котят, остававшихся зрительно депривированными до начала эксперимента (пунктирная линия), количество клеток, селективных по ориентации не увеличивается в течение этого периода (Blakemore & Van Sluysters, 1975).

На технику хирургического лечения катаракты у взрослых по началу возлагались большие надежды, в частности на то, что у этих больных восстановится нормальное зрение, позволяющее им жить полноценной жизнью. Действительно, в первое время после операции зрительный опыт этих людей обогащался новыми и необычайно яркими ощущениями, но в целом результат не оправдал ожиданий, — пациенты так и не обрели способности воспринимать мир как нормально видящие люди и уже не могли этому научиться.

Согласно литературным данным онтогенез нормального зрения включает в себя важные компоненты созревания и обучения, что осуществимо только в условиях нормальной стимуляции. Пациенты, чье зрение восстановили в зрелом возрасте, после операции порой впадали в депрессию и некоторые из них даже предпочитали жить как и раньше — в полной темноте, нежели сходить с ума от избытка хаотических зрительных ощущений, ворвавшихся в их жизнь (Gregory, 1970; Von Senden, 1960).

В онтогенезе зрительной системы не существует единого критического периода, охватывающего сразу все свойства зрения, но у разных свойств — свои периоды. Критический период развития чувствительности к ориентации стимула у кошек, к примеру, продолжается, вероятно, с первой по пятую недели жизни (Blakemore, Van Sluysters & Movshon, 1976). Формирование окулярных доминант происходит позднее, где-то между пятой и десятой неделями (Daw & Wyatt, 1976).

В целом критический период по каждому из типов клеток коры зависит от их уровня в зрительной системе: низкоуровневые клетки появляются быстрее, нежели клетки высокого уровня. Сия гипотеза уверенно «садится» на текущие данные, поскольку способность различать стимулы по их ориентации — это свойство клеток входных слоев коры, тогда как окулярные доминанты — это характеристика ее выходных слоев

Сказанное выше полностью касается как животных, так и людей. К примеру, при *катаракте* — помутнении хрусталика — критично важно то, когда возникла патология и когда была излечена: катаракта, возникшая после восьми лет, не дает никаких остаточных эффектов после удаления, поскольку критические периоды развития зрительной системы давно пройдены. Но врожденная катаракта обладает колоссальным деструктивным влиянием на зрение, если хирургически не удалена в первые несколько месяцев жизни (что выяснилось, как только появились первые эффективные методы ее оперативного лечения у младенцев).

(Shatz & Stryker, 1978). Отметим, что такой расклад вполне закономерен, поскольку высокоуровневые клетки могут развить свои реактивные свойства только после того, как таковые разовьются у клеток низкого уровня.

## 4.2 ПСИХОФИЗИЧЕСКИЕ КАНАЛЫ

Воистину один из интереснейших фактов, касаемых оптического этапа пространственного восприятия, является тот факт, что толкование функции клеток, открытых Хьюбелом и Визелем более 40 лет назад по-прежнему остается предметом активных споров: существует относительно небольшое число разночтений относительно точной формы рецептивных полей, а также разные взгляды на то, что важнее в их описании, но все это затмевается неистовыми спорами о функции: Что же эти клетки *делают*, как работают? До сего момента мы рассматривали лишь одну точку зрения, а именно, гипотезу о том, что клетки Хьюбела — Визеля являются детекторами краев и линий. Мы лишь упомянули о том, что существует альтернативная гипотеза, но также не дающая полного объяснения. Теперь настало время подробнее рассмотреть и саму эту теорию, и то, как она возникла из психофизики зрительного восприятия, и какое толкование зрительной обработки в V1 она предлагает.

Второй подход к оптическому этапу зрительного восприятия возник из ветви сенсорной психологии, именуемой *психофизика*. Как мы коротко упомянули в гл. 2 (и подробно рассмотрим в Приложении А), психофизики изучают количественные отношения между осознанными ощущениями (психология) и свойствами физического мира (физика), используя при этом психофизические методы исследования. Термин «психофизический» указывает на то, что психофизики, в отличие от физиологов, не фиксируют электрических событий в нейронах и не занимаются прямыми измерениями прочих видов нейронной активности. Они измеряют определенные параметры *психологической реакции* людей и животных при решении ими разного рода перцепционных задач и пытаются делать выводы о подлежащих механизмах из результатов этих измерений. К примеру, психофизик может интересоваться, насколько фотометрическая яркость пятна должна превысить таковую у его окружения, чтобы наблюдатель уловил едва заметное отличие между ними. Сие именуется *порогом* восприятия отличия.

Психофизик может интересоваться тем как этот порог зависит от таких факторов как размер пятна, фотометрическая яркость окружения или, скажем, время темновой адаптации наблюдателя. В Приложении А мы дадим описание и объяснение стандартных психофизических методов выявления различных порогов восприятия.

Для ответа на вопрос о сенсорных порогах и о том, как пороги эти зависят от прочих переменных, психофизики пытаются понять механизмы, лежащие в основе действий человека. Поскольку сии механизмы в обязательном порядке должны быть реализованы физиологически, должно существовать множество точек пересечения между психофизическим и физиологическим подходами к зрению. Чаще всего такие точки существуют. К примеру, при изучении цветового зрения имеет место устойчивая сходимость взаимоукрепляющих воззрений в этих двух областях знания (см. гл. 3), но в изучении пространственного зрения, такой сходимости гораздо меньше, чем хотелось бы. Действительно, мы увидим, что психофизические теории оптической обработки развивались в совершенно ином направлении, нежели физиологические теории детекции линий и краев.

В течение примерно 30 лет физиологическое сообщество трудится, опираясь на т.н.

*теорию пространственных частот*, доминирующую психофизическую теорию пространственного зрения, поскольку она способна объяснить наибольшее число важных и неожиданных результатов психофизических экспериментов. К сожалению, это довольно сложная и технократическая теория, и поэтому нам придется разобрать большое количество сопутствующего материала, дабы понять ее. Однако как только мы закончим эту работу, нашему взору предстанет весьма оригинальная концепция работы клеток Хьюбела — Визеля.

## 4.2.1 Теория пространственных частот

### 4.2.1.1 Синусоидальные волновые решетки

Как мы уже говорили выше, изучая пространственное зрение человека и животных, мы выбираем примитивные стимулы, которые могут послужить кирпичиками в конструкции стимула комплексного, сложного. К роли таких кирпичиков годны т.н. *синусоидальные волновые решетки*. На рис. 4.2.1, параллельно со своим яркостным профилем, показаны такие решетки. Каждая решетка состоит из чередующихся темных и светлых прутьев. Отметим, что пики яркостного профиля соответствуют светлым прутьям, а впадины — темным. Переход от светлых прутьев к темным мы назовем градуальным — плавным. При этом для полной спецификации синусоидальной решетки необходимо определить ее частоту, контраст, фазу и ориентацию.

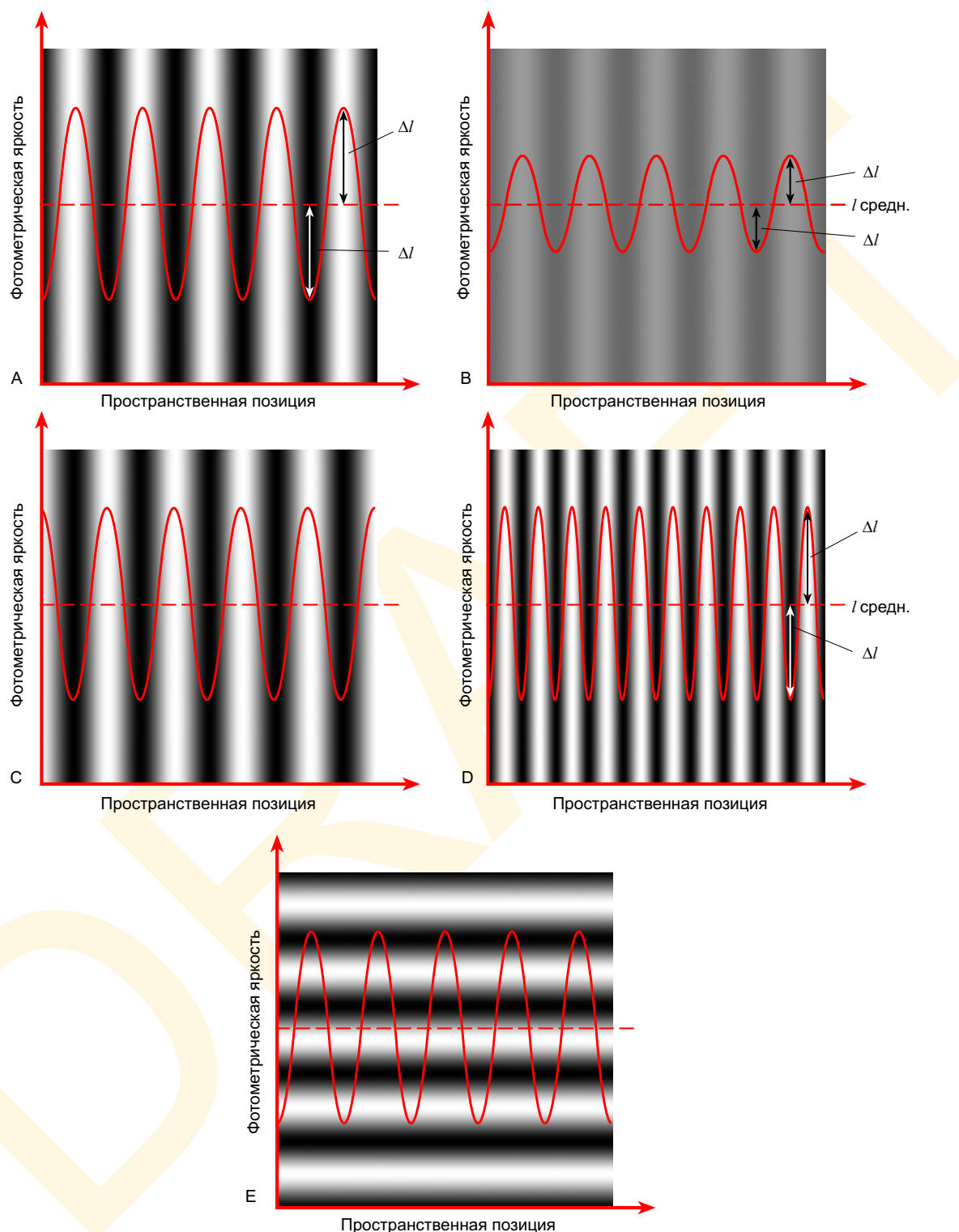
Также как и гипотеза о клетках-детекторах краев и линий теория пространственных частот в контексте оптического этапа зрительного восприятия базируется на том атомистическом предположении, что представление любого ретинального изображения, каким бы сложным оно ни было, есть совокупность множества пространственных примитивов — «атомов». Однако примитивы в теории пространственных частот весьма отличны от краев и линий, рассмотренных нами в предыдущем разделе. Здесь мы говорим о синусоидальных решетках — двумерных пространственно протяженных стимулах, фотометрическая яркость которых синусоидально меняется по одной пространственной размерности, но постоянна по перпендикулярной (рис. 4.2.1). Обратим внимание на то, что темные и светлые штрихи выглядят расфокусированными из-за того, что яркость по мере продвижения по решетке меняется плавно, но не резко (на что указывают плавные волны кривой графика).

Каждый примитив синусоидальной решетки характеризуется четырьмя параметрами: *пространственной частотой, ориентацией, амплитудой и фазой*.

**Частота.** Сравним две решетки во втором ряду рис. 4.2.1: в правой явно больше вариаций, чем в левой, и мы говорим, что пространственная частота правой решетки выше, чем у левой. Т.е., пространственная частота решетки — это показатель, определяющий ширину светлых и темных штрихов — при низкой частоте штрихи шире; при высокой — уже.

Пространственную частоту синусоидальной решетки обычно выражают в количестве циклов на градус зрительного угла (цикл/град), либо в количестве циклов на ту или иную пространственную единицу, например, на сантиметр (цикл/см). Как мы увидим далее, угловой вариант ощутимо практичнее.

**Контраст.** На рис. 4.2.1 В показана решетка низкого контраста параллельно со своим яркостным профилем; слева (рис. 4.2.1 А) дана решетка той же пространственной частоты, но большего контраста. Пунктирной линией обозначена средняя фотометрическая яркость решетки, то есть (среднее арифметическое между верхним и нижним



**Рис. 4.2.1** Синусоидальные решетки и соответствующие им яркостные профили. А — исходная решетка; В — решетка той же частоты, но меньшей амплитуды (меньшего контраста); С — решетка той же частоты но другой фазы; D — решетка большей частоты, но той же амплитуды; E — решетка той же частоты, той же амплитуды, но иной ориентации.

экстремумами). Несмотря на то, что средняя яркость ( $l_{ave}$ ) у обеих решеток одинакова, у нижней большая разница между пиками и средним значением, что указывает на больший контраст решетки.

Итак, контраст можно определить следующим уравнением:

$$contrast = \frac{\Delta l}{l_{ave}}$$

где  $\Delta l$  — отличие между пиковым и средним значениями, а  $l_{ave}$  — средняя фотометрическая яркость решетки (среднее арифметическое между максимальной и минимальной яркостями).

Несмотря на то, что приведенное уравнение весьма удобно в определении понятия «контраст», оно непрактично в измерении контраста. Практичнее замерять максимальное ( $l_{max}$ ) и минимальное ( $l_{min}$ ) значения яркости, и общепринятая формула выглядит так:

$$contrast = \frac{l_{max} - l_{min}}{l_{max} + l_{min}}$$

где:

$$l_{max} = l_{ave} + \Delta l$$

$$l_{min} = l_{ave} - \Delta l$$

Кстати говоря, заменив  $l_{max}$  и  $l_{min}$ , получим:

$$\frac{(l_{ave} + \Delta l) - (l_{ave} - \Delta l)}{(l_{ave} + \Delta l) + (l_{ave} - \Delta l)} = \frac{\Delta l}{l_{ave}}$$

**Фаза и ориентация.** О фазе синусоидальной решетки говорят в контексте ее сравнения с другой решеткой, к примеру: если две решетки одинаковой пространственной частоты *синфазны* — верхние и нижние экстремумы их яркостных профилей совпадут; если решетки находятся в *противофазе* — верхний экстремум одной решетки совпадет с нижним другой решетки.

Иногда о фазе решетки говорят в контексте ее позиции относительно некой эталонной точки. Измеряют фазу в углах, поэтому решетка, чья точка положительного перегиба совпадает с эталонной, имеет фазу  $0^\circ$  (т.н. *синусная фаза*); если в эталонной точке оказался максимум — говорят, что фаза равна  $90^\circ$  (*косинусная фаза*); если в эталонной точке оказалась точка отрицательного перегиба — фаза равна  $180^\circ$  (*антисинусная фаза*); и если на эталоне оказался минимум — фаза равна  $270^\circ$  (*антикосинусная фаза*).

Решетка С на рис. 4.2.1 отличается от решетки А своей фазой.

**Ориентация решетки** — это угол поворота штрихов решетки против часовой стрелки от вертикали.

На рис. 4.2.1 решетка Е отличается от всех остальных горизонтальной ориентацией.

**Анализ Фурье.** Через математический процесс, именуемый анализом Фурье, синусоиды определенной частоты, контраста, фазы и ориентации можно использовать для построения много более сложных стимулов. Хотя сие интуитивно и неочевидно, но любую ахроматическую сцену — дом, лес или лицо человека — можно представить как совокупность синусоид определенной частоты, контраста, фазы и ориентации.

Рассмотрим квадратно-волновую решетку рис. 4.2.2: ее яркостный профиль де-



монстрирует резкий переход от черных прутьев к белым, что часто называют *шаговым изменением* фотометрической яркости. Подобно любому другому пространственно сложному стимулу квадратную волну можно построить, собрав вместе соответствующие синусоиды — рис. 4.2.3. Синусоида той же пространственной частоты, что и у самой квадратной волны, именуется *базовой*. Синусоиды больших частот — *гармониками* (у третьей гармоники частота в три раза больше базовой, а контраст составляет одну треть; у пятой гармоники частота в пять раз больше базовой и в пять раз меньше контраст и т.д.). Добавляя к базовой волне только нечетные гармоники можно создать стимул, визуально неотличимый от квадратно-волнового (добавив бесконечное число нечетных гармоник можно построить даже физически идентичную квадратную волну). Создание квадратной волны и создание, скажем, лица человека — бесконечно далеки друг от друга, но оба принципа, однако, применимы и в первом и во втором случаях.

Итак, мы договорились полагать синусоидальные решетки примитивами (атомарными элементами) пространственного зрения. Однако же мы не оцениваем что-либо увиденное как состоящее из синусоидальных решеток. Ежели бы осознанное восприятие зрительных элементов было необходимым условием придания им статуса зрительных примитивов, то оно, в первую очередь, касалось бы линий и краев, нежели синусоидальных решеток, поскольку линии и края мы осознанно воспринимаем в естественных сценах. Однако нет никаких оснований предполагать, что на ранних этапах пространственного зрения примитивные элементы должны осознаваться. К примеру, мы не осознаем окраску отдельных мелких точек на поверхности предметов, тем не менее вызывающих отклик колбочек всех трех типов, и что несомненно является набором примитивов зрительной системы.

У выбора синусоидальных решеток в качестве примитивов существует прочное теоретическое обоснование, но это, скорее, формальное математическое основание, нежели экспериментальное. Логика рассуждений здесь основана на хорошо известной и широко применяемой математической *теореме Фурье*, названной так в честь выдающегося французского математика и физика Жана Батиста Жозефа Фурье, сформулировавшего эту теорему в 1822 г в работе «Аналитическая теория тепла».

В приложении к вопросу обработки ретинального 2D-изображения:

⇒ *Фурье-анализ* (или как его еще называют *гармонический анализ*) — это основанный на теореме Фурье метод, с помощью которого любое двухмерное оптическое изображение на сетчатке можно рассматривать как совокупность синусоидальных решеток отличных друг от друга по частоте, ориентации, амплитуде и фазе.

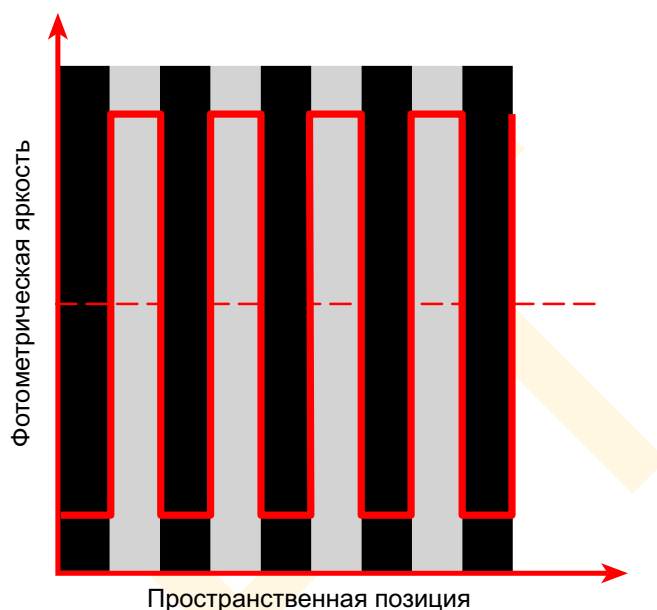
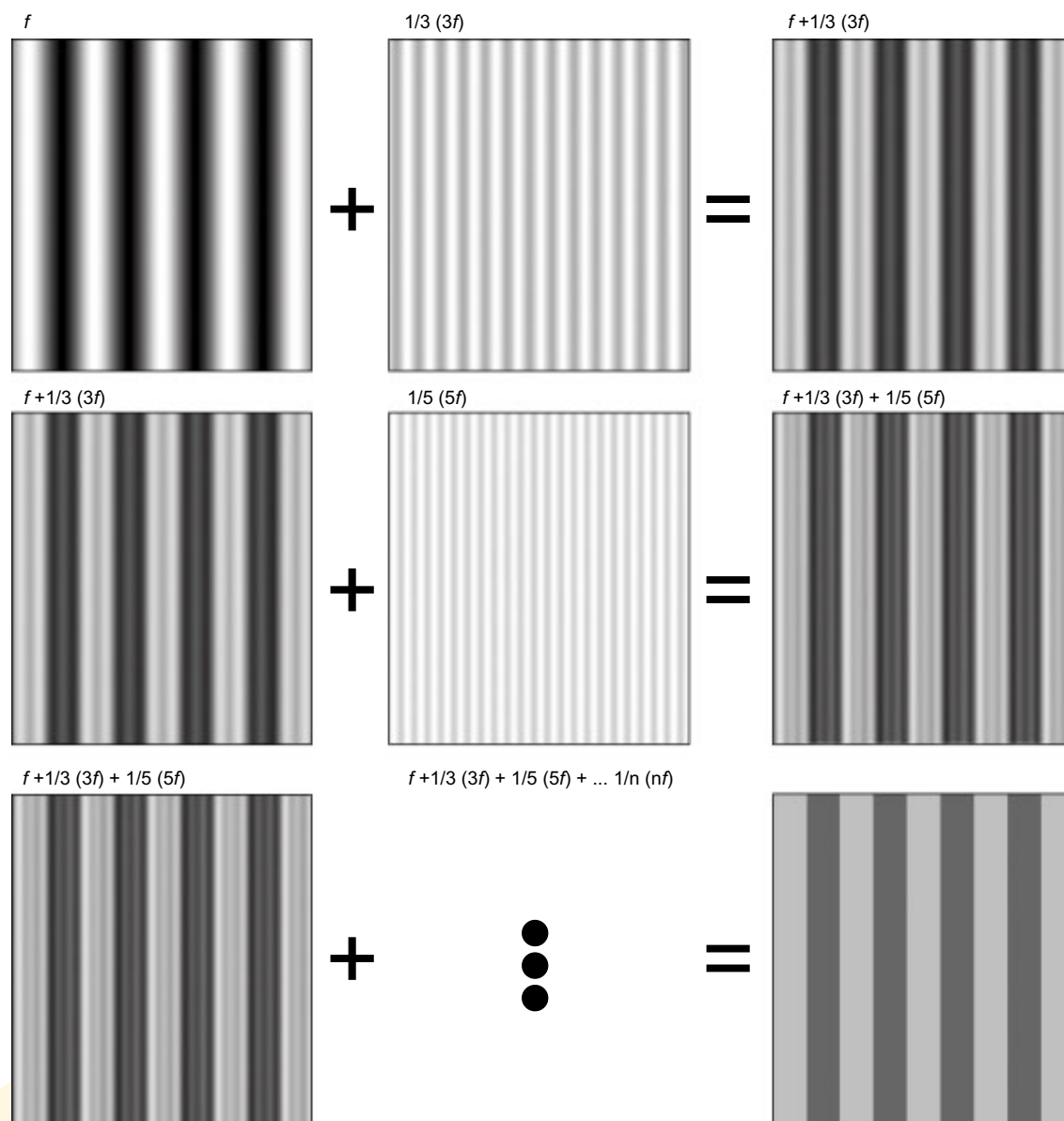


Рис. 4.2.2 Квадратно-волновая решетка (серо-черные полосы) и ее яркостный профиль (красная линия).



**Рис. 4.2.3** Создание квадратной волны из синусоидальных компонент. А — синусоидальная решетка с базовой частотой ( $f$ ) квадратной волны и яркостный профиль этой решетки. В — решетка с частотой третьей гармоники ( $3f$ ) и одной третьей амплитуды. Сложив решетки А и В, получаем решетку С. Добавив пятую гармонику ( $5f$ ) на одной пятой амплитуды получаем решетку D. Сложив все нечетные гармоники с соответствующими амплитудами и фазами получим квадратную решетку Е.

Два простых примера проиллюстрируют то, как синусоидальные решетки, объединяясь, образуют более сложные стимулы: на рис. 4.2.3 серия синусоидальных решеток одной ориентации и пространственных частот  $f$ ,  $3f$ ,  $5f$  и т.д., собраны вместе при определенных соотношениях амплитуды и фазы, дабы получить прямоугольную волну с острыми краями (а не с размытыми, как у исходных решеток); на рис. 4.2.4 показано, как две таких волны в разных ориентациях образуют клетчатый паттерн.

Разумеется, Фурье-анализ не ограничивается этими простыми, регулярными пат-

тернами — его можно применить к сложным ретинальным изображениям объектов, людей и даже полноценных сцен. Хотя мы не можем показать, как строятся такие сложные изображения из отдельных синусоидальных компонент (потребуется очень большое число решеток), мы можем продемонстрировать то, какой вид пространственной информации передается различными пространственными частотами. На рис. 4.2.5 дан портрет Гручо Маркса<sup>1</sup> вместе с двумя версиями этого портрета, несущих один — только низкие пространственные частоты, другой — только высокие. Четко видно, что низкие частоты в центре изображения передают его грубый пространственный компонент (светлые и темные области), тогда как высокие пространственные частоты на правом изображении несут информацию о тонких элементах портрета (таких как острые края и мелкие детали).

Фурье-анализ этого изображения состоит из двух частей: *спектра мощности* и *фазового спектра*. Спектр мощности характеризует амплитуду каждой слагающей изображение решетки определенной пространственной частоты и ориентации, тогда как фазовый спектр — фазу каждой решетки определенной пространственной частоты и ориентации. Когда все эти решетки собраны вместе, они воссоздадут оригинальное изображение.

Таким образом, Фурье-анализ — это универсальный метод декомпонирования сложных изображений на примитивные составляющие, поскольку доказано, что он приложим к любому изображению, в том числе и к ретинальному.

Отметим также, что Фурье-анализ можно «инвертировать», выполняя тем самым т.н. *Фурье-синтез (гармонический синтез)*, благодаря чему оригинальное изображение можно реконструировать из спектров мощности и фазовых спектров. Инвертируемость Фурье-анализа доказывает то, что спектры эти несут полную информацию об оригинальном изображении.

Применимы ли те же самые рассуждения к декомпозиции изображения на краевые

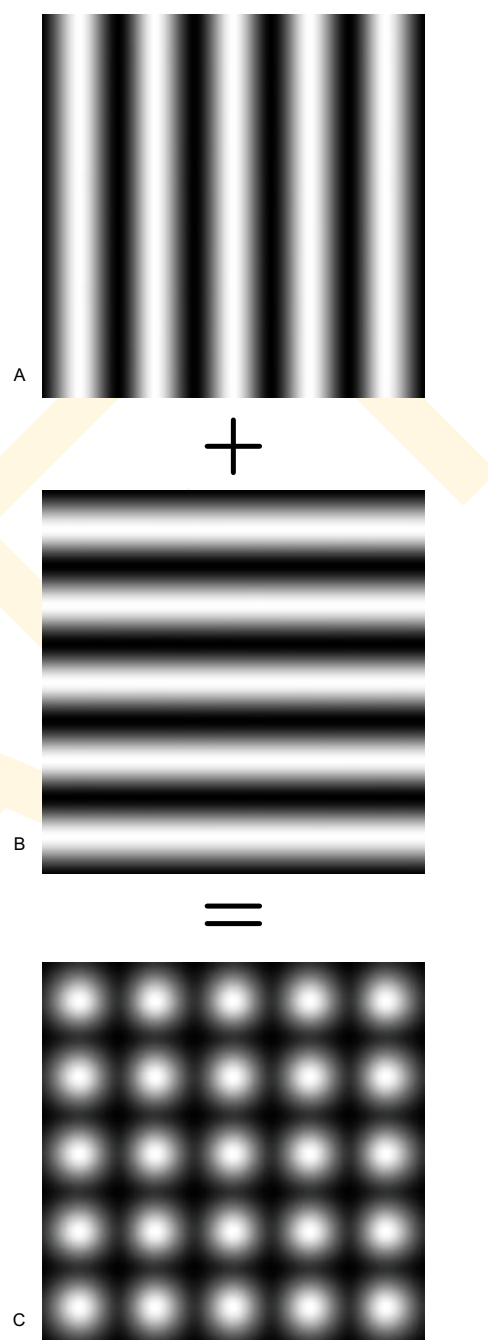
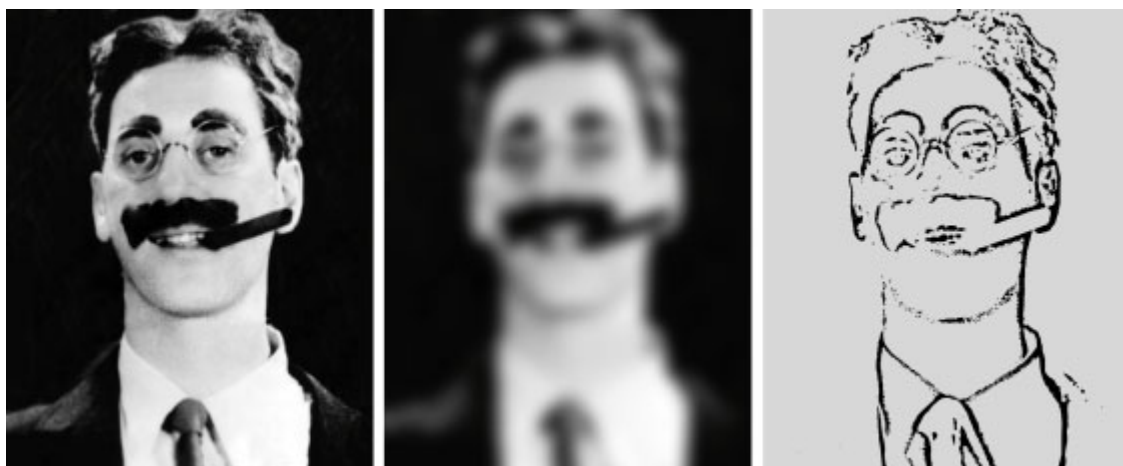


Рис. 4.2.4 Образование клетчатой решетки из двух решеток разных ориентаций.

<sup>1</sup> Гручо Маркс (Groucho Marx) знаменитый американский кинокомик первой половины XX века. Настоящее имя Джулиус Генри Маркс. — Прим. пер.



**Рис. 4.2.5** Пространственно-частотное содержимое сложного изображения: портрет Гручо Маркса (слева) разложен на низкочастотную составляющую (в середине) и высокочастотную (справа). Низкие частоты несут в себе общую картину светов и теней; высокие — информацию о локальных контрастах на краях объектов.

и линейные примитивы или к его синтезу из этих примитивов — непонятно, бо не существует теоремы аналогичной теореме Фурье, доказывающей, что линии и края могут выступать в роли пространственных примитивов. Однако как бы там ни было сама по себе математика не в состоянии привести убедительные аргументы того, что зрительная система выполняет нечто подобное Фурье-анализу ретинального изображения. Необходимы *эмпирические* доказательства — и они были получены. Рассмотрим некоторые из них.

#### 4.2.1.2 Каналы пространственных частот

Теория пространственных частот предполагает, что оптический этап зрительного восприятия толкуем в логике большого числа накладывающихся друг на друга *психофизических каналов*, отвечающих за трансляцию зрительной информации различных пространственных частот и ориентаций. Концепция психофизического канала требует некоторого объяснения, поскольку в каком-то смысле — это техническая конструкция. Однако ее можно понять интуитивно — по аналогии с каналами телевизионными.

Сигнал от всех телестанций непрерывно присутствует в окружающем нас пространстве: вещание идет за счет тех форм электромагнитной энергии, которых мы коротко коснулись в гл. 3 и частоты которых лежат далеко за пределами видимого спектра (см. рис. 3.1.2). Разные станции используют разные поддиапазоны частот, в результате чего удастся избежать паразитной интерференции сигналов. Очевидно, что важнейшим моментом в работе телевизора является его настройка на определенный канал — определенный поддиапазон длин волн, в котором идет вещание интересующей нас станции. То есть, мы говорим, что выбор нужного канала в телеприемнике — это запуск в работу внутреннего управляющего механизма, позволяющего селективно принимать сигналы в узком диапазоне частот (длин волн).

Концепция психофизического канала — это концепция гипотетического механизма в зрительной системе (чей фактический физиологический субстрат не описан), на-

строенного на ограниченный диапазон величин внутри некоего континуума. В сфере цветового зрения, к примеру, по версии Гельмгольца, существуют три канала, определяемых кривыми спектральной чувствительности гипотетических светочувствительных элементов трех классов (см. рис. 3.2.12). В теории пространственных частот, прилагаемой к науке о зрении, каждый канал определен пространственной частотой и ориентацией решеток, к которым зрение максимально чувствительно. То есть, пространственно-частотный подход к оптической обработке утверждает то, что зрительную систему можно толковать, как состоящую из множества частично перекрывающихся каналов, селективно настроенных на разные диапазоны пространственных частот и ориентаций.

Итак, теперь важно показать, что такая точка зрения имеет право на существование.

Одной из знаменательных вех истории теории пространственных частот в приложении к зрению, стала статья, опубликованная в 1969 г. британскими психофизиками Колином Блэйкмо и Фергусом Кэмпбеллом, в которой они доложили о результатах эксперимента, давшего убедительные доказательства существования в зрительной системе пространственно-частотных каналов. В результате эксперимента было показано, что если в течение длительного времени наблюдателю предъявляется синусоидальная решетка, его зрительная система селективно адаптируется к решеткам только данной ориентации и частоты (что было измерено специальными психофизическими методами). Поскольку опыт Блэйкмо — Кэмпбелла концептуально важен в контексте разговора о теории пространственных частот, мы рассмотрим этот эксперимент детально, но несколько позднее.

#### *4.2.1.3 Пространственная функция передачи модуляции хрусталика*

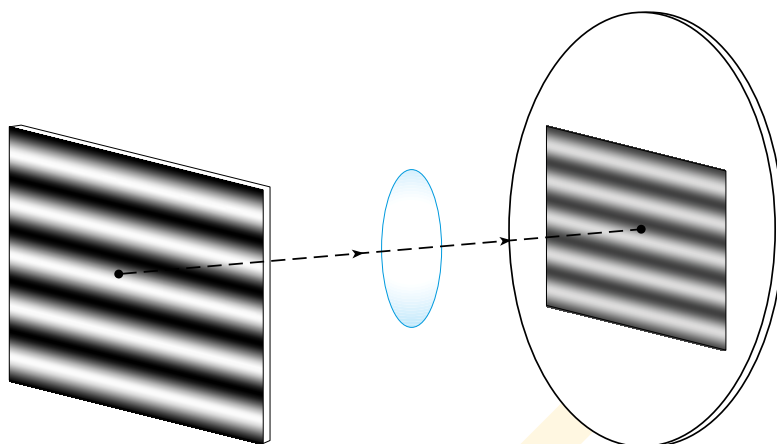
Пространственная функция передачи модуляции (sMTF — spatial modulation transfer function) свидетельствует о том, насколько добротна оптическая система передает информацию. Как определить эту функцию? Решетка определенной пространственной частоты и определенного контраста служит тестовым объектом для данной оптической системы (рис. 4.2.6): контраст проекционного изображения (в его наилучшей фокусировке) измеряется, а затем делится на контраст тестового объекта. Полученное отношение говорит нам о том, насколько хорошо оптическая система передает информацию об объекте (т.е. его контраст). У идеальной оптической системы (коих, разумеется, не существует) контраст изображения объекта будет тем же, что и у самого объекта. В реальности оптические системы всегда несколько снижают контраст.

Описанную процедуру повторяют по определенному диапазону пространственных частот — от низких до высоких.

Результатом выступает пространственная функция передачи модуляции (sMTF), аналогичная показанной сплошной линией на рис. 4.2.7. О такой кривой можно говорить как о представляющей качество изображения как функцию от пространственной частоты. Кривая, показанная на рис. 4.2.7 — это характеристика некоей оптической системы: по низким и умеренным пространственным частотам изображение передается с высокой точностью (малое снижение качества). В зависимости от аберраций, свойственных той или иной системе, изображение будет терять в качестве на больших пространственных частотах. Те же аберрации дадут сравнительно меньший эффект на низких и средних пространственных частотах.

Каково влияние оптической дефокусировки на sMTF? На рис. 4.2.7 сплошной линией показана sMTF некоей оптической системы при ее идеальной фокусировке, а

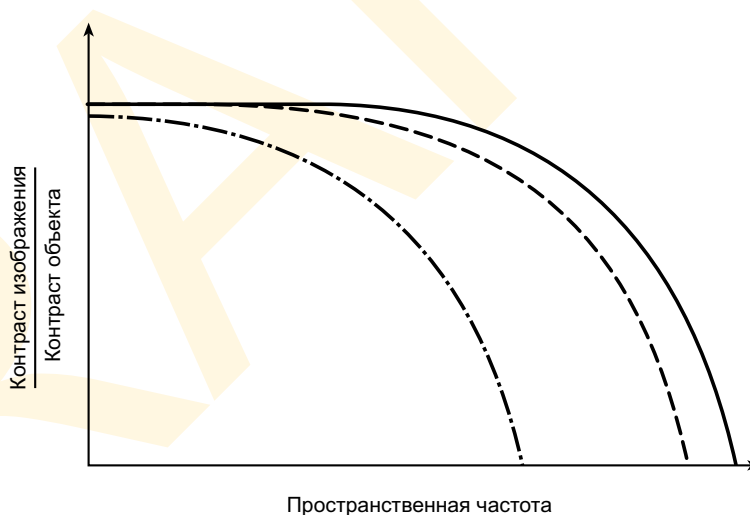




**Рис. 4.2.6** Чтобы диагностировать функцию передачи модуляции оптической системы на данной пространственной частоте, с помощью этой системы на экран проецируют синусоидальную решетку требуемой пространственной частоты, а затем сравнивают контраст проекционного изображения с контрастом оригинала.

пунктирными — передача модуляции той же системой по мере ухода системы в нерезкость. Отметим, что падение качества, в первую очередь, проявляет себя на высоких пространственных частотах и меньшей степени на средних и низких.

Предположим, что мы покрыли линзы очков прозрачным лаком для ногтей, отчего они сделались полупрозрачными. Оптическое изображение, сформированное такими линзами, сильно страдает в качестве: функция передачи модуляции такой оптической системы дает общее падение контраста на всех пространственных частотах (пунктирно-точечная кривая на рис. 4.2.6). Причина сего эффекта в избыточном светорассеянии поверхностью линз.



**Рис. 4.2.7** Сплошная линия — пространственная функция передачи модуляции (sMTF) сфокусированной оптической системы. Пунктирная линия — то же, но не в фокусе. Пунктирно-точечная линия — sMTF полупрозрачной оптики.

#### 4.2.1.4 Психофизическая контрастно-сенситивная функция зрительной системы человека

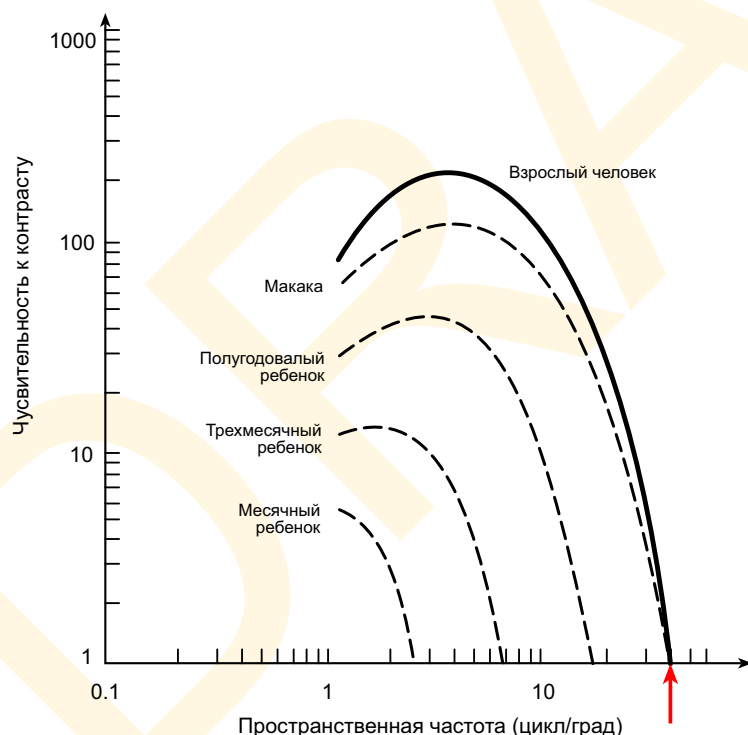
Функцию передачи яркостной модуляции зрительной системой часто именуют *контрастно-сенситивной функцией (CSF)*, поскольку измеряется чувствительность, а не контраст: испытуемого просят смотреть на экран монитора, на котором «показана»

пространственная решетка, контраст которой ниже порогового, то есть решетка эта не видна и экран выглядит равномерным. Экспериментатор слегка повышает фотометрический контраст решетки до тех пор, пока решетка не станет заметной.

Величиной, обратной пороговому контрасту, является чувствительность зрения к решетке. Пороги, промеренные по большому числу пространственных частот, дадут кривую контрастной чувствительности как функции от пространственной частоты, то есть — контрастно-сенситивную функцию.

Базовая цель эксперимента Блэйкмо — Кэмпбелла (1969), состояла в оценке влияния адаптации наблюдателя к решеткам определенных пространственных частот на пространственное восприятие. Техника эксперимента заключалась в измерении чувствительности к этим решеткам до и после адаптации.

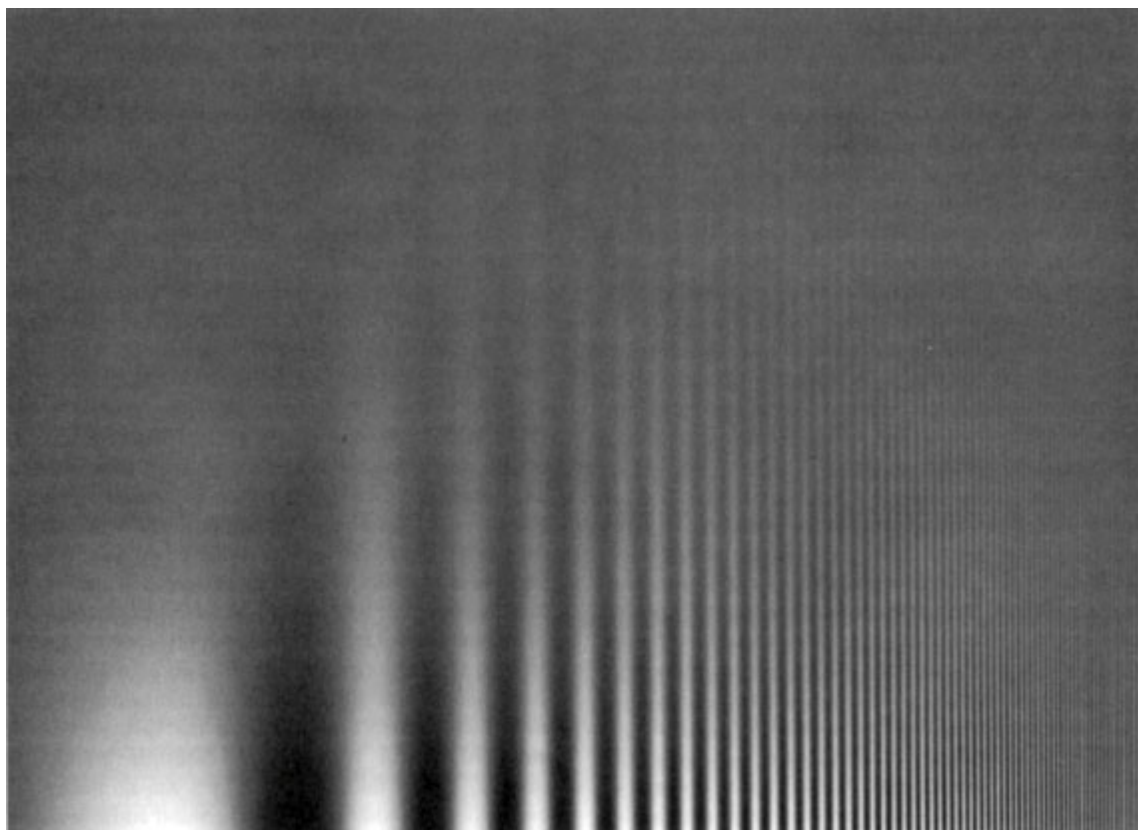
Стандартное измерение того, насколько наблюдатель чувствителен к решеткам различных частот именуется *психофизической диагностикой* контрастно-сенситивной функции (функцией чувствительности к контрасту) зрительной системы. Контрастно сенситивная функция строится нахождением минимального фотометрического контраста, при котором наблюдатель начинает едва различать разницу между решеткой и равномерно-серым полем, то есть определением порога, при котором весьма и весьма низкоконтрастная синусоидальная решетка перестает выглядеть равномерным полем и начинает восприниматься собственно решеткой. Порог измеряется по множеству решеток разных пространственных частот: от очень низких (широкие размытые штрихи), до высоких (тонкие размытые линии).



**Рис. 4.2.8** Сплошная линия — типичная контрастно-сенситивная функция зрительной системы взрослого человека. Пунктирные линии — контрастно-сенситивные функции детей первого года жизни и макаки. Красная стрелка — отсечка на высоких частотах. Отметим, что координатные оси логарифмированы. Ось абсцисс представляет 100% контраст.

Простейший и самый быстрый способ измерения контрастных порогов — это т.н. *метод управления*, при котором наблюдатель плавно поворачивает ручку повышения контраста решетки определенной пространственной частоты (отображаемой экраном компьютерного монитора) до тех пор, пока не начнет различаться ее полосчатая структура. Процедура повторяется по множеству решеток разных пространственных частот.

Типичная психофизическая контрастно-сенситивная функция взрослого человека дана на рис. 4.2.8: сие полосовая функция, дающая пик чувствительности примерно при 4 циклах/град и ее падение по обеим сторонам от этого



**Рис. 4.2.9** Демонстрация формы контрастно-сенситивной функции. Пространственная частота решетки плавно растет слева направо, а контраст — сверху вниз. Хорошо видно, что внешняя поверхность полосатого фрагмента иллюстрации в целом аппроксимирует кривую рис. 4.2.8

значения (Campbell & Robson, 1968). Человек различает решетку с пространственной частотой 4 цикла/град с контрастом меньшим, чем необходим для различения решеток прочих пространственных частот.

Свою собственную контрастно-сенситивную функцию мы можем в общих чертах отследить с помощью рис. 4.2.9, на котором даны синусоидальные полосы с пространственной частотой, растущей слева направо и фотометрическим контрастом, растущим сверху вниз. На уровне порогового контраста наша способность различать решетку пропадает, а высота, на которой полосы сливаются в сплошную серую массу, укажет на чувствительность нашего зрения при данной пространственной частоте. Если рассматривать изображение с расстояния 70-75 см, внешний контур полосчатой фигуры окажется весьма схож с графиком рис. 4.2.8.

Итак, CSF зрительной системы человека демонстрирует то, что наибольшая чувствительность приходится на срединные частоты, т.е. порядка 4-5 циклов на градус зрительного угла. К сравнению, пунктиром на рис. 4.2.8 даны несколько контрастно-сенситивных функций, построенных у детей и обезьян. Отметим, что новорожденные менее всего чувствительны к контрасту между штрихами, особенно на высоких пространственных частотах (Atkinson, Braddick & Moar, 1977).

Если CSF измерена в условиях низкой освещенности (т.н. *скотопических* условиях), то контрастно-сенситивная функция падает почти отвесно, в особенности на высо-

ких частотах. Сие указывает на то, что ночью, когда работают только палочки, острота зрения у человека намного ниже дневной.

**Отсечка на высоких частотах.** Отсечка зрительной чувствительности на высоких пространственных частотах ограничивает способность зрительной системы к различению мелких деталей, даже если их контраст равен 100%. Если продолжать повышать пространственную частоту решетки даже при 100% ее контраста, настанет момент, когда решетка перестанет восприниматься как решетка. У молодых здоровых людей высокочастотная отсечка находится примерно на уровне 60 цикл/град (показано стрелкой на рис. 4.2.8).

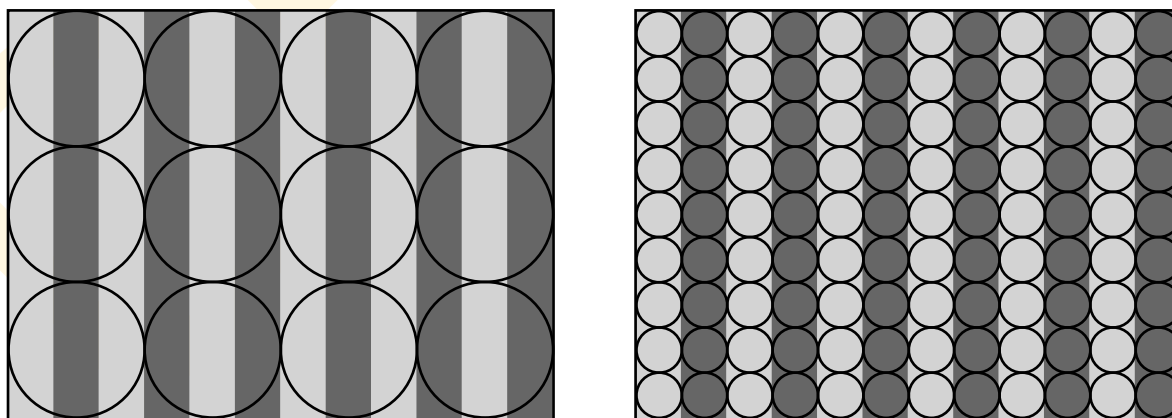
Итак, отчего же чувствительность зрительной системы падает на высоких пространственных частотах?

Одна из причин уже обсуждалась — оптические ограничения: любая оптическая система, включая глаз, дает такие ограничения из-за оптических aberrаций. Сие относится даже к идеальной фокусировке.

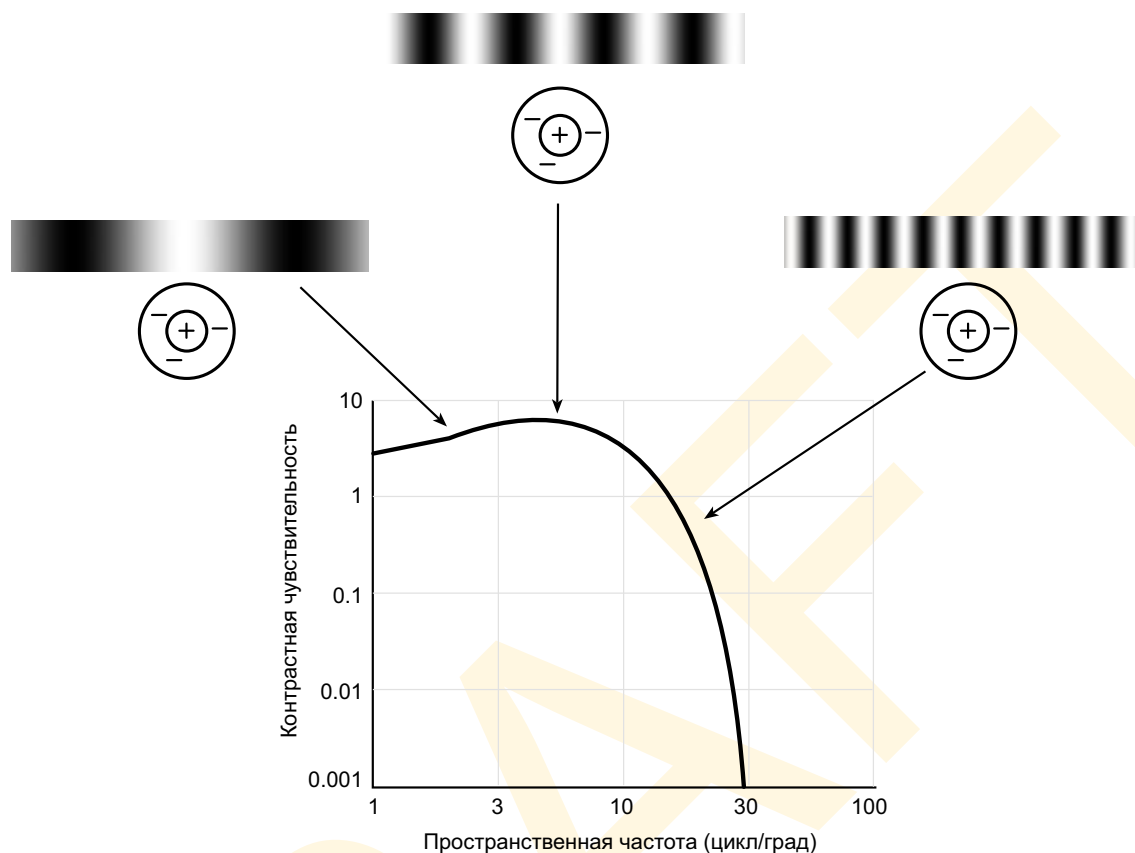
Вторая причина — это пространственная плотность залегания ретинальных фоторецепторов. Рассмотрим схематичную иллюстрацию паковки фоторецепторов на низкой и высокой плотностях (рис. 4.2.10): на каждую из матриц положена решетка. Если сие предполагает, что каждый фоторецептор суммирует весь падающий на него свет, то все равно почти невозможно вообразить, как грубая матрица слева может разрешить эту решетку. Более мелкая матрица может однако справиться с этой задачей, поскольку рецепторы упакованы намного плотнее, позволяя светлым прутьям решетки попадать на чередующиеся ряды рецепторов. Фактически, рис. 4.2.10 — это наглядная иллюстрация правила Найквиста (Willams, 1986).

Теперь рассмотрим вопрос с позиций нейронной доктрины.

Коль скоро отклик линейной системы порой удобнее характеризовать в показателях отклика этой системы на стимуляцию, описанную гармоническими функциями, то для лучшего понимания механизма нейронного кодирования информации имеет смысл измерить отклик нейронов на гармонические функции, определяя при этом величину контраста стимула, необходимую для выявления уровня уверенного отклика нейрона. Когда контрастный паттерн не оказывает эффективного влияния на нейрон,



**Рис. 4.2.10** Пространственно-частотная решетка, спроецированная на грубую матрицу фоторецепторов (слева) и на плотную матрицу (справа). Плотная матрица в состоянии разрешить эту решетку, поскольку светлые полосы решетки ложатся через колонку рецепторов.



**Рис. 4.2.11** Контрастно-сенситивная функция, описывающая чувствительность нейрона к гармоническому стимулу. Данный пример иллюстрирует то, что линейный нейрон с рецептивным полем класса оп-центр сильнее всего отвечает на стимул средней пространственной частоты, чьи светлые штрихи попадают в центр рецептивного поля, а темные — на оппозитное этому центру окружение. Когда пространственная частота стимула низка, то сигналы от центра и от окружения противоположны друг другу и чувствительность минимальна. Когда пространственная частота стимуляции высока, то отклик усредняется центром, вновь понижая чувствительность.

По отклику на гармонический стимул мы можем диагностировать пространственную структуру рецептивного поля нейрона.

мы должны предъявить зрению стимул большего контраста, дабы добиться отклика. Отметим, что когда стимулирующий паттерн четко подходит к рецептивному полю нейрона, то даже весьма небольшой контраст дает устойчивый отклик.

⇒ Минимальную величину контраста, необходимую для получения отклика нейрона называют *пороговым контрастом*. Величину обратную пороговому контрасту именуют *чувствительностью к контрасту*.

На рис. 4.2.11 мы видим, как ганглиозная клетка с рецептивным полем класса центр-окружение, откликается на косинусоидальный паттерн, пик которого пришелся на самый центр рецептивного поля нейрона. Когда пространственная частота слишком мала, то яркая полоса стимула покрывает как возбуждающий центр, так и тормозящее окружение, и поэтому устойчивый отклик мал. Наиболее эффективной окажется



ся такая пространственная частота, при которой яркая полоса решетки ляжет на возбуждающий центр вытянутого рецептивного поля, а темная — на тормозящее окружение. Такая частота идеально согласуется с рецептивным полем, и мы наблюдаем четкий нейронный отклик.

Если пространственная частота решетки окажется еще выше, то те ее фрагменты, что больше и меньше среднего попадут как на возбуждающие, так и на тормозящие участки. Совокупный эффект от такой стимуляции невелик, поэтому мы говорим, что стимулы высоких пространственных частот малоэффективны.

Итак, просуммируем отклик нейрона на стимулы разных пространственных частот в определенном диапазоне, используя т.н. *контрастно-сенситивную функцию*. Различные аспекты этой функции обеспечивают нас информацией о пространственном характере рецептивного поля нейрона: наиболее эффективная пространственная частота дает информацию об общем размере рецептивного поля; область падения чувствительности на низких пространственных частотах дает информацию о мощности оппозитного окружения; наконец, падение чувствительности на высоких пространственных частотах описывает размер центра рецептивного поля. Отметим, что нейроны с малым центром в рецептивных полях лучше откликаются на стимулы высоких пространственных частот, нежели нейроны с крупным центром рецептивного поля.

На рис. 4.2.12 показана контрастно-сенситивная функция «линейной» парвоклетки коленчатого тела макаки. Отметим, что рецептивные поля этих клеток неотличимы от рецептивных полей миджет-ганглионаров сетчатки.<sup>1</sup>

Итак, мы говорим, что:

⇒ Контрастно-сенситивная функция ретинальной ганглиозной клетки — это альтернативный вариант представления рецептивного поля этой клетки.

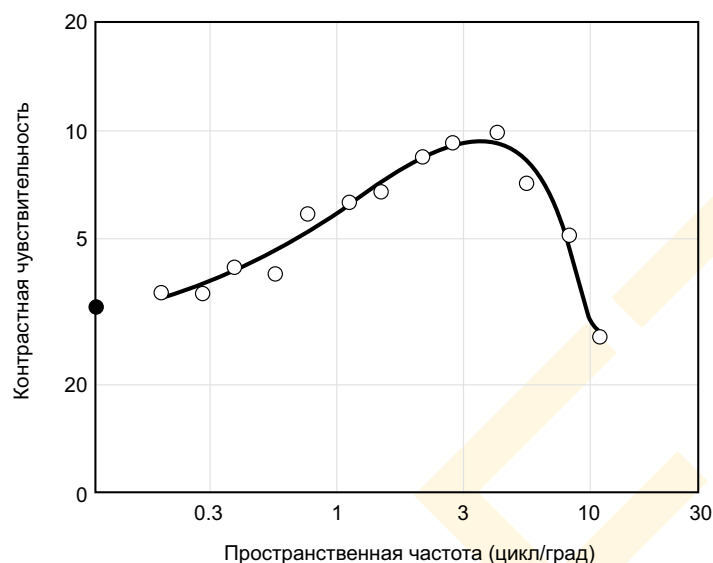
Контрастно-сенситивные функции обычно однопиковые и носят название *полосовых*. Пиковая частота этих функций именуется *центральной частотой*.

#### 4.2.1.5 Значимость контрастных паттернов

Как мы уже говорили выше, в первых исследованиях нейронного отклика на световую стимуляцию в роли входной переменной рассматривали фотометрическую яркость стимула  $i$ . Однако в данном случае от линейно-системного подхода толку мало — если мы хотим использовать линейные методы при характеристике откликов ганглиозных клеток сетчатки, важно зафиксировать средний уровень и в качестве входного сигнала рассматривать *контраст стимула* ( $a$ ).

Ставя эксперименты в логике контраста, мы не отнюдь не придаем нелинейности линейной системе: нелинейным поведение нейронов становится заметным, когда мы

<sup>1</sup> Для общего описания линейного рецептивного поля мы должны измерить отклик нейрона с использованием как синусоидальных, так и косинусоидальных контрастных паттернов. Рецептивное поле ретинальной ганглиозной клетки можно диагностировать, используя лишь косинусоид с пиком в центре, поскольку рецептивные поля ретинальных ганглионаров, как говорят — *четно-симметричны*. Мы говорим, что функция четносимметрична, если  $f(x) = f(-x)$ . Функция будет нечетно-симметричной, если  $f(x) = -f(-x)$ . Когда рецептивное поле четносимметрично, оно дает нулевой отклик на любой нечетно-симметричный стимул, поэтому измерять мы должны отклики лишь на четносимметричную стимуляцию. Для ганглиозной клетки сетчатки контрастно-сенситивная функция — это исчерпывающее описание рецептивного поля.



**Рис. 4.2.12** Диагностика чувствительности к контрасту стимуляции парвонеона ЛКТ макаки. Данный нейрон лучше всего отвесает на стимуляцию пространственной частоты порядка 5 цикл/град. Отметим, что оппозитное окружение понижает чувствительность к низкочастотным паттернам. Каждая решетка смещалась по сетчатке с такой скоростью, что в каждую точку сетчатки проецировалось 5.2 цикла решетки в секунду. Черный кружок на оси ординат — это контрастная чувствительность к пространственно равномерному паттерну, мигавшему с частотой 5.2 Гц.

По материалам Derrington & Lennie, 1984.

сравниваем измерения с разными средними уровнями. Но если организовать измерения контрастных откликов при одном фиксированном среднем уровне, то мы можем использовать линейные методы для характеристики рецептивного поля при колебаниях кверху и книзу от этого уровня. Для полнейшего описания поведения нейрона мы должны объединить нейронные отклики по множеству средних уровней интенсивности. Сие вынуждает нас признать наличие в нейронном отклике определенных нелинейностей.

Итак, большинство современных исследователей зрения полагают контраст ключевым экспериментальным параметром. Линейности, о которых мы будем говорить в текущей главе обычно измеряют в домене контраста при фиксированном среднем уровне интенсивности стимуляции. Если мы изменим этот уровень, то обнаружим, что свойства линейной системы изменились. Сие есть фундаментальная нелинейность в совокупном поведении линейных биологических систем. Отметим, что добротная теория должна учитывать вариабельность локальных линейных откликов при общей средней яркости.

### 4.2.2 Ретинальная световая адаптация

Изменения характеристик системы по мере изменения среднего уровня стимуляции именуется *зрительной адаптацией*. Мы можем измерить эффекты зрительной адаптации у отдельных нейронов, начиная с сетчатки.

Зрительная адаптация — это один из наиболее важных аспектов науки о зрении.

Рассматривая стимул в логике среднего уровня интенсивности и логике контрастных паттернов, мы отделяем эффекты зрительной адаптации от локальных контрастных эффектов. Формулировка наших входных экспериментальных данных таким путем — это очень важный момент, определяющий то, как в дальнейшем мы будем рассуждать о многих и многих аспектах зрения.

Итак, изучение рецептивных полей мы разделили на два этапа:

— во-первых, мы начали с изучения нейронного отклика к контрастным паттернам, предъявленным на равномерном среднем фоне. Контрастные паттерны колеблются по яркости вверх и вниз от среднего уровня. Путем исследования откликов на эти стимулы, мы смогли успешно применить линейные методы к спецификации преобразования оптического сигнала в нервные импульсы.

Когда мы проводим измерения по контрасту, но не по абсолютной интенсивности, тесты на линейность имеют наибольший шанс на успех. В значительной степени это объясняется тем, что диапазон фотометрических яркостей контрастных стимулов в ретинальном изображении ограничен. К примеру, пик интенсивности 100% синусоидального паттерна лишь в два раза выше средней интенсивности фона. Диапазон контрастов, кодируемых в типичных ретинальных изображениях не превышает двух порядков.

В течение дня однако, разброс абсолютных световых интенсивностей ретинальных изображений обычно держится в районе шести порядков, но зрительные пути не могут оставаться линейными в таком огромном диапазоне. Когда мы исследуем отклик на контрастные стимулы, мы обращаем внимание лишь на локальный линейный отклик общей нелинейной системы.

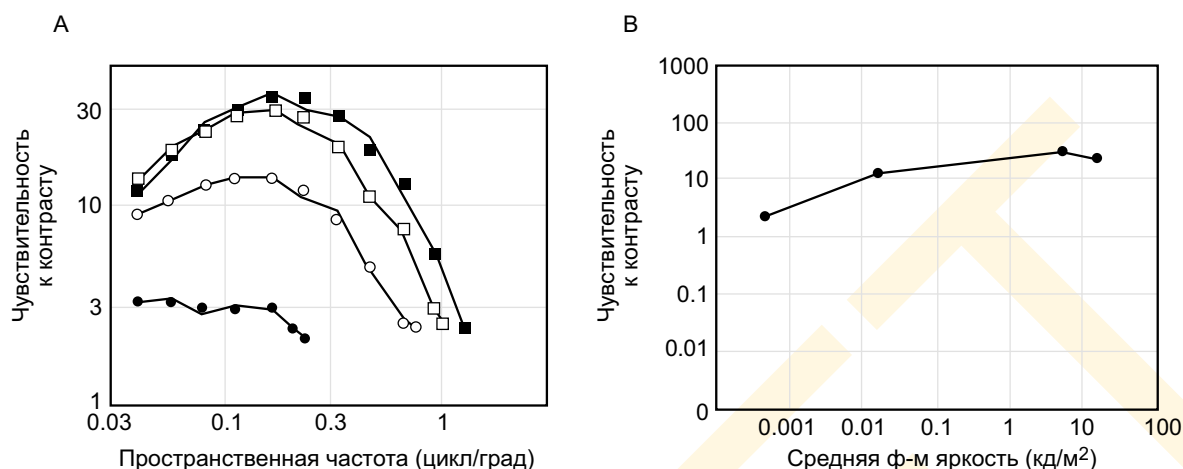
— во-вторых, для полного понимания свойств нейронов мы должны также проанализировать и то, как их отклики меняются по мере изменения среднего уровня фотометрической яркости фона. Изменение нейронного и психофизического откликов как функция от средней фотометрической яркости именуется *зрительной адаптацией*. В данном разделе мы рассмотрим то, как локальные линейные измерения, разработанные нами на базе понятия «контраст» следует расширить, когда речь идет о зрительных путях в широком диапазоне уровней среднего сигнала.

#### 4.2.2.1 Чувствительность к контрасту: зависимость от средней яркости

На рис. 4.2.13 даны несколько контрастно-сенситивных функций, полученных от ганглиозных клеток кошачьей сетчатки, и измеренных в широком диапазоне средних интенсивностей: мы видим, что когда средняя яркость стимула относительно низка, нейрон отвечает слабее на относительно высокие пространственные частоты.<sup>1</sup> На низких уровнях средней яркости функции демонстрируют слабое полосовое поведение, что указывает на легкий эффект тормозящего окружения. По мере роста фотометрической яркости фона, контрастно-сенситивные функции меняют форму: пониженная чувствительность к стимулам низких пространственных частот становится заметной, указывая на то, что влияние оппонентного окружения более выражено.

Изменение контрастно-сенситивной функции при изменении средней яркости —

<sup>1</sup> В целом, в сравнении с приматами, пространственное разрешение зрения у кошек относительно низкое, что обусловлено морфофизиологически и подтверждено путем реактивных (психофизических) измерений. Диапазон пространственных частот ганглиозных клеток сетчатки кошек много ниже нормального диапазона остроты зрения приматов.



**Рис. 4.2.13** Контрастно-сенситивная функция ретиальной ганглиозной клетки домашней кошки, измеренная при различных уровнях средней фотометрической яркости тестовых решеток. А — кривые, соответствующие контрастно-сенситивным измерениям на уровнях средней яркости (сверху вниз) в 16, 5.1, 0.016 и 0.0005  $\text{cd/m}^2$ . В — контрастная чувствительность к решетке с частотой в 0.2 цикл/град, которая меняется не более чем в 10 раз при изменении среднего уровня более чем на пять порядков.

По материалам Enroth-Cugell & Robson, 1966.

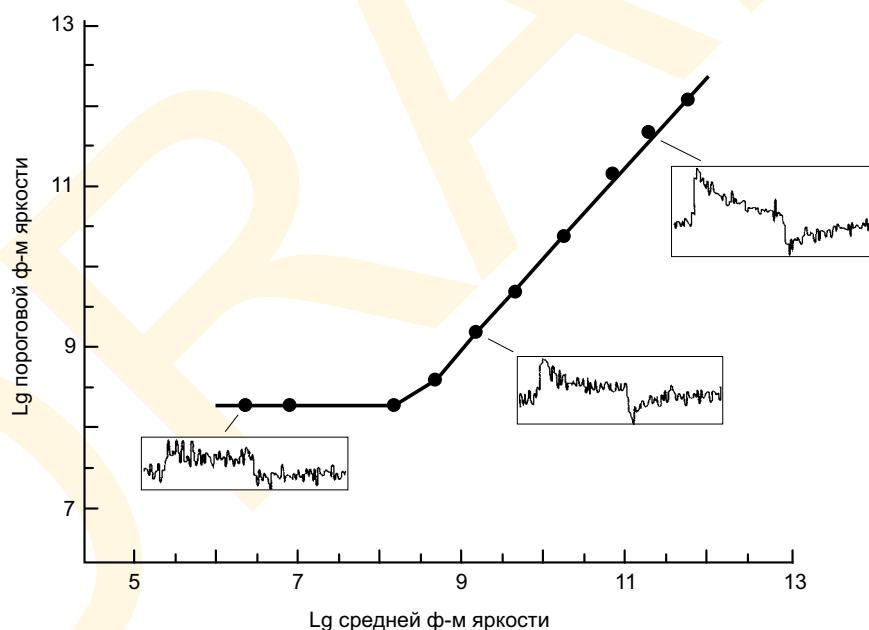
это ключ к пониманию того, как адаптация компенсирует эти изменения: на низких уровнях средней яркости нейроны попросту суммируют воздействия от фотонов, бомбардирующих рецептивное поле — при этом ингибирующее влияние окружения невелико, отчего невелико падение низкочастотной порции контрастно-сенситивной функции; на высоких уровнях средней яркости, на которых общее число фотонов существенно выше, пространственное рецептивное поле нейрона становится чувствительнее к стимулам, чья интенсивность варьирует внутри рецептивного поля. Наличие оппозитного окружения означает то, что нейрон не может более суммировать отклики на каждый фотон — вместо этого нейрон откликается больше, когда имеются вариации интенсивности внутри пространственного рецептивного поля.

На высоких уровнях средней фотометрической яркости контрастно-сенситивные функции в целом весьма похожи. К примеру, три наибольших средних яркости, использованных при построении графиков рис. 4.2.13, отличаются друг от друга с коэффициентом 100. Тогда как в диапазоне низких пространственных частот, контрастно-сенситивные функции, измеренные на этих уровнях отличны друг от друга лишь с коэффициентом 3. На рис. 4.2.13 В показано относительно небольшое изменение в контрастной чувствительности при среднем уровне фона путем сравнения клеточной чувствительности к контрастному паттерну с пространственной частотой 0.2 цикл/град на разных уровнях средней яркости. Когда средняя интенсивность меняется более чем на пять порядков, чувствительность к контрасту этого паттерна меняется всего лишь на порядок.

Относительное постоянство в чувствительности к контрасту — лишнее доказательство того, что отклик нейрона теснее связан именно с контрастом стимула, нежели с абсолютным уровнем его интенсивности. И это значит, что переменная контраста может явиться ключевой стимульной переменной, представляемой через нейронную активность.

Контраст ретинального изображения и его интенсивность тесно связанные величины. Относительная константность чувствительности к контрасту предполагает, что чувствительность нейрона к абсолютному уровню световой интенсивности меняется при изменениях уровне в среднем. Рис.4.2.14 демонстрирует эту изменчивость: по оси ординат отложен логарифм необходимого приращения световой интенсивности мигания, вызывающего фиксированный прирост частоты пульсации от уровня спонтанной активности. По оси абсцисс отложен логарифм среднего уровня интенсивности. Поскольку наклон растущей порции графика близок к единице мы можем сделать вывод, что порог растет примерно пропорционально уровню средней яркости стимула. Напомним, что взаимоотношения между порогом и средним уровнем стимуляции впервые описал Вебер (см. раздел 1.3.2.5).

По мере изменения средней интенсивности стимула, открываются всё новые и новые аспекты нейронного отклика: на вкладышах к графику даны перистимульные временные гистограммы пороговых приращений. Время, которое длится отклик к инкрементной вспышке меняется при изменении среднего уровня ее интенсивности, несколько усиливаясь при остром сигнале на подъеме и спуске. Таким образом, мы наблюдаем то, что темпоральный отклик нейрона, подобно пространственно-контрастной чувствительности, меняется при смене среднего уровня интенсивности. Вероятно, биологический смысл адаптации состоит в том, чтобы в полном объеме реализовать высокое качество сигнала на высоких уровнях световой интенсивности.



**Рис. 4.2.14** Пороговая чувствительность как функция от средней яркости, измеренная у ретинальной ганглиозной клетки домашней кошки. Порогом полагается приращение яркости, дающее устойчивый пик пульсации нейрона. Тестовая вспышка предъявлялась на большом стационарном фоне. Логарифм пороговой яркости тестовой вспышки линейно растет по мере роста логарифма средней яркости. Наклон примерно равен единице. Следовательно, пороговая яркость пропорциональна средней яркости. Вставки -- перистимульные гистграммы, демонстрирующие активность ганглиозной клетки, которая меняется по мере изменения уровня яркости.

По материалам Enroth-Cugell et al., 1977



#### 4.2.2.2 Сравнение с психофизической контрастной чувствительностью

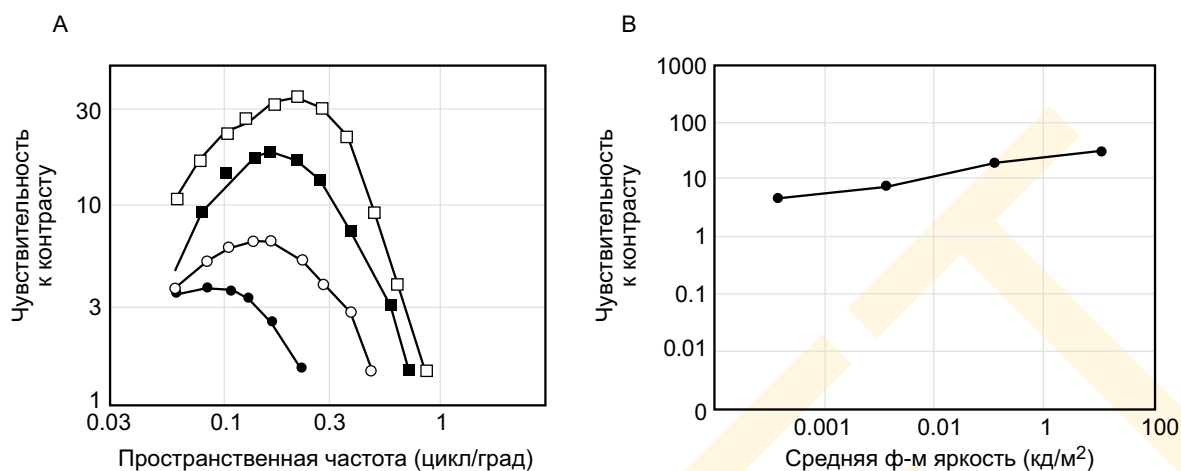
Чувствительность к контрасту и чувствительность к абсолютному уровню яркости стимула — это взаимодополняющие друг друга показатели. Когда порог к абсолютному уровню пропорционален среднему, как то прогнозирует закон Вебера, чувствительность к контрасту будет строго постоянной. В большинстве случаев, однако, веберовский закон — это лишь грубая аппроксимация реальной физиологии. Рассмотрим, к примеру, контрастно-сенситивные функции рис. 4.2.14: в тех ее участках, где закон Вебера точен — все контрастно-сенситивные функции должны совпасть друг с другом. Но очевидно, что этого не происходит, более того, качество веберовской аппроксимации зависит от пространственной организации стимула — приближение выше по низким пространственным частотам и ниже по высоким.

Логика, основанная на данном типе аппроксимации — это часть тех трудностей, с которыми сталкивается читатель, изучающий биологию сенсорных систем: обсуждаемые нами законы — далеко не столь точны как законы физические, и, чаще, это вопрос оценки того адекватен ли тот или иной закон биометрическим данным и насколько он согласуется с принципиальной точкой зрения на функцию нейрона или зрительного пути? В отношении того, что контраст — это ключевая переменная, кодируемая ганглиозными клетками сетчатки научным сообществом достигнут консенсус. Отчасти консенсус этот строится на основе данных по нервным откликам, что мы рассматриваем в текущей главе, отчасти он зависит от анализа, который мы проведем позднее и при котором будем говорить о наиболее важных сигналах в функции зрения.

Еще один момент, который нам следует разобрать — это связь тех или иных свойств контрастно-сенситивных функций, измеренных на нейронном уровне, с характером психофизической реакции особи.

Физиологические контрастно-сенситивные функции ганглиозных клеток сетчатки имеют свои копии в *психофизических* контрастно-сенситивных функциях. Американские исследователи Татьяна Пастернак и Вильям Мериган (1981) измерили последние у кошек при различных средних интенсивностях стимуляции (рис. 4.2.15) и обнаружили, что психофизические контрастно-сенситивные функции — суть прямые проекции нейронных, к примеру: если средний уровень интенсивности стимула растет, контрастно-сенситивные функции, измеренные психофизическим способом, становятся более полосовыми. При этом, если средний уровень меняется в диапазоне шести порядков, чувствительность к контрасту варьирует в рамках всего лишь одного порядка. Следовательно, имеет место четкая количественная согласованность между реактивными чувствительностями отдельных ганглиозных клеток сетчатки и психофизической реакцией особи.

Согласованность между нейронными измерениями и измерениями психофизическими демонстрирует то, что информация, представленная в откликах отдельных нейронов, аналогична информации используемой кошкой, когда она выполняет психофизическую оценку, обусловленную наличием контрастного паттерна. Сие не значит однако, что отклики отдельных ганглиозных клеток сетчатки управляют психофизикой особи. Мы уже показали, к примеру, что информация о длине волны, закодированная фоторецепторами, не эквивалентна информации, которой пользуется наблюдатель, выполняя трихроматическое уравнивание на колориметре. Также сие означает, что когда мы формулируем наше сознательное решение о достижении цветового равенств, мы не имеем доступа к информации, закодированной колбочками. То есть, согласование между нейронным и психофизическим откликами весьма абстрактно. Мы вернемся к сравнениям такого типа, когда начнем изучать зрительную



**Рис. 4.2.15** Психофизическая контрастно-сенситивная функция домашней кошки. А — измерение при среднем уровне яркости в 16 кд/м² (белые квадратики), 0.16 кд/м² (черные квадратики), 0.0016 кд/м² (белые кружки) и 0.000016 кд/м² (черные кружки). Индивидуальные психофизические контрастно-сенситивные функции демонстрируют те же качественные свойства, что и контрастно-сенситивные функции отдельных ганглиозных клеток. По мере роста уровня средней яркости контрастно-сенситивные функции становятся все более полосовыми. В — при изменении уровня средней яркости более чем на шесть порядков контрастная чувствительность к решетке с пространственной частотой в 0.3 цикл/град меняется всего лишь в 10 раз.

кору и психофизику человека. Вспомним, что наилучший метод сравнения нейронного и психофизического откликов — это сравнительный анализ информации, доступной на той или иной точке зрительного пути с информацией, которой располагает человек, принимая то или иное зрительное решение.

#### 4.2.2.2 Пространственная (паттерн-) адаптация

Давайте вернемся к эксперименту Блэйкмо — Кэмпбелла. После диагностики CSF у каждого испытуемого, экспериментаторы адаптировали наблюдателей к решетке той или иной пространственной частоты, предъявляя каждую в течение нескольких минут. После чего проводилось повторное измерение порога по каждой пространственной частоте. Длительная экспозиция вынуждала зрительную систему наблюдателей адаптироваться, т.е. понижать свою чувствительность к стимуляции (см. раздел 1.1.3.1), но при этом только к стимулам пространственной частоты и ориентации сходной с таковыми у адаптирующей решетки.

**Адаптация по частоте.** На рис. 4.2.16 А дана постадаптационная CSF, демонстрирующая селективное понижение чувствительности по частоте адаптирующей решетки. Решетки больших и меньших пространственных частот не влияли на характер адаптации.

Степень адаптации можно измерить, построив график отличия постадаптационной CSF от исходной (рис. 4.2.16 В).

Эффект пространственной адаптации (паттерн-адаптации) можно оценить, сравнив контрастно-сенситивные функции до и после адаптации: сплошная кривая демонстрирует нам контрастно-сенситивную функцию до адаптации; после адаптации к си-

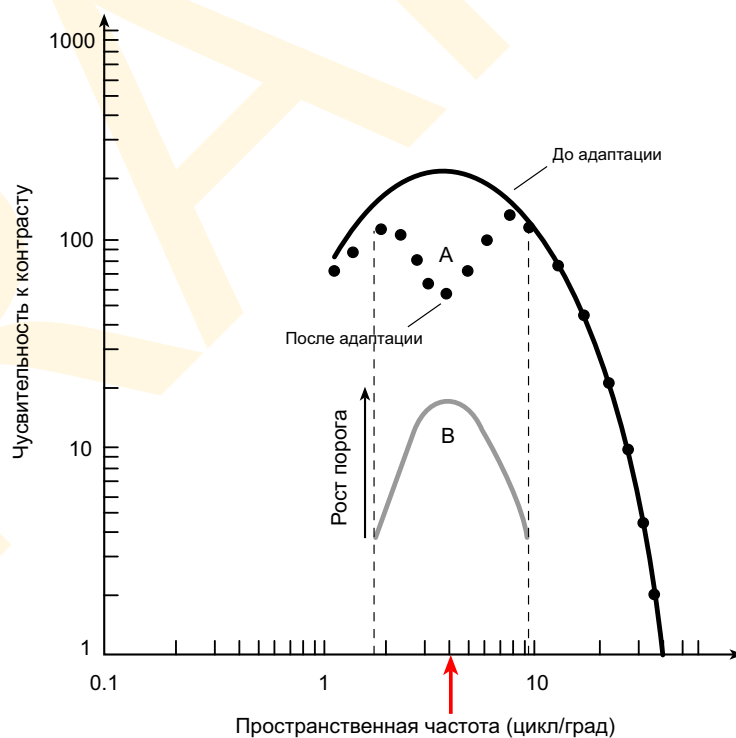
нусоидальной решетке в течение нескольких минут (выполненной по уже знакомой схеме) контрастная чувствительность наблюдателя к стимулам примерно той же пространственной частоты, что и у адаптирующего паттерна, снижается, тогда как чувствительность к прочим контрастным паттернам остается без изменений.

Когда данный эксперимент повторяют, используя пространственные решетки других частот, снижается чувствительность к тем паттернам, чья пространственная частота близка к адаптирующей (Blakemore & Campbell, 1969).

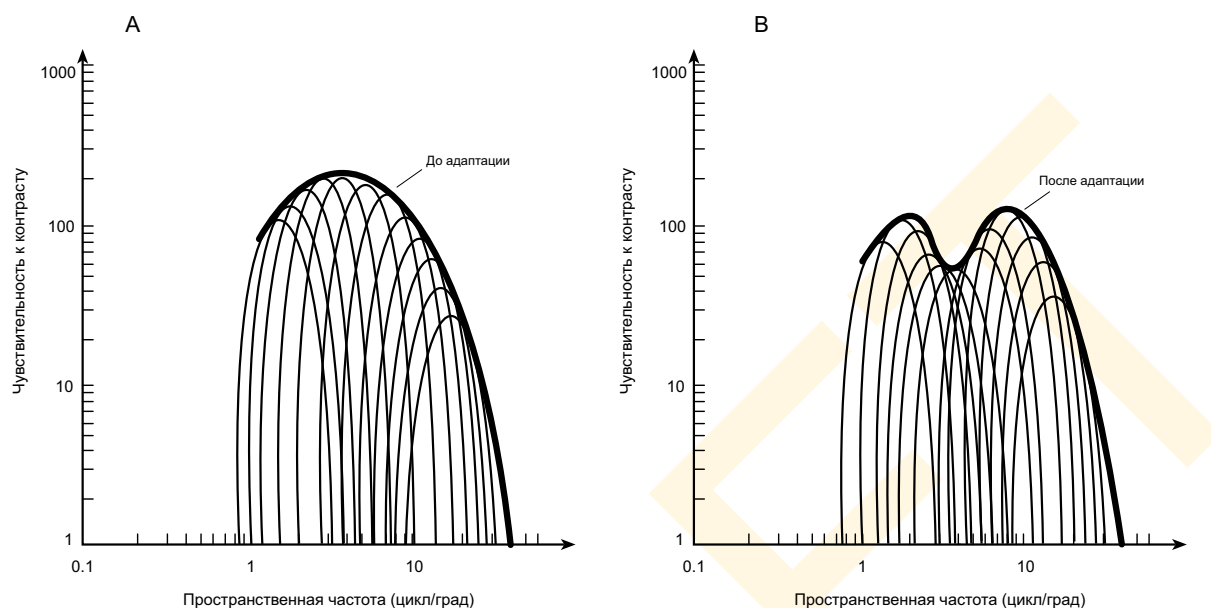
Итак, селективная адаптация — это один из маркеров психофизического канала: каждый канал адаптируется до такой степени, которая отражает его чувствительность к адаптирующему стимулу. Таким образом, адаптация понижает чувствительность только лишь в малом диапазоне пространственных частот, не затрагивая при этом весь диапазон.

Результаты эксперимента Блэйкмо — Кэмпбелла проще всего толкуемы концепцией пространственно-частотных каналов, которая позиционирует то, что исходно измеренная широкополосная контрастно-сенситивная функция представляет собой фактически совокупность вкладов множества перекрывающихся узкополосных каналов, каждый из которых чувствителен к стимулам различных диапазонов пространственных частот (см. рис. 4.2.17 А).

Когда адаптирующая решетка предъявляется наблюдателю в течение длительного периода, то каналы, чувствительные к этим пространственным частотам *утомляются*, то есть «устают» и отвечают не столь активно. Сие утомление представлено справа на рис. 4.2.17 В пониженной чувствительностью каналов в районе пространственной частоты адаптирующей решетки. Таким образом, общая контрастно-сенситивная функция после адаптации в районе адаптирующей частоты имеет «ямку», поскольку после адаптации специфические каналы, ответившие на решетку, лежащую в данном диапазоне частот, стали менее чувствительны к тем же (или подобным) стимулам. Стоит отметить, что сей эффект неприменим к простым побуждениям, поскольку глаза наблюдателя все время



**Рис. 4.2.16** Чувствительность к контрасту до и после адаптации к синусоидальной решетке. Сплошная черная кривая — нормальная контрастно-сенситивная функция до адаптации. Черные точки — данные, полученные после адаптации к решетке с пространственной частотой 8 цикл/град (отмечена красной стрелкой). Серая кривая демонстрирует отличие между двумя измерениями, то есть рост порога (снижение чувствительности).



**Рис. 4.2.17** Гипотеза совокупного действия пространственно-частотных каналов. Предполагается, что результирующая форма контрастно-сенситивной функции (жирная линия на левом графике) может быть следствием наложения множества пространственно-частотных каналов (тонкие параболы). Предполагается также, что пролапс (проседание) контрастной чувствительности, возникающий после адаптации (жирная линия на правом графике), является следствием селективной адаптации по каналу, близкому к адаптирующей частоте.

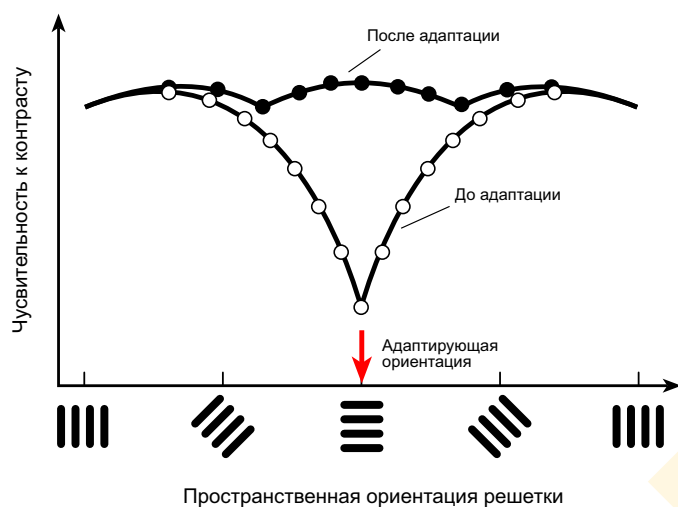
«ощупывают» решетку в период адаптации, полностью смазывая любой пстобраз.

**Адаптация по ориентации.** Селективная адаптация имеет место и в отношении ориентации решеток (Blakemore & Campbell, 1969; Blakemore & Nachmias, 1971). Мы уже демонстрировали последствие такой адаптации в гл. 1 (рис. 1.1.3.2). Чтобы измерить величину этого эффекта в пороговых показателях, вначале мы должны определить чувствительность каждого испытуемого к синусоидальным решеткам той или иной пространственной частоты во многих ориентациях. По выполнении сего, испытуемого адаптируют к единичной решетке определенной ориентации. Чувствительность к каждой ориентации затем переопределяется измерением постадаптационных порогов к решеткам той же пространственной частоты, но иных ориентаций.

На рис. 4.2.18 дан результат такого эксперимента по решетке низкой пространственной частоты, четко указывающий на понижение чувствительности по ориентациям близким к адаптирующей: мы видим, что образуется «ямка», аналогичная такой же как и у пространственной CSF.

#### 4.2.2.4 Пространственно-частотные последствия

Наряду с тем, что синусоидальные решетки определенных частоты и ориентации дают специфические эффекты адаптации — они также рожают специфические последствия, в чем нам поможет убедиться Рис. 4.2.19: во-первых посмотрим на правую его часть и удостоверимся, что верхние и нижние решетки совершенно одинаковы; затем сосредоточимся на левых решетках примерно на одну минуту, без излишнего напряжения удерживая внимание на горизонтальной полоске между ними (в ре-



**Рис. 4.2.18** Селективная адаптация к ориентации решеток. Чувствительность к решетке дана как функция от ориентации этой решетки до адаптации (белые кружки) и после (черные). Адаптирующая ориентация отмечена стрелкой.

По материалам Bradley, Switkes & De Valois, 1988).

ланию штрихами разных пространственных частот. Продолжительная стимуляция решетками с широкими прутьями утомляет нейроны, селективно реагирующие на стимулы низких пространственных частот верхней части поля зрения. Аналогично продолжительная стимуляция решеткой с тонкими прутьями утомляет клетки, селективно отвечающие на стимулы высоких частот в нижней части поля зрения. В итоге, две физически идентичных решетки справа начинают выглядеть по-разному, поскольку «высокочастотные» клетки после адаптации оказались относительно более чувствительны к низкочастотным решеткам, и «низкочастотные» клетки — к высокочастотным решеткам. Эти отклонения в чувствительностях приводят к сдвигу паттерна активности в противоположных направлениях в восприятии верхней и нижней решеток справа и вызывают при этом ощущение того, что прутья справа у верхнего образца истончены, а у нижнего — утолщены.

Читатель может без труда перенести давешние рассуждения на *ориентационно-специфические последствия*, часто именуемые *уклон-специфическими* или просто *тилт-последствиями* (рис. 1.1.3). Объяснение в логике нейронного утомления, меняющего паттерн активности в ответ на идентичные стимулирующие решетки, применимо и здесь, за исключением того, что обсуждаемые клетки селективно настроены не на разные пространственные частоты, но на разную ориентацию решеток.

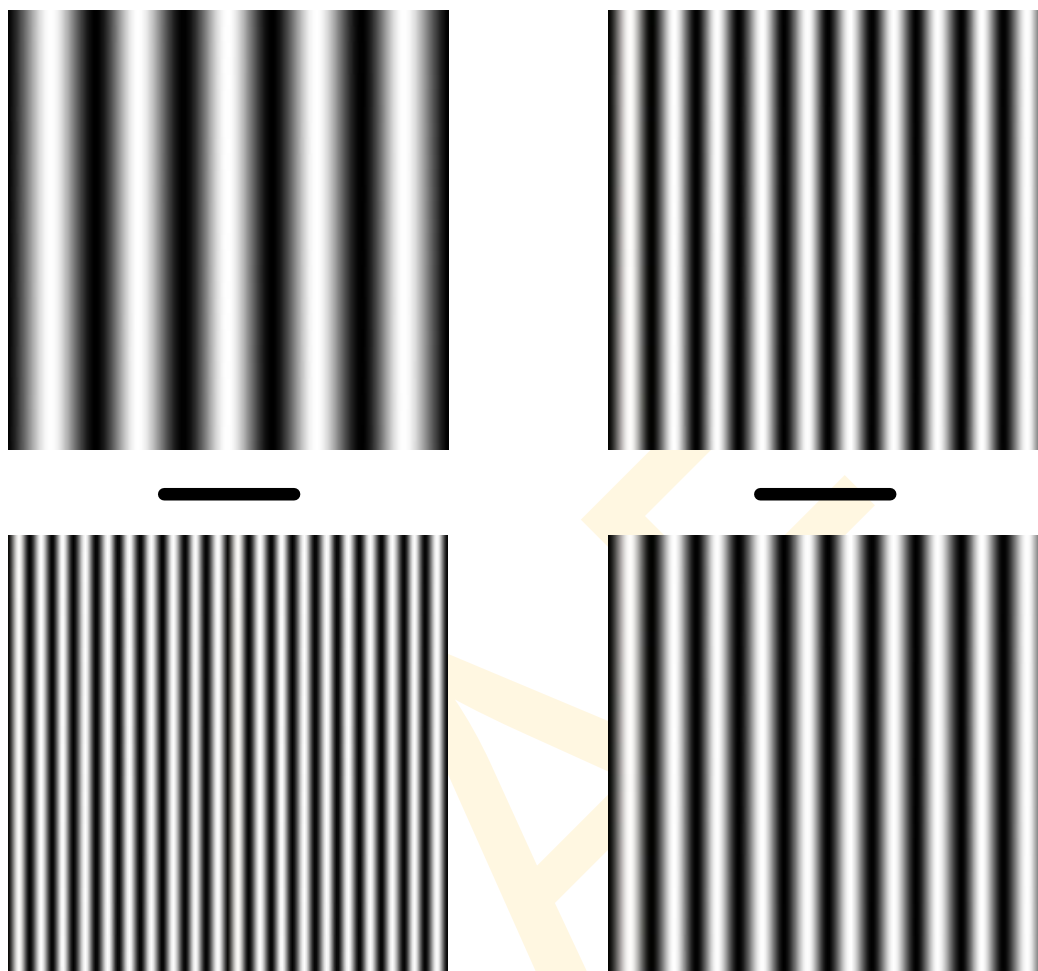
#### 4.2.2.5 Пороги чувствительности к синусоидальным и квадратно-волновым решеткам

Обширные психофизические исследования по распознаванию и различению решетчатых стимулов подтвердили правомочность применения теории пространственных частот к толкованию оптического этапа пространственного зрения. Среди них не-

зультате чего верхняя решетка адаптирует нижнюю половину сетчатки, а нижняя — верхнюю); после чего переведем взгляд на горизонтальную полосу справа. Мы видим теперь, что верхняя решетка явно увеличила свою частоту, то есть ее полосы стали выглядеть тоньше, а их паковка — плотнее, чем у нижней решетки. И это несмотря на то, что обе решетки физически остались прежними. Следовательно, восприятие выдало два различных последствия в ответ на адаптацию решетками разной пространственной частоты.

Стандартное объяснение пространственно-частотных последствий весьма схоже с таковым у хроматических постобразов, с той только разницей, что в процессе участвуют клетки, селективно настроенные на стиму-





**Рис. 4.2.19** Демонстрация последействия при пространственно-частотной адаптации. Убедимся, что обе решетки справа физически идентичны. затем адаптируемся порешеткам слева, удерживая взгляд на левой горизонтальной полоске примерно в течение минуты. Затем переведем его на центральную полоску справа и сравним визуальную ширину штрихов у правой верхней и правой нижней решеток.

обычайно интересны эксперименты, которые с целью изучения отличий между детекцией синусоидальных и квадратно-волновых решеток провели психофизики Фергус Кэмпбелл и Джон Робсон (1968). Исследователи выполнили ряд т.н. контринтуитивных прогнозов, основанных на теории пространственных частот, и обнаружили, что прогнозы эти оказались на удивление точны.

Основой прогнозирования послужила гипотеза о том, что квадратно-волновая решетка должна быть представлена на начальных этапах зрительного восприятия не как самостоятельный цельный стимул, но как совокупность множества синусоидальных решеток разных пространственных частот и амплитуд.

Теория пространственных частот утверждает в частности, что квадратно-волновая решетка с частотой  $f$  и амплитудой  $a$  декомпозируется на синусоидальную решетку с частотой  $f$  и амплитудой  $a$  плюс еще одну решетку с частотой  $3f$  и амплитудой  $a/3$ , плюс третью решетку с частотой  $5f$  и амплитудой  $a/5$  и т.д. (рис. 4.2.3).

Первый прогноз, проверенный Кэмпбеллом и Робсоном состоял в том, что порог де-

текции квадратно-волновой решетки окажется ровно таким же как и порог детекции решетки синусоидальной той пространственной частоты ( $f$ ), что является т.н. *опорной (наинизшей) частотой квадратной волны*. Рассуждение простое: порог детекции квадратно-волновой решетки окажется преодоленным всякий раз, когда любая из этих синусоидальных компонент перейдет свой собственный независимый порог. Компонент опорной (базовой) частоты преодолеет порог первым, поскольку его амплитуда много выше таковой у всех высших гармоник квадратной волны ( $3f$ ,  $5f$ ,  $7f$  и т.д.).

Кэмпбелл и Робсон проверили свои прогнозы, определив контрастный порог (т.е. минимальную амплитуду), при которой квадратно-волновая решетка начинает выделяться из равномерного серого поля, фотометрическая яркость которого равна средней фотометрической яркости решетки (см. Приложение А). С точки зрения испытуемого задача сводилась к попытке углядеть в сером поле хоть какую-то полосчатость. Исследователи обнаружили, что при такой визуальной задаче по квадратно-волновой решетке порог ровно такой же как и по решетке синусоидальной, чья пространственная частота равна опорной частоте квадратной волны.

Сие открытие в первый момент показалось неожиданным, бо яркостный градиент квадратно-волновой решетки много более крутой (т.е. намного быстрее меняется от светлого к темному), нежели у решетки синусоидальной. С этой точки зрения наиболее очевидный прогноз был бы таков, что у квадратно-волновой решетки порог должен оказаться ниже, чем у синусоидальной, поскольку яркостный градиент круче. Но данное рассуждение оказалось неверным — разницы в порогах выявлено не было.

Кэмпбелл и Робсон также проанализировали контрастный порог, при котором наблюдатели начинают отличать синусоидальную и квадратно-волновую решетки одинаковой пространственной частоты. В этом случае испытуемым предъявлялись два штриховых паттерна одновременно. Теория пространственных частот и здесь дает уверенный прогноз: контрастный порог визуального отличия синусоидальной решетки от квадратно-волновой должен быть таким же как порог отличия равномерного серого поля от синусоидальной решетки с пространственной частотой третьей гармоники ( $3f$ ) квадратной волны. Сей прогноз также основан на гипотезе, что зрительная система декомпонирует квадратно-волновую решетку на серию синусоидальных компонент: опорную ( $f$ ) и нечетные гармоники ( $3f$ ,  $5f$ ,  $7f$  и т.д.). Ежели так, то отличие между квадратно-волновой решеткой и синусоидальной при одной и той же опорной частоте имеет место только в нечетных гармониках и чтобы уловить это отличие одна из этих гармоник должна перешагнуть свой независимый порог контраста. Поскольку третья гармоника квадратной волны обладает наибольшей амплитудой ( $a/3$ ), чем все прочие нечетные гармоники, она должна перейти свой порог первой, и этот порог должен быть пересечен на том же контрасте, на котором третью гармонику можно отличить от равномерного поля.

Результаты эксперимента вновь продемонстрировали контролепригодность теории пространственных частот и точность ее прогнозов. И стоит отметить, что ни одна из существующих теорий не может выполнить их точнее. Со временем выполнили множество других экспериментов по проверке теории пространственных частот было выполнено, и, в результате, теория эта в последние десятилетия стала доминирующей психофизической теорией начального этапа пространственного зрения.

#### 4.2.2.6 Онтогенез пространственно-частотных каналов

Психофизические исследования зрения детей показали, что малыши видят мир радикально отлично от взрослых и как минимум по ряду моментов. Исследования эти



**Рис. 4.2.20** Имитация восприятия женского лица взрослым человеком (слева) в сравнении с таковым у младенца (справа). Хорошо видно, что из правого изображения исключены высокие пространственные частоты.

строились обычно в парадигме зрительных предпочтений, которой мы коснулись в разделе 3.2.4 (Fantz, 1958, 1965), но с синусоидальными решетками в роли стимулов. Ребенку обычно на одной стороне поля зрения предъявляли синусоидальную решетку, а на другой его стороне — серое поле. Оба стимула были уравненными по средней фотометрической яркости, так что разница имела место лишь в уровне модуляции между темными и светлыми прутьями решетки. Если ребенок не видит отличий, то на разглядывание каждого из стимулов он тратит примерно одно и то же время. Если ребенок разницу видит, то очевидно, что решетке он посвятит больше времени, поскольку она визуальнее интереснее. Меняя контраст решетки и измеряя время, потраченное малышом на разглядывание стимула, были измерены контрастно-сенситивные функции младенцев разных возрастов (рис. 4.2.8) (Atkinson et al., 1977).

Опыты показали, что дети в целом менее чувствительны к решеткам и наибольшее отличие имеет место на высоких пространственных частотах. Перцепционное значение этих фактов проиллюстрирует рис. 4.2.20: слева показан результат восприятия женского лица взрослым человеком; справа — как (вероятно) видят его дети, у которых набор воспринимаемых пространственных частот еще достаточно ограничен. Размытие контуров — следствие детской нечувствительности к пространственно высокочастотной стимуляции.

#### 4.2.2.7 Физиология пространственно-частотных каналов

Психофизические каналы — это гипотетические механизмы, выведенные умозрительно из психофизических измерений, но не из физиологии нервной системы. То есть, психофизические каналы информационной обработки относятся к марровому алгоритмическому уровню описания, но не его имплементарному уровню. Однако ежели каналы сии реальны, они должны иметь какой-то, как говорят, *анатомо-физиологический субстрат*. Наша задача разобраться в том, каков он?

Ответ на данный вопрос есть теория функции клеток Хьюбела — Визеля в стриарной коре. Существует весьма серьезные доказательства тому, что клетки эти способны к выполнению *локального пространственно-частотного анализа* информации о ретинальном изображении, поступившей из колленчатого тела. Анализ «локален» потому что рецептивные поля стриарных клеток пространственно ограничены несколькими градусами зрительного угла.

Сие безусловно сильное упрощение в сравнении с теоретической бесконечностью набора синусоидальных решеток, на которых базируется классический Фурье-анализ. Это даже еще большее упрощение, чем применение крупных решетчатых стимулов ( $10^\circ$  и более) в психофизических исследованиях.

Локальный, кусочный пространственно-частотный анализ выполним посредством дробления стимула на мелкие синусоидальные решетки, плавно «исчезающие» по мере удаления от центра рецептивного поля (рис. 4.2.21). Такая организация рецептивных полей (именуемая *функцией Габора* или *вейвлетом*) построена умножением синусоидальной решетки на колоколообразную гауссову кривую. Одномерный яркостный профиль данной функции показан на рис. 4.2.21 А совместно с полным двухмерным представлением, демонстрирующим то, как световая интенсивность меняется по габоровой зависимости.

Как было показано в разделе 4.1.2, Рассел де Валоис и Карен де Валоис выполнили полный мэппинг рецептивных полей клеток V1, получив бесспорное доказательство множественности участков возбуждения и торможения (см. рис. 4.1.23).

Дабы укрепить связь между нашим представлением о функции этих клеток и теорией локальных пространственных частот Де Валоисы, а также Дуэйн Альбрехт и Лайза Торелл (1982) измерили пространственно-частотную настройку простых и сложных стриарных нейронов. Исследователи обнаружили, что множество клеток «заточены» на весьма узкий диапазон пространственных частот, как следовало бы ожидать, если бы они были внедрены в локальные пространственно-частотные каналы мозга. На рис. 4.1.42 дан пример построения частотно-сенситивной характеристики клеток стриарной коры макаки.

Уровень «заточки» корковых клеток очевидно падает по сторонам от пиковой частоты (De Valois et al., 1982). Мы видим, что «высокочастотные» клетки в целом настроены тоньше, селективнее «низкочастотных». Простые клетки очевидно склонны к более острой «заточке», нежели клетки сложные, хотя в целом отличие невелико.

Сходную картину наблюдаем и в отношении настройки клеток по ориентации стимула: некоторые клетки отвечают на стимуляцию только теми решетками, чья ориентация близка к приоритетной для данных клеток, тогда как другие дают равновеликий отклик на решетки любой ориентации.

Как видим пространственно-частотные и ориентационные характеристики пороговых клеток коррелируют друг с другом: клетки низкой частотной селективности низкоселективны и в отношении пространственной ориентации решеток и наоборот — частотно специализированные клетки высокоселективны и по ориентации стимулов (De Valois & De Valois, 1988).

Дальнейшие физиологические исследования показали, что корковое расположение клеток, настроенных к различным пространственным частотам весьма систематично, в частности, внутри гиперколонки они явно расположены по направлению, перпендикулярному направлению расположения у клеток селективных (De Valois & De Valois, 1988). То есть размерность пространственной частоты и размерность ориен-

тации в буквальном смысле задают базис двумерного пространства внутри каждой гиперколонки.

Де Валоисы показали, что размерность пространственной частоты и размерность пространственной ориентации в стриарной коре кошек лежат в декартовой системе координат (рис. 4.2.22 А), но у обезьян, полагали исследователи, архитектура немного иная: клетки, настроенные на ориентацию стимула и на его пространственную частоту, расположены в полярном координатном пространстве (рис. 4.2.22 В), то есть — ориентация представлена направлением от центра гиперколонки, а пространственная частота — расстоянием от центра.

Итак, хотя доказательство того, что простые и сложные клетки в V1 способны к локальному анализу пространственных частот ретиальной стимуляции выглядит весьма эффектно, оно, к сожалению, не универсально. Вместе с тем, применение теории локальных пространственных частот привело к ряду очень интересных и важных открытий в контексте функции нейронов стриарной коры и выступает в роли весьма серьезной альтернативы концепции Хьюбела — Визеля.

Интересно и познавательно отследить отношения между обеими теориями. Теории эти взаимно-конкурентны, поскольку их функциональное исполнение различно: концепция пространственных частот утверждает, что нейроны V1 не являются «детекторами» тех или иных свойств ретиального изображения, в частности, линий и краев, но играют роль анализаторов (или как иногда говорят — *фильтров*), декомпозирующих

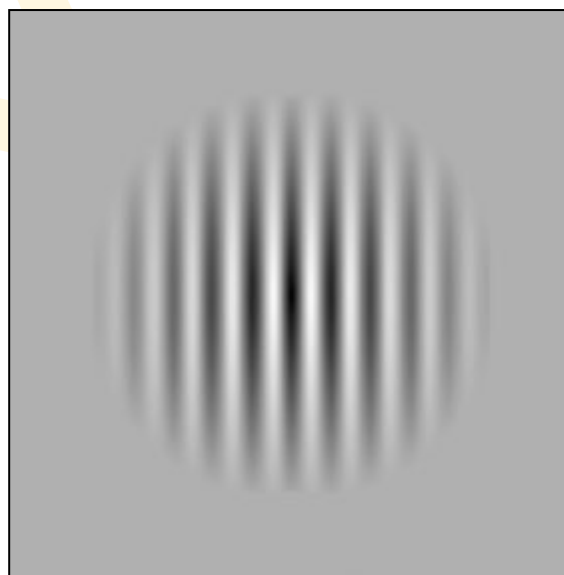
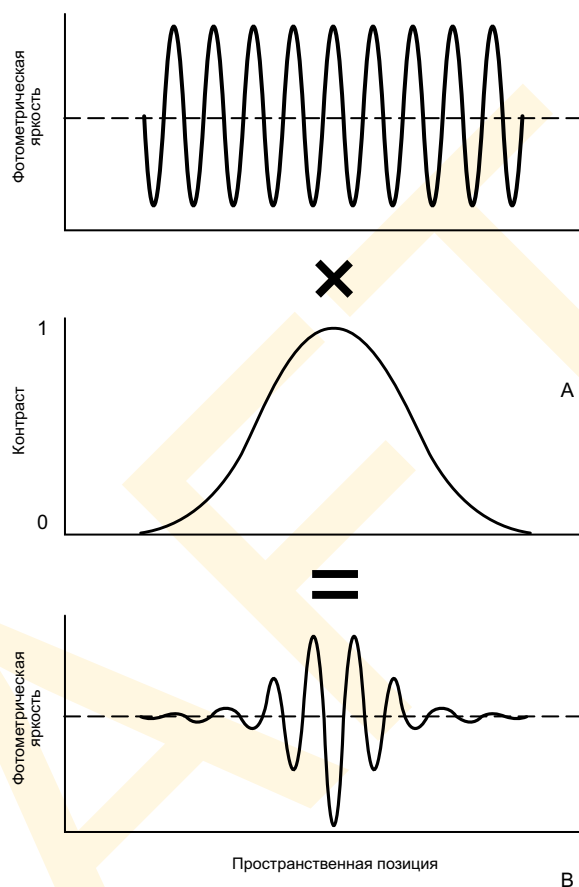
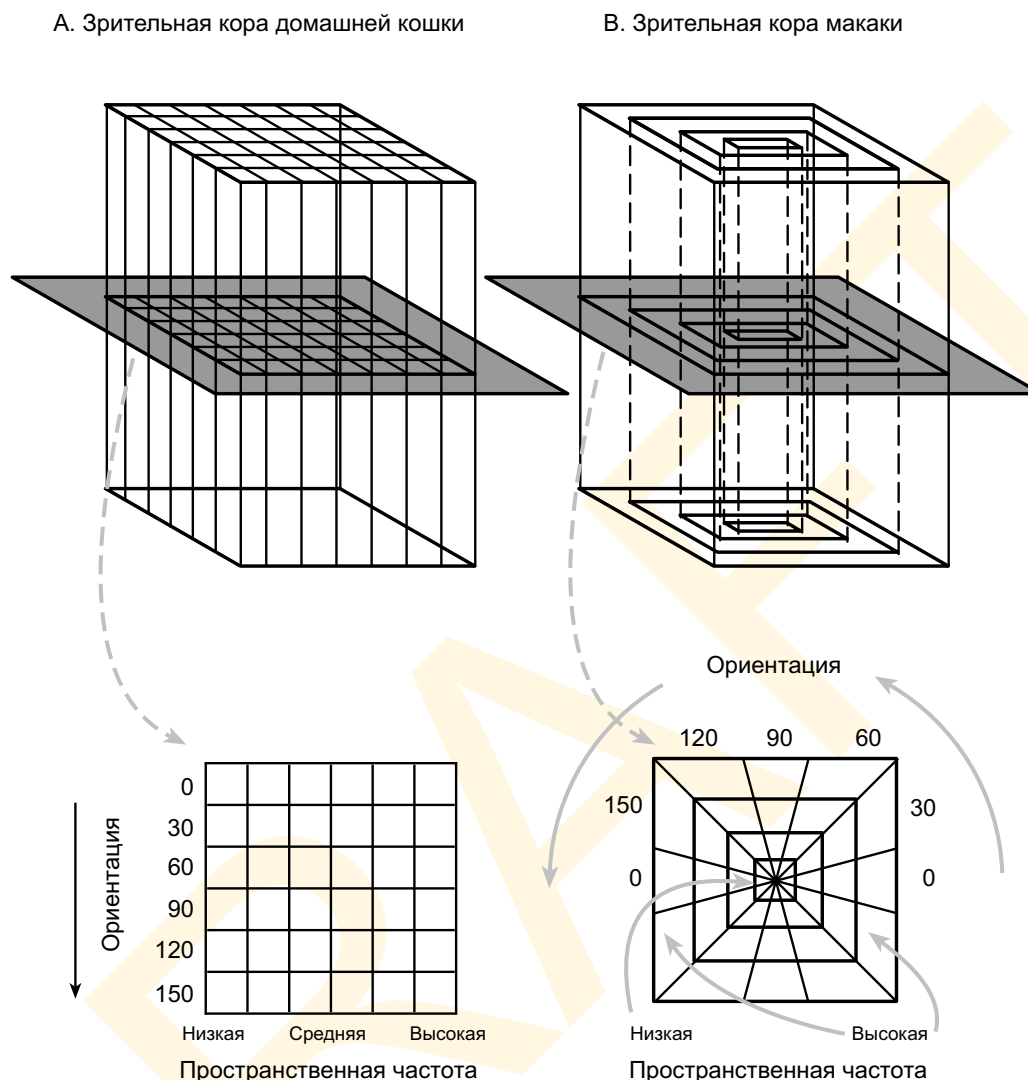


Рис. 4.2.21 А — габорова функция, построенная умножением синусоиды на гауссову кривую. В — результирующая картина яркостей.





**Рис. 4.2.22** Две модели корковой гиперколончатой архитектуры коры домашней кошки и макаки, включающие в себя размерность пространственной частоты. В модели кошачьей коры размерность пространственной частоты и размерность ориентации ортогональны в декартовом пространстве; в модели коры макаки эти размерности радиальны в полярном пространстве.

По материалам De Valois & De Valois, 1988.

изображение на отдельные примитивы, что позволяет кратчайшим путем описать любое изображение. При этом данная точка зрения отнюдь не исключает существования в зрительной системе клеток — детекторов краев и линий. Скорее, она просто переносит их на следующий уровень, где соответствующие габоровы фильтры идентифицируют край или линию.

Таким образом, теория локальных пространственных частот потенциально совместима концепцией краевой детекции, но вне тезиса о том, что детекторы эти воплощены в каких-то клетках V1. К сему противоречию мы вернемся позднее — после того как разберем то новое, что дадут нам количественные подходы к обработке ретинального изображения.

### 4.3 КОЛИЧЕСТВЕННЫЕ МОДЕЛИ ПСИХОФИЗИЧЕСКИХ КАНАЛОВ

Количественные исследования происходящего на оптическом этапе зрительного акта шли по нескольким направлениям. Частично они были посвящены изучению механизмов детекции естественных компонент ретинального изображения, в частности таких как края и линии. Многие из известных специалистов — приверженцев данного подхода — трудились в команде по изучению вопросов зрительного восприятия Массачусетского технологического института, в частности: Дэвид Марр, Томазо Погжио, Элен Хилдрет, Симон Ульман и многие другие. Их исследования в основном были нацелены на компьютерную реализацию марровых опорных скетчей, которых мы коротко коснулись в гл. 2. Команда получила ряд очень важных результатов, касаемых количественного и алгоритмического описания (в марровом смысле) проблемы детекции линий и краев. Отметим, что работа плотно соотносима с догадкой Хьюбела и Визеля о том, что функция стриарных клеток — в детекции краев и линий.

Вместе с тем, имеются альтернативные точки зрения и одна из них отстаивает т.н. *фильтровый подход* к начальным этапам пространственного зрения, подход, основанный преимущественно на теории пространственных частот (см. предыдущий раздел). Исследователи — приверженцы фильтрового подхода, такие, в частности, как Адельсон и Берген (1985), Хигер (1988), Кёндеринк и Ван Дорн (1976а), а также Малик (Jones & Malik, 1992; Malik & Perona, 1990) изучали количественные аспекты применения т.н. *мультиориентированных и мультишкальных фильтров* (габоровы функции), выделяющих из ретинального изображения пространственные примитивы, которыми зрительная система оперирует на более высоких уровнях обработки зрительной информации.

Третья группа исследователей образовала лагерь коннекционистов, занявших исключительно оригинальную позицию по отношению к вопросу пространственной обработки ретинальных изображений. Они применили т.н. *методы полнофункционального обучения* (например, метод обратного распространения, описанный в Приложении В), позволяющие нейронным сетям «самопрограммироваться» для выполнения четко определенных перцепционных задач. Единоразово обучив сеть отрабатывать некое задание, исследователи подвергали такую сеть тщательной проверке. Одна из сетей строилась из нейроноподобных элементов с «рецептивными полями», которые оказались на удивление схожи с рецептивными полями стриарных клеток Хьюбела — Визеля. Результаты указали на некоторые новые и неожиданные возможности функционирования корковых нейронов вне теории детекции линий и краев, а также теории пространственных частот.

Однако дабы выдержать единую методическую линию изложения, в данном разделе мы обсудим модели пространственного зрения, основанные на теории пространственных частот, в следующем поговорим о проблемах компрессии пространственной информации с целью ее эффективной передачи по проводящим путям в высшие отделы нервной системы, после чего перейдем к разговору о марровой концепции опорных скетчей. Завершит текущую главу попытка синтеза действующих теорий пространственного зрения.

#### 4.3.1 Паттерн-чувствительность

В текущем разделе мы рассмотрим способы диагностики чувствительности зрительной системы человека к пространственным и временным паттернам, а также во-

просы моделирования этой чувствительности. И, хотя мы уже касались этих вопросов ранее, сейчас мы расширим наш анализ, описав психофизические исследования, призванные выявить то, как зрительная система в целом детектирует и различает пространственно-временные паттерны стимуляции.

Сразу отметим, что в научной литературе, посвященной пространственному зрению, в основном описаны эксперименты по детекции и различению, но не опыты, в которых наблюдатель рассказывает о том, как выглядит та или иная вещь. И тому есть минимум три причины.

Во-первых, множество систем визуализации (телевизионных, печатающих и пр.) способны воспроизводить изображения, которые весьма схожи с оригиналом, но не точно такие же. Вопрос визуальной аппроксимации изображением оригинальной сцены крайне важен и часто является основополагающим в работе разработчиков систем визуализации. Следовательно, имеется определенный интерес к созданию полноценной теории, прогнозирующей то, что изображение и сцена (или два изображения) окажутся визуально подобны друг другу. Такая теория очень поможет конструкторам воспроизводящих аппаратов — она даст ответ на вопрос, изображения какого характера данное устройство будет воспроизводить максимально близко к оригиналу? Для создания такой теории нужны пороговые эксперименты и эксперименты на различение.

Во-вторых, многие авторы полагают, что пороговая задача и задача различения могут играть специальную роль в анализе нейрофизиологических механизмов зрения.

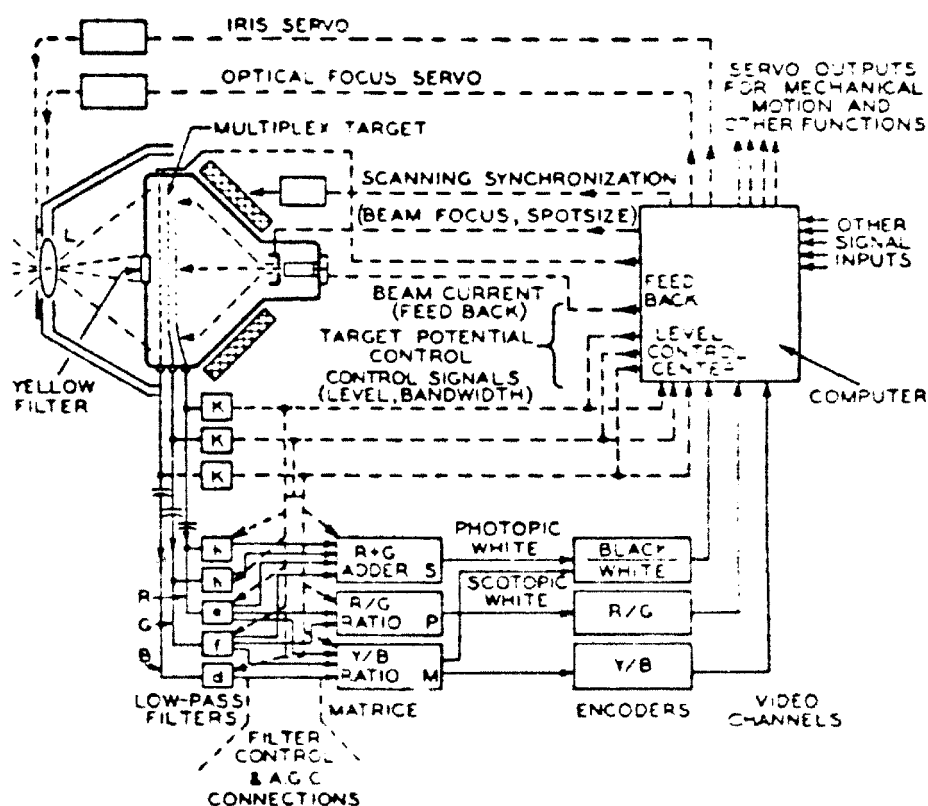
Предположим, что нервная система построена из множества компонент анализирующих пространственную световую картину на сетчатке; следовательно, мы должны проанализировать эти компоненты, дабы понять, каков их вклад в восприятие пространственных паттернов. Показатель порога дает нам наибольший шанс выделить эти компоненты, поскольку на уровне порога лишь наиболее чувствительные механизмы реализуют видность объекта. Если показатель порога восприятия зависит от стимуляции единичного зрительного механизма или малого числа этих механизмов, то пороговые исследования дают в руки физиологов весьма мощный инструмент: на уровне порога мы можем выделять различные фрагменты зрительных путей.

И все же у нас нет полной уверенности в том, что приведенное обоснование целиком адекватно пафосу пороговых измерений. Впрочем, кое в чем нам уже удалось разобратся (De Valois & De Valois, 1988; Graham, 1989), и многие из основополагающих идей, используемых в толковании представления ретинального изображения, а также в сфере машинного зрения, выросли из работ по детекции и распознаванию пространственных паттернов.

### 4.3.2 Теория моноразрешающего представления

Вопрос прогнозирования чувствительности зрительной системы человека к контрастным паттернам ничуть не менее важен, чем все остальное уже нами пройденное: свойства ретинального изображения, цветовое восприятие, отклики отдельных нейронов и пр. Перед нами встала очень серьезная задача: научиться уверенно прогнозировать чувствительность к самым разнообразным пространственным контрастным паттернам по измерениям чувствительности к небольшому числу пространственных стимулов.

В 1956 г. немецкий исследователь Отто Шейд подал идею о том, как выполнять такие прогнозы и как построить фотоэлектрический аналог зрительной системы. Идея



**Рис. 4.3.1** Шейдова электронная модель зрительной системы человека. Отто Шейд сконструировал зрительный имитатор, прогнозирувавший чувствительность зрительной системы человека к контрастным паттернам. Аппарат состоял из множества узлов, имитирующих зрительные пути, и, вероятно, являлся одной из первых (если не самой первой) моделью человеческого зрения.

По материалам Schade, 1956.

состояла в применении некоего устройства, прогнозирующего то, какие минимальные изменения в параметрах визуализации будут распознаваться зрительной системой человека. К примеру, при разработке новых компьютерных мониторов и телевизоров конструкторы, разумеется, стремятся повысить пространственное разрешение устройства. Для ответа на вопрос будет ли это отличие визуально заметным, Шейд предлагает механизм прогнозирования чувствительности зрительной системы человека к изменениям, вносимым разработчиками в новый монитор.

На рис. 4.3.1 дана принципиальная схема шейдового «компьютерного глаза», дабы читатель мог убедиться, что ученый создал мощную модель, исполнявшую множество зрительных функций, в частности таких как формирование оптического «ретинального» изображения, преобразование и интеграцию сигналов от отдельных «колбочек» разных типов, адаптацию, пространственную интеграцию, отрицательную обратную связь и даже некоторые виды коррекции, интерпретации и корреляции сигналов с информацией хранящейся в памяти. Словом, работа Шейда послужила предтечей компьютерных моделей зрения, предложенных спустя время Дэвидом Марром (1982).

#### 4.3.2.1 Нейронное изображение

Базовым положением шейдовой концепции (и первым, что следует обсуждать в разговоре о пространственном зрении) является тезис о том, что мы вправе просуммировать эффекты от множества зрительных компонент, используя единичное представление, ныне называемое *нейронным изображением*. Оптическое ретинальное изображение и пара нейронных схематично даны на рис. 4.3.2. Идея состоит в том, что существует группа нейронов, чьи отклики, взятые в совокупности, реализуют полноценный захват визуальной информации из ретинального изображения. Отклики группы нейронов со схожими рецептивными полями и отличные друг от друга лишь тем, что поля эти центрованы по разным позициям, образуют нейронное изображение.

На рис. 4.3.2 показано, как концепция нейронного изображения позволяет нам реконструировать изображение ретинальное из нейронных откликов: в нескольких местах рисунка нейронные отклики представлены в виде обычного яркостного изображения; интенсивность в каждой точке этих изображений представляет отклик единичного нейрона, рецептивное поле которого центрировано по соответствующей точке изображения ретинального. Светлые участки представляют рост отклика в ответ на стимуляцию, темные участки — падение.

Предположения, нами сделанные в отношении свойств рецептивных полей нейронов, входящих в состав нейронного изображения, позволяют нам прогнозировать нейронное изображение, используя линейные методы. Положим, к примеру, что рецептивные поля группы нейронов идентичны друг другу за исключением пространственной позиции их центров. Предположим также, что распределены эти поля равномерно. В этом случае мы можем рассчитать мэппинг оптического ретинального изображения на отклики этих нейронов, используя методику т.н. *шифт-инвариантного линейного мэппинга (свертку)*.

Нейронные изображения с рис. 4.3.2 иллюстрируют ту мысль, что разные популяции нейронов представляют разные типы информации.

Одно из нейронных изображений передает грубую структуру оригинала (показатели этого изображения получены путем локального усреднения оригинальных яркостей). Среднюю локальную интенсивность, представленную в этом изображении, можно использовать для разных целей, например, для регулировки уровня адаптации к вариациям освещенности сцены.

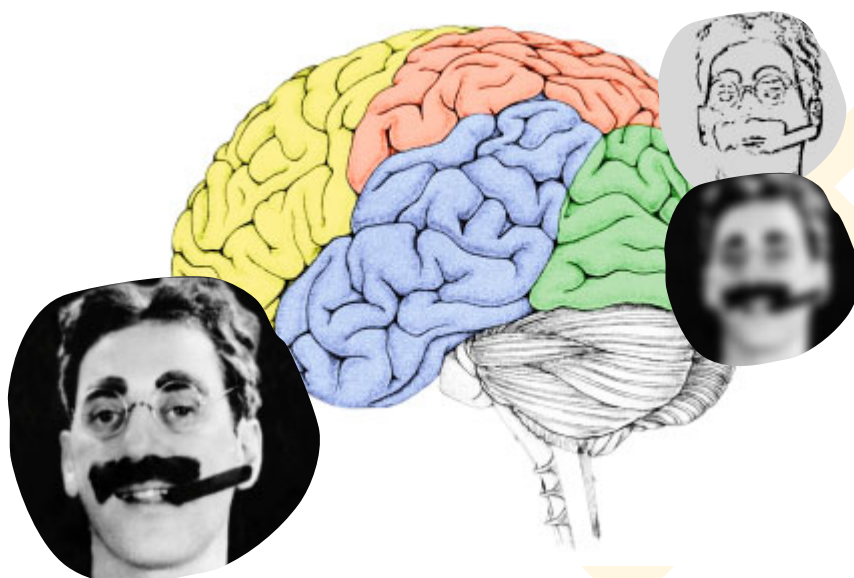
Второе нейронное изображение, представляющее картину краев, можно использовать в работе системы распознавания.

Иными словами, в зрительной коре существует множество самых разнообразных нейронных изображений, представляющих разные аспекты изображения ретинального, включая хроматику, движение и объем. Разумеется, информация, представленная в корковом нейронном изображении зависит от преобразований сигнала, имеющих место по всей протяженности зрительного тракта, включая расфокусировку хрусталиком, дискретизацию колбочковой мозаикой, колбочковое кодирование и т.д.

#### 4.3.2.2 Шейдова теория моноразрешающего представления

Шейдова теория паттерн-чувствительности сформулирована в контексте фовеального зрения, т.е. зрения, основанного на информации, захваченной центральной ямкой сетчатки. Шейд предположил, что фовеальную пространственную чувствительность можно прогнозировать по информации, извлекаемой из единого общего нейронного изображения. Он предположил также, что по данному участку поля зрения реле-





**Рис. 4.3.2** Нейронное изображение представляет собой совокупную активность группы нейронов. Оптическое ретинальное изображение слева кодируется множеством отдельных популяций нейронов мозга. Интенсивность в каждой точке нейронного изображения (справа) представлена активностью единичного нейрона в популяции. Каждая из нейронных популяций может кодировать различные свойства оригинала. Нейронное изображение справа сверху представляет картину краев; изображение справа внизу кодирует локальную среднюю яркость, которая может играть адаптирующую роль. Прочие изображения (не представленные на иллюстрации) могут нести информацию об окраске, движении и объеме.

вантное нейронное изображение можно представить шифт-инвариантным преобразованием изображения ретинального, подобно нейронному изображению, показанному снизу на рис. 4.3.2. В данном разделе мы поговорим о роли шейдовой гипотезы в науке о зрении и эмпирической апробации этой теории.

Мы знаем, что нейронное изображение, целиком покрывающее поле зрения, на самом деле не может быть шифт-инвариантным. Ранее мы рассматривали результаты исследований, наглядно демонстрирующих то, что в центральной ямке намного больше фоторецепторов и ганглиозных клеток, чем на периферии сетчатки и что в зрительной коре представлению фовеальной информации отведено намного больше места, нежели информации периферической. Следовательно, нейронное изображение может иметь шифт-инвариантное представление только в относительно малом участке поля зрения, скажем, внутри центральной ямки или небольшом фрагменте периферии.

Итак, действительно, имеет смысл начинать с разговора именно о фовеальной модели распознавания пространственно-контрастных паттернов: во-первых, теория оказывается довольно простой, потому что мы можем уйти от сложностей связанных с неоднородностью поля зрения; во-вторых, поскольку глаза наши все время находятся в движении — центральная ямка сетчатки, это главный источник пространственной информации.

Существует несколько способов применения гипотезы шифт-инвариантных нейронных изображений, в прогнозировании контрастной чувствительности зрения. Пожалуй, наиболее толковой представляется идея, о которой мы говорили ранее: если

мэппинг оптического изображения на изображение нейронное шифт-инвариантен, то характер этого мэппинга задается формой рецептивного поля. При этом мы помним, что у шифт-инвариантного нейронного изображения существует лишь одна возможная базовая форма рецептивного поля как примитива этого изображения — рецептивные поля нейронов, формирующих нейронное изображение, разнятся лишь своей пространственной позицией.

Стоит сказать, что аналогия между шифт-инвариантными вычислениями и рецептивными полями нейронов весьма практична, но при этом мы должны помнить, что рассуждаем в логике реактивных (психофизических) измерений, но не реальных рецептивных полей. Следовательно, удобнее выражать наши оценки, используя несколько более абстрактный язык линейных преобразований. При этих преобразованиях линейное рецептивное поле эквивалентно т.н. *ядру свертки* шифт-инвариантного мэппинга. Гипотеза шифт-инвариантности указывает на то, что для оценки характера нейронного изображения мы должны позиционировать ядро свертки, параметры которого определяют, какая информация представлена в нейронном изображении, а какая нет.

Начнем с анализа одномерных стимулов — вертикальных синусоид, варьирующих только по  $x$ -направлению. Если на входе мы используем только одномерные стимулы, то мы можем позиционировать лишь одномерное рецептивное поле и записать шифт-инвариантное преобразование, проецирующее одномерный контрастный стимул ( $a_x$ ) на одномерное нейронное изображение ( $n_x$ ):

$$n_x = \sum_{(y)} l_{x-y} a_y$$

где  $l_x$  — одномерное рецептивное поле. Данное преобразование мы можем выразить и в матричной форме:

$$\begin{pmatrix} n_0 \\ n_1 \\ \dots \\ n_{N-1} \\ n_N \end{pmatrix} = \begin{pmatrix} l_0 & \dots & l_{-N/2} & \dots & l_{-N} \\ \dots & \dots & \dots & \dots & \dots \\ l_{N-2} & \dots & l_0 & \dots & l_{N/2} \\ \dots & \dots & \dots & \dots & \dots \\ l_N & \dots & l_{N/2} & \dots & l_0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_{N-1} \\ a_N \end{pmatrix}$$

Матрица системы весьма проста: ряды и колонки полностью равны рецептивному полю (т.е. ядру свертки) за исключением сдвига или инверсии. Следовательно, позиционировав ядро свертки, мы сможем прогнозировать результат преобразования из оптического контрастного изображения на сетчатке в изображение нейронное.

Общая схема прогнозирования пространственной чувствительности зрительной системы выглядит следующим образом:

— во-первых, мы оцениваем чувствительность к набору гармонических контрастных паттернов, и результатом является контрастно-сенситивная функция зрительной системы наблюдателя. Поскольку специфика взаимоотношений между гармоническими функциями и шифт-инвариантными линейными системами уже описана нами ранее, мы будем использовать контрастно-сенситивную функцию для позиционирования ядра свертки шифт-инвариантного линейного преобразования из оптического изображения на сетчатке в изображение нейронное ( $l_x$ );

— во-вторых, полученное ядро мы используем для расчета нейронного изображе-

ния и прогнозирования чувствительности зрительной системы к прочим одномерным контрастным паттернам. Сей завершающий этап поможет нам оценить добротность нашей теории.

Итак, мы начинаем понимать, что должны сделать ряд дополнительных предположений, прежде чем сможем использовать контрастно-сенситивные данные для установки ядра свертки.

#### 4.3.2.3 Пространственные контрастно-сенситивные функции

Шейд диагностировал психофизическая контрастно-сенситивные функции, прося наблюдателей оценить видность синусоидальных паттернов различного контраста. То есть, шейдовы испытуемые ловили минимальный фотометрический контраст решеток, при которых последние становятся едва различимыми. Из-за ряда оптических и нейронных факторов наблюдатели оказались неравночувствительны ко всем пространственно-частотным решеткам: порог контраста, как выяснилось, зависит от пространственной частоты паттерна.

Чтобы лучше понять характер шейдовых экспериментов (1956), обратимся к его же тексту:

«Тестовая решетка плавно усиливалась повышением электромодуляции на определенной частоте и одновременным наблюдением за показаниями модулометра. В тот момент, когда испытуемый начинал различать ее, он подавал сигнал, а экспериментатор фиксировал показание прибора. Модуляция возвращалась в ноль и вновь повышалась на той же пространственной частоте. Усреднив 10-15 попыток мысленно и записав среднее непосредственно на миллиметровку, контрастно-сенситивную функцию удавалось получить в течение короткого времени. Все нестыковки при этом проявлялись сразу и устранялись дополнительными наблюдениями.»

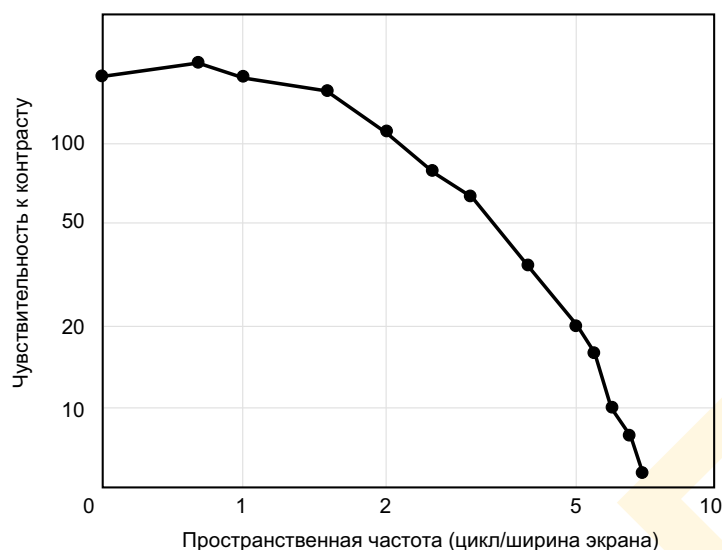
Контрастно-сенситивная функция, выведенная Шейдом, дана на рис. 4.3.3: по абсциссам отложена пространственная частота в циклах *на ширину экрана*; по ординатам — чувствительность к контрасту, а именно  $\log(1/c) = -\log c$ , где  $c$  — это контраст решетки на уровне порога различения.

Отметим, что у контрастно-сенситивной функции имеются два замечательных признака:

— во-первых, чувствительность падает по мере роста пространственной частоты стимулирующей решетки. Сей эффект велик, но не удивителен, поскольку мы уже знаем о множестве различных компонентов в зрительных путях, нечувствительных к стимулам высокой пространственной частоты (оптическое размытие хрусталиком понижает контраст решеток высокой пространственной частоты, ганглиозные клетки с рецептивными полями типа центр/окружение менее чувствительны к высоким частотам и пр.).

— во-вторых (что уже несколько странно), мы не наблюдаем роста чувствительности на низких пространственных частотах, более того — даже ее легкую потерю на самых низких частотах. Оптическая система глаза не понижает чувствительность к низким частотам сцены, так что снижение это целиком обязано нейронным факторам, и рецептивные поля типа центр/окружение — единственно возможное объяснение этого эффекта.

Шейдовы измерения выполнялись с применением т.н. *монотонных* или *дрейфующих решеток*. В 1966 г. поминавшийся ранее Джон Робсон (1966; Kelly, 1961) выполнил дополнительные измерения, используя мигающие *контрастно-реверсивные ре-*



**рис. 4.3.3** Психофизическая контрастно-сенситивная функция, выведенная Шейдом. Чувствительность к контрасту дана как функция от пространственной частоты в циклах на ширину экрана (а не на градус, как принято).

По материалам Schade, 1956.

$$[1.0 + a \cos(2\pi f_t t) \cos(2\pi f_x x)]m$$

Отметим, что интенсивность всегда положительна, тогда как пространственно-временной контраст паттерна — это:

$$a \cos(2\pi f_t t) \cos(2\pi f_x x),$$

и он может быть как положительным, так и отрицательным.

Как видим из рис. 4.3.3 пространственная чувствительность к низким частотам падает, когда измерение проводится на низких временных частотах (1 Гц). На высоких временных частотах, с которыми мы постоянно сталкиваемся благодаря быстрым глазным движениям, снижения чувствительности к низким пространственным частотам не отмечается. Как мы увидим далее, контрастно-сенситивная функция меняется при смене иных параметров стимуляции, в частности таких как средний уровень фотометрической яркости и спектральный состав стимула.

#### 4.3.2.4 Психофизическая линейно-расставленная функция

Вернемся теперь к вопросу применения контрастно-сенситивных данных при вычислении ядра свертки ( $l_x$ ). Поскольку ядро это задает ряды и колонки шифт-инвариантного линейного преобразования, ядро свертки часто называют *психофизической линейно-расставленной функцией* (по аналогии с *оптической линейно-расставленной функцией*).

Несмотря на то, что до сего момента мы постоянно пользовались теорией линейных систем, некоторые специфические моменты измерений требуют от нас применения не-

шетки.

Контрастно-реверсивные решетки — это гармонические пространственные паттерны с гармонически меняющейся амплитудой. К примеру, предположим, что средний уровень освещенности равен  $m$ . А интенсивность

контрастно-реверсивного стимула на пространственной частоте  $f_x$ , временной частоте  $f_t$  и при контрасте  $a$  равна:

сколько иных подходов. Вычисления, задействующие

контрастно-сенситивные функции для прогнозирования чувствительности ко всем возможным контрастным паттернам — не исключение.

Итак, теперь необходимо определить, что нам нужно для позиционирования

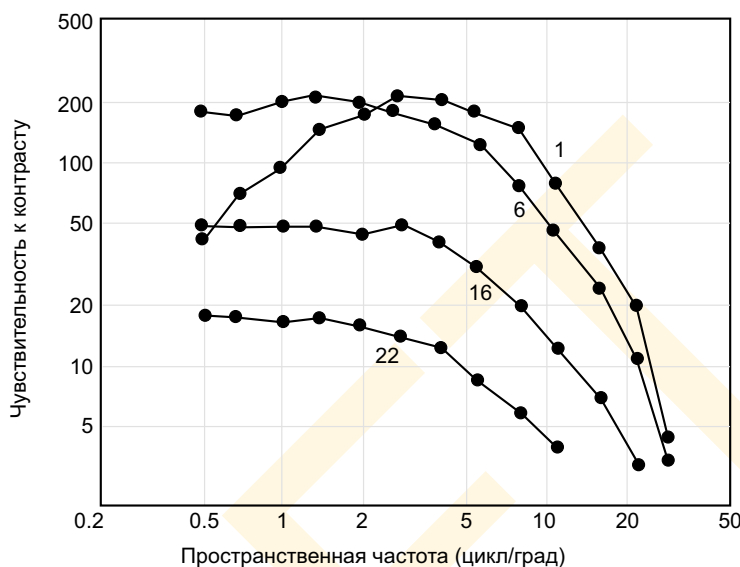
линейно-расставленной психофизической функции. Главная линейная проблема иллюстрируется матричным представлением (см. ранее), где входной стимул дан как вектор-столбец, описывающий одномерный пространственно-контрастный образец. Матрица описывает то, как стимул преобразуется в нейронное изображение. Напомним, что наша задача — по малому числу измерений получить коэффициенты этой матрицы.

Мы уже решали подобную проблему ранее, но в данном случае у нас специальная задача. В прежних попытках установить линейное преобразование мы могли описать и сам стимул, и отклик на него, однако когда мы выполняем контрастно-сенситивную диагностику, мы не можем анализировать выходное нейронное изображение. Мы лишь измеряем пороги на входе, и, следовательно, на руках у нас весьма и весьма ограниченный объем информации.

Поскольку мы предположили, что происходящее — это шифт-инвариантный мэппинг, в нашем распоряжении уже есть кое-какая информация о нейронном изображении: коль скоро входной стимул представляет собой гармоническую функцию — на выходе должна быть гармоническая функция той же частоты. Однако мы не знаем при этом ни амплитуды, ни фазы гармонической функции в гипотетическом нейронном изображении. Для установки линейно-расставленной психофизической функции мы должны сделать ряд дополнительных предположений о свойствах нейронного изображения. Из этих предположений мы выведем его амплитуду и фазу.

Понадобятся два дополнительных предположения:

— во-первых, мы допускаем, что пространственная фаза нейронного изображения та же, что и пространственная фаза входного пространственного контрастного паттерна. В частности, когда входной стимул представляет собой одномерную косинусоиду —  $\cos(2\pi fi/N)$ , мы предполагаем, что выходной паттерн нейронного изображения —



**Рис. 4.3.4** Темопральные вариации, меняющие форму пространственной контрастно-сенситивной функции. Функции, показанные на этой иллюстрации строились по контрастно-реверсивным мишеням на нескольких временных частотах. На низких временных частотах (1 Гц) контрастно-сенситивная функция представляет собой полосовой фильтр. На высоких (22 Гц) — фильтр нижних частот.

По материалам Robson, 1966.



это шкалированная копия паттерна входного:  $n_i = a_i \cos(2\pi fi/N)$ . Шкалирующий коэффициент ( $a_i$ ) зависит от частоты входного сигнала.

— во-вторых, мы должны выставить амплитуду нейронного изображения при детекции порога. Амплитуду нейронного изображения следует относить к видности (различимости) стимулирующего паттерна, и мы должны прописать ряд свойств, которые следует связать с этой различимостью. К примеру, независимо от того растет или падает внесенное сигналом изменение, любые отклонения от частоты спонтанной пульсации должны быть детектируемы. Плюс к тому, детектируемость зависима скорее от результирующей картины откликов по всему нейронному изображению, нежели от отклика единичного нейрона.

Мерой обоих параметров явится т.н. *квадрат размерности вектора* откликов нейронного изображения. Квадрат размерности вектора нейронного изображения ( $d^2$ ) определен уравнением:

$$d^2 = \sum_{i=1}^N n_i^2$$

Данное уравнение удовлетворяет обоим нашим требованиям поскольку, во-первых, знаки отдельных нейронных откликов ( $n_i$ ) не имеют значения (потому что входные показатели нейронного изображения возведены в квадрат), и во-вторых, поскольку уравнение включает в себя отклики различных нейронов.

Ежели мы предположим, что на пороге детекции контраста все нейронные изображения имеют одну и ту же размерность вектора, то мы сможем специфицировать множество гармонических функций в нейронном изображении. Следовательно, в данный момент наших допущений достаточно для спецификации полного нейронного изображения, то есть получения психофизической линейно-расставленной функции.

На рис. 4.3.5 показаны три линейно-расставленных функции, полученные при этих допущениях. На рис. 4.3.5 А дан психофизический линейный расстав, вычисленный из шейдовых данных (отметим, что пространственная размерность не откалибрована). На рис. 4.3.5 В и 4.3.5 С показаны психофизические расставы, вычерченные в показателях градусов зрительного угла и полученные из робсоновских измерений с однопериодными и шестипериодными контрастно-реверсивными функциями.

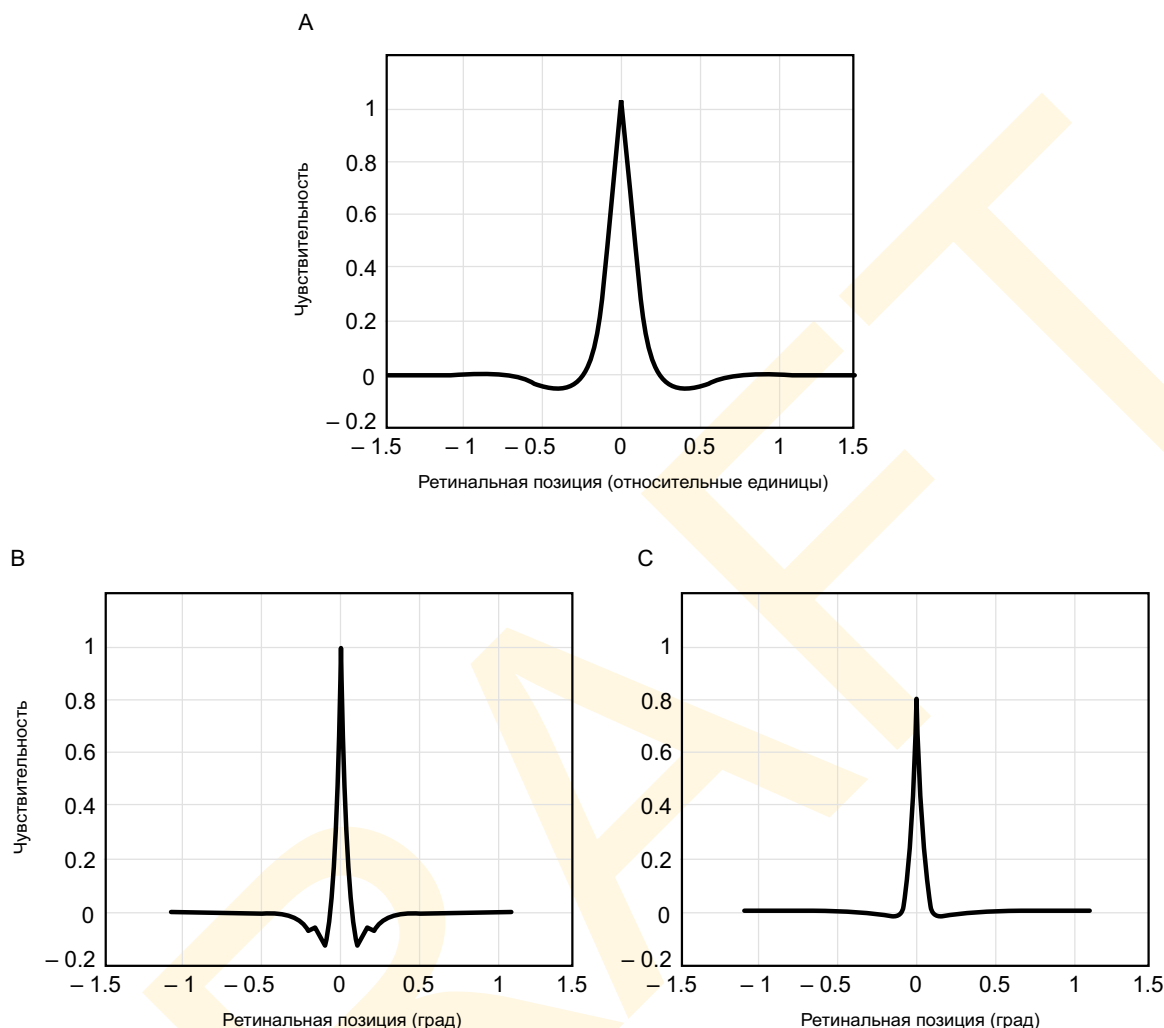
Следует сказать, что не существует единой линейно-расставленной функции, применимой ко всем условиям стимуляции (то, как меняется линейно-расставленная функция при изменении условий стимуляции, мы обсудим чуть позднее).

Шейд (1956) показал, что общую форму психофизической линейно-расставленной функции можно описать, используя отличия между двумя гауссовыми функциями. Такое описание использовали Родик (1965), а также Энрот-Кугель с Робсоном (1966) при моделировании рецептивных полей ретинальных ганглиозных клеток, и отметим, что достигнутое ими соответствие между психофизической линейно-расставленной функцией (полученной при психофизических тестах) и функциями рецептивных полей ганглиозных клеток сетчатки (полученных физиологически) весьма воодушевляет.

#### 4.3.2.5 Теоретический аспект: статические нелинейности

Вспомним, что когда нам потребовалось определить ядро свертки ретинального изображения, дабы получить шейдово гипотетическое нейронное (не прибегая при этом к прямому нейронному измерению, нам пришлось сделать ряд допущений):

1. Нейронные изображения — плод шифт-инвариантного линейного кодирования.



**рис. 4.3.5** Психофизическую линейно-расставленную функцию можно вывести из контрастно-сенситивной функции. А — линейно-расставленная функция, полученная из шейдовых данных. По оси абсцисс даны относительные единицы (потому что Шейд дал пространственную частоту в относительных единицах — см. рис. 4.3.3). В и С — соответственно линейно-расставленные функции, построенные по робсоновским контрастно-реверсивным решеткам с частотой 1 Гц и 6 Гц (по оси абсцисс отложены градусы зрительного угла).

2. Сдвиг фазы при этом кодировании равен нулю.

3. Правило размерности вектора определяет видность.

Однако, взятая в целом, сие есть *нелинейная* теория пространственной чувствительности: в самом деле, несмотря на то, что нейронное изображение — это линейное представление входной оптической информации (более того — шифт-инвариантное представление), правило размерности вектора, связывающее нейронное изображение с его клеточным воплощением — нелинейно. Мы можем убедиться в этом, отметив, что когда вектор стимуляции  $s_1$  имеет длину  $d_1$ , а вектор  $s_2$  длину  $d_2$ , то длина суммы векторов  $s_1 + s_2$  не равна  $d_1 + d_2$ . Таким образом, даже когда  $s_1$  и  $s_2$  на полупороге, сумма этих векторов не достигает порога детекции.

Расчет длины вектора выполняется после линейных вычислений и с учетом той са-

мой статической нелинейности. Однако сие относительно простая нелинейность и поэтому прогнозы незатруднительны. В следующем разделе мы рассмотрим некоторые из этих прогнозов, равно как и методику их экспериментальной проверки.

#### 4.3.2.6 Экспериментальная проверка

Контрастно-сенситивные функции сами по себе не являются тестом шейдовой теории, если не считать того, что они в целом похожи на правду, поскольку весьма напоминают рецептивные поля зрительных нейронов. Но коль скоро мы сделали множество разных допущений, важно найти главные параметры теории, которую мы хотим проверить экспериментально и тем самым лишний раз убедиться в ее практической применимости.

Гармонические функции играют особую роль в такой проверке и минимум по двум причинам:

- приняв тезис о шифт-инвариантности, видим, что ежели стимул гармонический, то и нейронное изображение гармоническое;
- гармонические функции на различных частотах ортогональны одна другой. С геометрических позиций ортогональность означает, что векторы ориентированы перпендикулярно один другому; с алгебраических — мы скажем, что два вектора  $\mathbf{a}$  и  $\mathbf{b}$  ортогональны тогда, когда  $0 = \sum a_x b_x$ . Синусоида и косинусоида ортогональны одна другой и в любой паре гармонических функций на разных частотах эти функции взаимно ортогональны. Для проверки шейдовой теории мы будем использовать оба этих свойства, объединив их с правилом размерности вектора.

Предположим, мы воспроизвели стимул равный сумме двух синусоид на частотах  $f_i$  и контрастами  $c_i$  для  $i=1, 2$  (рис. 4.3.6). Согласно шифт-инвариантной концепции нейронное изображение этих двух синусоид — это взвешенная сумма этих синусоид. Каждая синусоида шкалируется коэффициентом ( $s_i$ ), что определяет качество передачи стимула шифт-инвариантной системой. Квадрат размерности вектора нейронного изображения, получаемый квадратом суммы двух синусоид — это:

$$d^2 = \sum_{i=1}^N \left[ c_1 s_1 \sin(2\pi f_1 i/N) + c_2 s_2 \sin(2\pi f_2 i/N) \right]^2$$

Квадрат суммы можно представить как:

$$d^2 = \left[ c_1 s_1 \sum_{i=1}^N \sin(2\pi f_1 i/N) \right]^2 + \left[ c_2 s_2 \sum_{i=1}^N \sin(2\pi f_2 i/N) \right]^2 + 2c_1 c_2 s_1 s_2 \sum_{i=1}^N \sin(2\pi f_1 i/N) \sin(2\pi f_2 i/N)$$

Поскольку синусоиды разных частот — это ортогональные функции, третий компонент равен нулю, то есть, остается:

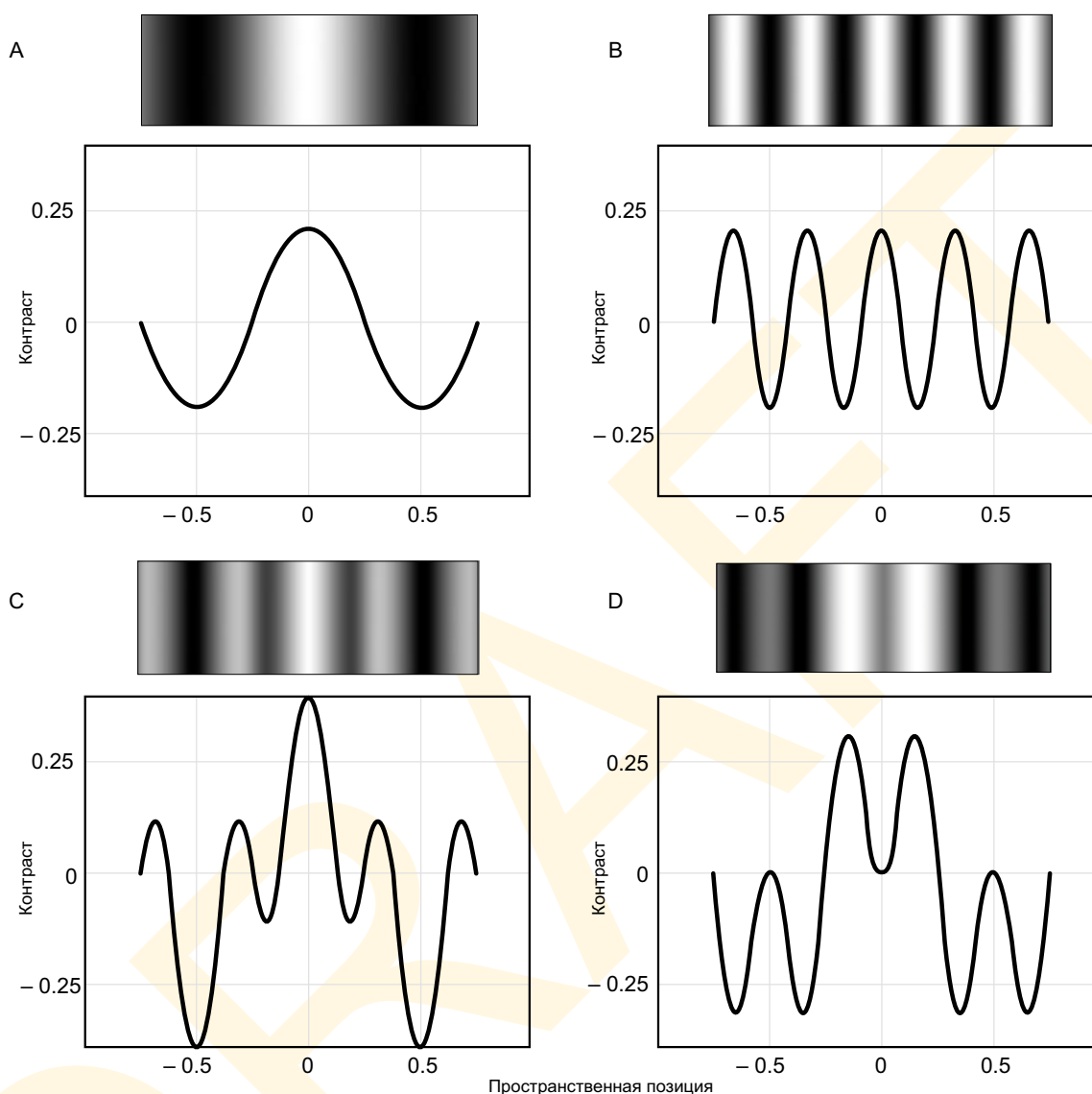
$$d^2 = \left[ c_1 s_1 \sum_{i=1}^N \sin(2\pi f_1 i/N) \right]^2 + \left[ c_2 s_2 \sum_{i=1}^N \sin(2\pi f_2 i/N) \right]^2$$

Теперь мы можем сгруппировать некоторые компоненты уравнения, дабы получить новое:

$$d^2 = (c_1 a_1)^2 + (c_2 a_2)^2,$$

где  $a_i$  — это константа, равная  $s_i \sum_{i=1}^N \sin(2\pi f_i j/N)^2$ .

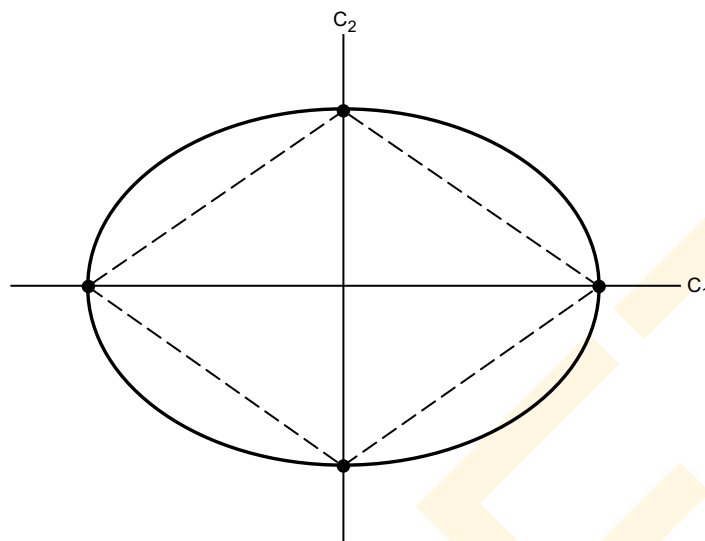
Последнее уравнение говорит нам о том, что контрасты обеих синусоид ( $c_1, c_2$ ) долж-



**Рис. 4.3.6** Смеси пространственно-контрастных паттернов. А и В — косинусоидальные стимулы с пространственной частотой 1 и 3 цикл/град. С — результат смешения этих стимулов при таком относительном смещении фаз, когда пики складываются; D — когда вычитаются.

ны быть на уровне порога. На рис. 4.3.7 дано графическое представление этих прогнозов: оси графика представляют уровни контраста двух синусоидальных компонент, использованных в смеси. Решение уравнения дает нам график, именуемый *контуром детекции*.

Как видно из рис. 4.3.7 результирующее уравнение — это уравнение эллипса, чьи главные оси совпадают с осями графика. Две неизвестных — шкалирующие коэффициенты  $a_i$  — относятся к длине главных осей. Следовательно, ежели мы шкалируем контраст синусоидальных решеток так, что порог контраста по каждому синусоидальному компоненту произвольно берется за единицу, прогнозируемый контур детекции явится окружностью (Graham & Nachmias, 1971; Nieksen & Wandell, 1988).



**Рис. 4.3.7** Эксперимент по смещению пространственных решеток позволяет тестировать модели пространственной чувствительности. Мы диагностируем видность тест-смеси, у которой контрасты синусоидальных компонентов равны  $C_1$  и  $C_2$ . Пороги детекции смесевой пары в разных комбинациях смеси дают т.н. *контур детекции*. Шейдова гипотеза прогнозирует то, что контур детекции — это овал, лежащий на осях графика (жирная линия). Если бы пиковый контраст стимула определял чувствительность зрения к смесевой паре, то контур детекции имел бы форму ромба (пунктирная линия).

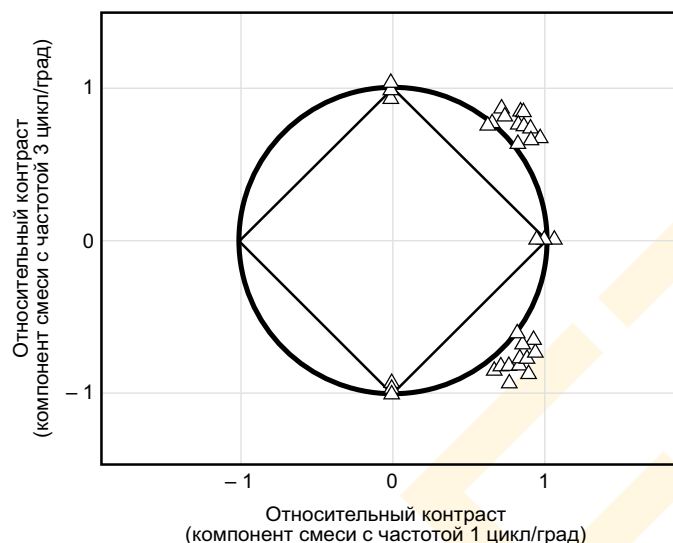
Существует множество альтернативных теорий, например: если мы предположим, что порог определен пиковым контрастом паттерна, то контур детекции стимула получит ромбовидную форму (рис. 4.3.7). Важным моментом является то, что контур детекции зависит от базовой теории. Коль скоро прогнозы, выполняемые на основе теории Шейда, очевидны и четки — мы можем использовать их же для проверки этой теории.

В 1978 г. американцы Норма Грэхем и Жакоб Начмиас (1978, см. также Graham & Nachmias, 1971) определили чувствительность к смеси двух синусоидальных решеток с пространственной частотой 1 цикл/град и 3 цикл/град. Исследователи диагностировали пороги, используя тщательно продуманную психофизическую процедуру установки порогов, получившую название *двухинтервального принудительного выбора*.

Идея состояла в том, что каждый сеанс разбивался на два временных периода, обычно отделяемых друг от друга звуковым сигналом. Таким же сигналом отмечалось и окончание сеанса. Тестовый стимул предъявлялся в одном из двух временных интервалов, а наблюдатель, глядя на экран, должен был определить, в каком интервале был предъявлен тестовый стимул. Когда контраст тестового паттерна очень низок, то наблюдателю приходилось оперировать малым объемом информации и результат был скорее случайным. Когда контраст высок, наблюдатель почти всегда верно определял временной интервал. Следовательно, когда контраст тестового стимула растет, результат варьирует от 0.5 до 1.0.

Отметим, что рабочий уровень порога — величина достаточно капризная, но по сугубо техническим причинам, описанным в своей статье, Грэхем определила порог как уровень контраста при котором наблюдатель дает верный ответ с вероятностью 0.81. Как видим по рис. 4.3.8 данные Грэхем не сели точно на контур детекции, прогнози-





**Рис. 4.3.8** Пороги восприятия пространственных смесей, диагностированные с использованием смесевой решетки (частота компонентов 1 и 3 цикл/град). Пороговые значения легли за пределы контура детекции, предсказанного шифт-инвариантной гипотезой и правилом размерности вектора.

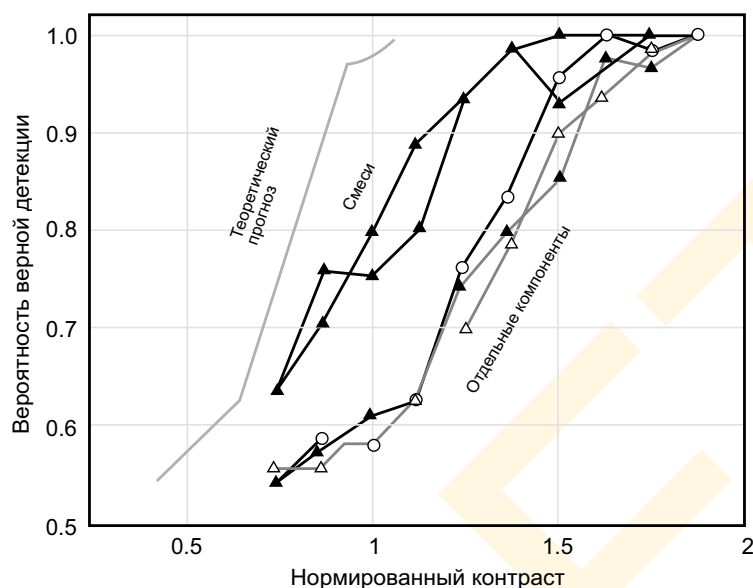
По материала Graham et al., 1978).

руемый шейдовой теорией моноразрешающего представления: пороги, диагностированные при наклоне стимулирующей решетки в  $45^\circ$  оказались в основном за пределами контура и при этом несколько выше ожидаемых: теория прогнозирует, что пороговые контрасты отдельных компонентов должны уменьшаться в 1.414 раза, тогда как на деле они уменьшились лишь в 1.2 раза. Впрочем, результат типичен такого рода экспериментам.

Однако является ли отличие значимым? Смысл нашей концепции в том, чтобы по малому числу пространственных паттернов прогнозировать чувствительность ко всем остальным пространственным стимулам. Если мы видим, что прогноз не срабатывает, когда мы диагностируем порог всего лишь по двум тестовым стимулам, мы должны всерьез задуматься: теория должна справляться с декомпозицией произвольного паттерна на сумму множества синусоидальных решеток, а затем прогнозировать чувствительность зрения к смеси компонент. Если мы наблюдаем несрабатывание даже всего лишь по двум компонентам, то как же будет работать теория по трем компонентам и более?

На рис. 4.3.9 приведены данные по измерениям чувствительности к комбинации трех синусоидальных решеток. В данном случае информация представлена графически в виде т.н. *психометрических функций*, и зависимость реакции от контраста очевидна. Точки, соединенные серыми линиями отображают вероятность правильной детекции наблюдателем отдельных синусоидальных решеток с частотами 1.33, 4.0 и 12 цикл/град. Отметим, что контрасты шкалированы по абсциссам, а шкалирующие коэффициенты подобраны так, чтобы все три кривых совместились друг с другом.

Видность двух паттернов, образованных смесями трех синусоидальных решеток, чьи соотношения контрастов подобраны так, чтобы паттерны визуально уравнились между собой — показаны сплошными линиями.



**рис. 4.3.9** Эксперимент по трехкомпонентному пространственному смешению. Вероятность верной детекции при двухинтервальном принудительном выборе дана как функция от нормированного контраста. Серые линии справа показывают детекцию трех простых синусоидальных решеток с частотами 1.33, 4 и 12 цикл/град. Две черных линии — вероятность детекции смеси этих компонент в синфазе и в противофазе. Серая линия слева — чувствительность, прогнозируемая шифт-инвариантной моделью и правилом размерности вектора.

По материалам Graham, 1989.

Опять же: коль скоро входные сигналы — это синусоиды или суммы синусоид, мы можем прогнозировать результат, базируясь на идее шифт-инвариантности нейронного изображения и правиле размерности вектора: нейронное изображение суммы трех синусоидальных решеток будет взвешенной суммой трех нейронных изображений. Прогноз порога различения смеси трех паттернов показан сплошной линией в левой части рис. 4.3.9. Правило размерности вектора прогнозирует, что вероятности правильной детекции будут одинаковы в обеих фазах.

Наконец, стоит сказать, что модельный прогноз нельзя назвать уж совсем бестолковым: взаимоотношения фаз решеток не оказывают существенного влияния на детекцию порога смеси трех паттернов. Но последняя не столь легко различима, как то прогнозирует теория: контраст трехкомпонентной смеси понижен в 1.4 раза в сравнении с их индивидуальными порогами, тогда как теоретический прогноз уменьшает контраст в 1.73 раза. В целом, базовая теория дает ряд весьма добротных рычагов прогнозирования, но точность количественных прогнозов, разумеется, падает по мере роста организации стимулов.

#### 4.3.2.7 Предварительный итог

Мы начали формулировку психофизической теории с простого количественного тезиса о шифт-инвариантности, за которой следует статическая нелинейность. И хотя такая формулировка — это колоссальное упрощение реальной ситуации, для начала она нам вполне подойдет. Нервная система очень сложна и несет в себе множество элементов, исполняющих разного рода «вычисления». Несмотря на то, что шейдовы по-

пытки передать все нюансы нейронного представления весьма вдохновляющи, в то же время они еще очень наивны, поскольку большей частью нейронные представления невероятно сложно организованы. Однако же мы можем использовать психофизические модели для исследования тех аспектов нейронного представления, что наилучшим образом предсказывают психофизическую реакцию особи. Сравнивая реактивные данные с данными нейронными, мы можем выделить главные функциональные элементы нейронного представления, нужные для решения разного рода зрительных задач.

Несмотря на то, что шифт-инвариантная концепция — это далеко не истина в высшей своей инстанции, тем не менее она организует ход нашей мысли и указывает на то, какие эксперименты мы должны поставить. Умозрительный психофизический расклад аналогичен рецептивным полям некоторых периферических нейронов. Плюс к тому, нельзя сказать, что шифт-инвариантная теория идеально согласуется с экспериментальными данными, более того: она лишь портит дело, когда речь идет о сложных зрительных стимулах, таких как изображения естественных сцен. Короче говоря: нам нужны новые теоретические подходы и новые эксперименты.

### 4.3.3 Теория мультиразрешающего представления

Шейдова теория моноразрешающего представления не дает точных прогностических данных, однако же она и не настолько плоха, чтобы отказаться от нее полностью. Поэтому мы намерены искать пути совершенствования этой теории, выделив ее лучшие качества.

Современные концепции пространственного зрения не ограничиваются одним лишь линейным кодированием, состоящим из набора шифт-инвариантных линейных преобразований. У каждого шифт-инвариантного преобразования имеется свое собственное ядро свертки и, следовательно, формируется свое собственное нейронное изображение. Итак, мы будем обращаться к данным, представленным отдельными шифт-инвариантными представлениями, как к составляющим общей теории. При этом нам нужно будет выбирать ядра свертки, ассоциированные с каждым шифт-инвариантным линейным преобразованием и следующей за ней статической нелинейностью. Также нам необходимо описать то, как выходные сигналы по разным компонентам ретинального изображения объединяются, формируя общий сигнал детекции.

#### 4.3.3.1 Пространственная адаптация в логике мультиразрешающих представлений

Мотивом к созданию теории мультиразрешающего представления явилось то, что скопилась внушительная коллекция эмпирических наблюдений, одно из которых — обсуждавшийся уже нами ранее феномен пространственной адаптации (рис. 4.2.19).

Результаты оценки уровней пространственной адаптации подталкивают нас к уходу от идеи моноразрешающего представления и принятию концепции представлений *мультиразрешающих*. В логике мультиразрешений нейронное представление строится из отдельных *субизображений*, селективно чувствительных к узкому диапазону пространственных частот и ориентаций. Такого деления зрительной информации можно добиться, используя различные ядра свертки, каждое из которых выделяет из изображения свою пространственную частоту. Сие вполне реализуем нервной системой за счет многообразия свойств рецептивных полей, в частности тех, что в великом

множестве встречаются у простых клеток зрительной коры: вспомним, что рецептивные поля этих клеток обладают как ориентационными, так и пространственно-частотными свойствами.

Итак, поскольку отдельные компоненты ретинального изображения предполагают представление различных пространственных частот, мы скажем, что:

⇒ Нейронное изображение — это мультиразрешающее представление изображения ретинального.

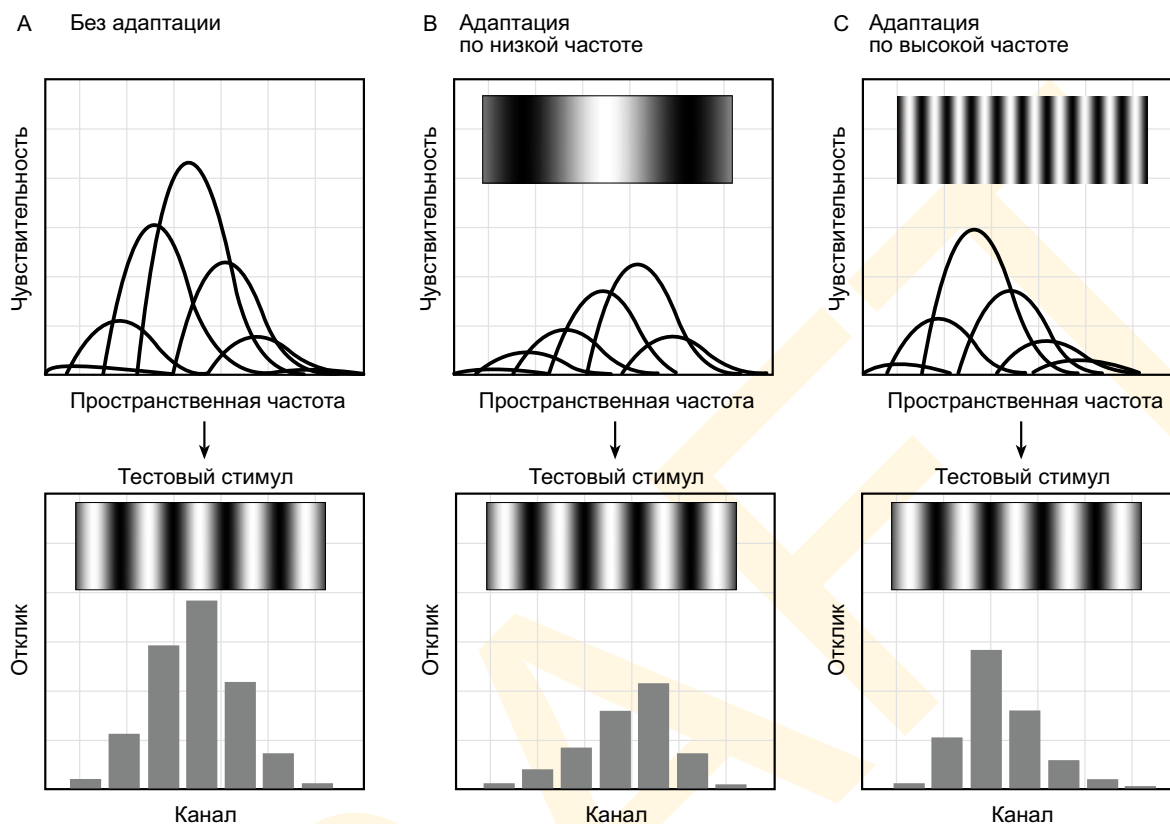
Мультиразрешающими представлениями легко объяснить феномен пространственной адаптации (рис. 4.3.10). Предположим, некая зрительная система обычным путем кодирует изображение, задействуя комплект шифт-инвариантных компонент этого изображения и чья контрастно-сенситивная кривая дана на рис. 4.3.10 А. До адаптации каждый компонент изображения представлял квадратную волну с амплитудой, зависящей от частоты этой волны и чувствительности канала. Решетка на рис. 4.3.10 А внизу, показывает амплитуду нейронных субизображений как реакцию на тестовый стимул до адаптации.

Адаптация к квадратной волне низкой частоты подавляет отклик в некоторых нейронных субизображениях (рис. 4.3.10 В). Следовательно, паттерн откликов к стимулу тестовой частоты, следующий после адаптации, будет таким, как показан в нижней части рис. 4.3.10 В. Мы видим, что картина откликов теперь состоит из сигналов, какие были бы при неадаптированном ответе на квадратно-волновой паттерн большей пространственной частоты. Сие есть объяснение тому наблюдению, что после адаптации волной низкой пространственной частоты, тестовая решетка выглядит сдвинутой в сторону больших пространственных частот.

На рис. 4.3.10 С дана картина чувствительностей, формирующаяся после адаптации к квадратно-волновому стимулу высокой пространственной частоты (вверху) и картина того, как меняются амплитуды откликов (внизу). В данном случае паттерн откликов — это паттерн неадаптированного кодирования стимула низкой пространственной частоты.

Согласно идее мультиразрешающего представления пространственная адаптация уменьшает или даже полностью умяляет вклад определенного множества нейронов в результирующий отклик системы, нарушая баланс активности и вызывая изменения в этом отклике. Возбуждение, возникающее в ответ на высокочастотную стимуляцию после адаптации к стимулу низкой пространственной частоты, относительно выше обычного (т.е. неадаптированного) и ведет к тому, что тестовая решетка воспринимается более плотной (т.е. с большей пространственной частотой). И наоборот: возбуждение в ответ на низкочастотную стимуляцию, возникающее после адаптации к стимулу высокой пространственной частоты относительно выше обычного, а результирующий перцепт обладает меньшей пространственной частотой (т.е. прутья решетки выглядят шире).

В целом, эмпирические наблюдения за паттерн-адаптацией свидетельствуют о том, что синусоидальные или квадратно-волновые адаптирующие решетки влияют на чувствительность к стимулам лишь примерно той же пространственной частоты. Эти наблюдения утверждают тезис о том, что адаптирующие компоненты ретинального изображения могут формироваться на разных пространственных частотах.



**рис. 4.3.10** Мультиразрешающая модель способна истолковать некоторые аспекты паттерн-адаптации. А — неадаптированная стимуляция. Результирующая картина откликов выведена из совокупности нейронных субизображений, каждое из которых представляет собой наибольший отклик по той или иной пространственно-частотной полосе (каналу). Пространственная селективность каждого из субизображений показана вверх, а амплитуда нейронного кодирования тестового стимула по каждому из субизображений — на диаграмме внизу. В — после адаптации по стимулу низкой пространственной частоты (показан на вставке) чувствительность нейронов, отвечающих за определенные субизображения, понижается. Совокупная картина откликов на тестовый стимул по всем каналам будет схожей с неадаптированной стимуляцией решеткой высокой пространственной частоты. С — после адаптации паттерном высокой пространственной частоты (показан на вставке) совокупный отклик окажется схож с неадаптированным откликом на решетку низкой частоты.

### 4.3.3.2 Различение решеток и маскировка

В дополнение к экспериментам по пространственной адаптации существует ряд прочих экспериментальных наблюдений, усиливающих позиции концепции мультиразрешающего представления в зрительном восприятии. Эпохальным событием явилась в свое время статья Фергуса Кэмпбелла и Джона Робсона (1968), в которой авторы доложили об экспериментах по детекции и различению, поставленных с применением квадратно-волновых решеток (и прочих пространственно-периодических стимулов).

Вспомним, что квадратные волны, как и все периодические стимулы, можно выразить в виде взвешенной суммы синусоидальных компонент, используя при этом дис-



кретные ряды Фурье. Квадратную волну —  $\text{sq}(x)$  — которая колеблется между  $+1$  и  $-1$  с частотой  $f$ , можно представить в виде суммы синусоидальных компонент:

$$\text{sq}(x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin(2\pi(2n+1)fx)$$

Квадратная волна частоты  $f$  равна сумме серии синусоид нечетных частот —  $f$ ,  $3f$ ,  $5f$  и т.д. Амплитуда синусоид падает при росте частоты: амплитуда  $3f$ -синусоиды — это одна треть амплитуды  $f$ -синусоиды; амплитуда синусоиды  $5f$  — одна пятая и т.д. Когда общий контраст квадратной волны очень низок, то амплитуда показателей высокого порядка предельно мала и ее можно проигнорировать. На низких контрастах для создания паттерна, восприятие которого будет сильно схожим с подлинной квадратной волной, потребуются всего лишь одна или две дополнительных синусоиды. На низких контрастах квадратно-волновую решетку плотно аппроксимирует паттерн, описываемый следующим уравнением:

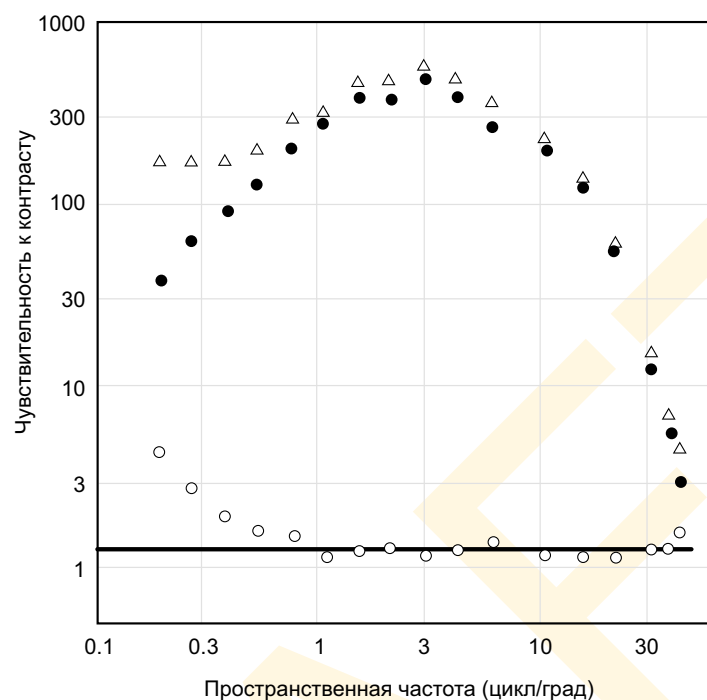
$$\text{sq}(x) \approx \frac{4}{\pi} \left[ \sin(2\pi fx) + \frac{1}{3} \sin(2\pi(3f)x) + \frac{1}{5} \sin(2\pi(5f)x) \right]$$

Проверяя концепцию мультиразрешающего представления, Кэмпбелл и Робсон пользовались квадратные волны (и прочие периодические решетки) разными способами:

— во-первых, исследователи определили уровень наименьшего контраста, при котором наблюдатель начинает различать квадратно-волновой паттерн. Отметим, что амплитуда компонента наименьшей частоты, которая, напомним, именуется *опорной* или *базисной частотой*, равна  $4/\pi$ . Поскольку контраст на опорной частоте наибольший, и поскольку чувствительность к паттернам с частотой выше 1 цикл/град растет, Кэмпбелл и Робсон показали, что нейроны, чьи рецептивные поля четко соответствуют базовой частоте, первыми сигнализируют о наличии квадратной волны. Ежели в определении видности квадратной волны сие действительно наиболее значимый показатель, то порог контраста должен составлять  $4/\pi$  раза от порога контраста синусоидальной решетки той же частоты. На рис. 4.3.11 даны контрастно-сенситивные функции по синусоидальной и квадратно-волновой решеткам и отношения контрастных чувствительностей. Как и прогнозировалось по паттернам с частотой выше 1 цикл/град, отношение контрастов на уровне порога восприятия 1.28, то есть, примерно  $4/\pi$ ;<sup>1</sup>

— во-вторых, Кэмпбелл и Робсон оценили то, насколько устойчиво наблюдатели различают квадратные волны и синусоиды. В этих экспериментах испытуемым предъявляли квадратную волну и синусоиду одной частоты —  $f$ . Две решетки были выставлены на соотношение контрастов  $4/\pi$ , гарантируя при этом, что контрасты базового компонента квадратной волны и синусоиды строго одинаковы. Испытуемые управляли контрастом обоих паттернов, удерживая данное фиксированное соотношение контрастов, до тех пор пока квадратная волна и синусоида не начинали едва-едва

<sup>1</sup> Эксперимент Кэмпбелла — Робсона — это частный случай ранее описанного эксперимента по смесевой стимуляции. Результаты показывают, что частоты от  $3f$  и выше ничего не дают наблюдателю в плане детекции пространственных квадратных волн. На вопрос о том, как объединяются компоненты различной частоты, ответ дает эксперимент по смесевой стимуляции, выполненной Грэхем и Начмиасом (1971).



**рис. 4.3.11** Чувствительность к контрасту, измеренную по квадратно-волновым решеткам с частотой, превышающей 1 цикл/град, можно прогнозировать по контрасту на базовой частоте. Белые треугольнички и черные кружки демонстрируют чувствительность к контрасту квадратных и синусоидальных волн соответственно. Белые кружки — отношение контрастных чувствительностей на каждой из частот. Черная линия проходит по значению  $4/\pi \approx 1.273$  — амплитуде на базовой частоте единично-контрастной квадратной волны.

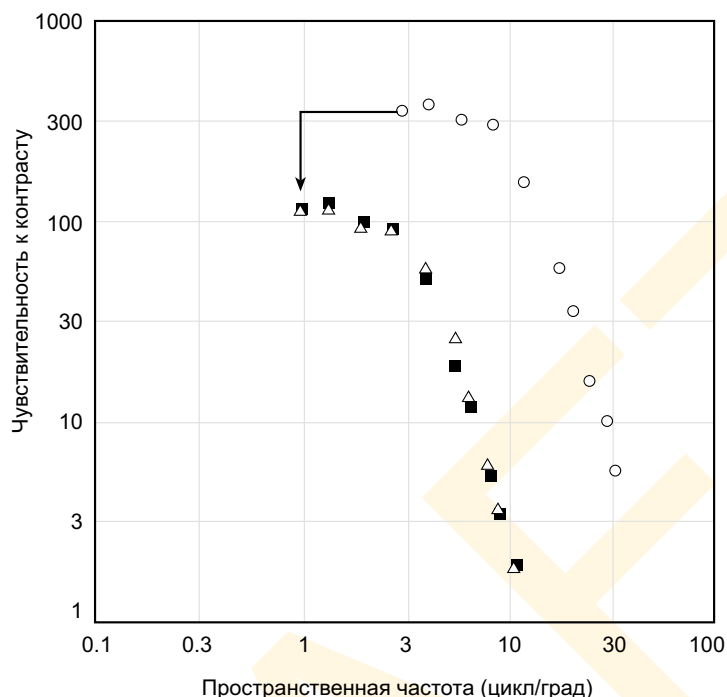
По материалам Campbell & Robson, 1968.

отличаться друг от друга. Поскольку контраст на опорной частоте квадратной волны оставался равным таковому у синусоиды, различимость строилась только на компонентах с частотами от  $3f$  и выше.

Кэмпбелл и Робсон обнаружили также, что наблюдатели начинают видеть отличие между квадратной волной и синусоидой, когда контраст третьей гармоники достигает порогового уровня.

Выводы исследователей основаны на данных, показанных на рис. 4.3.12: сплошными кружками отмечена контрастно-сенситивная функция; полыми кружками — контраст квадратно-волнового стимула, когда перцепт последнего только-только начинает отличаться от синусоиды. Четко видно, что контраст квадратной волны, необходимый для различения двух паттернов, превышает контраст, необходимый для собственно детекции квадратной волны. Однако же мы можем объяснить этот факт, приняв во внимание контраст  $3f$ -компонента квадратной волны: вспомним, что контраст этого компонента составляет одну треть контраста квадратной волны и поэтому, сдвинув данные по квадратной волне (полые кружки) влево с коэффициентом 3 по пространственной частоте и понизив контраст также в три раза, мы компенсируем оба эти фактора. Полые треугольнички показывают результат такого сдвига. Мы видим при этом, что он совпал с результатами исходных замеров контрастной чувствительности.

Из совпадения сдвинутых данных по распознаванию с данными контрастно-сенситивных измерений мы делаем вывод о том, что квадратно-волновой стимул начинает



**рис. 4.3.12** Различие между синусоидальными и квадратно-волновыми решетками становится заметным тогда, когда третья гармоника квадратной волны достигает собственного независимого порога. Белые треугольнички — контрастно-сенситивные функции. Белые кружки — уровень контраста, при котором квадратную волну уже можно отличить от базовой частоты. Черные квадратики — данные по различению квадратной волны сдвинутой с коэффициентом 3 по частоте и контрасту.

По материалам Campbell & Robson, 1968.

отличаться от синусоиды тогда, когда контраст  $3f$ -компонента квадратной волны выходит на уровень собственного порога детекции.

#### 4.3.3.3 Маскировка и фасилитация

Данные Кэмпбелла — Робсона по визуальном различению хорошо согласуются с концепцией мультиразрешающего представления. Предположим, например, что базовая и третья гармоники кодированы разными нейронными субизображениями. Поскольку амплитуда базового компонента у квадратной волны и синусоиды одна и та же, наблюдатель не может воспользоваться этой информацией для различения этих стимулов. Когда контраст третьей гармоники переходит собственный независимый порог, наблюдатель начинает отличать квадратно-волновой паттерн от синусоидального.

Однако же, несмотря на то, что идеология мультиразрешающих представлений четко согласуется с этими результатами, мы должны лишний раз убедиться в добротности доказательств. Одна из наиболее общих схем такова: перед наблюдателем ставится задача отличить друг от друга два пространственных паттерна —  $S$  и  $S + \Delta S$ , когда  $\Delta S$  — на собственном пороге детекции стимула  $S$ . Сие есть феномен, о котором Кэмпбелл и Робсон доложили, когда  $S$  и  $\Delta S$  представляли собой стимулы, существенно раз-

нящиеся по пространственной частоте. Будет ли наблюдатель всегда отличать  $S$  от  $S + \Delta S$ , когда  $\Delta S$  на своем собственном пороге? Ответ — нет.

На самом деле, случай, описанный Кэмпбеллом и Робсоном весьма редок. В большинстве ситуаций две решетки  $S$  и  $S + \Delta S$  не различаются наблюдателем даже тогда, когда  $\Delta S$ , предъявленный отдельно, четко виден. То есть, мы говорим, что стимул  $S$  *маскирует* стимул  $\Delta S$ . Существуют также ситуации, когда  $S$  и  $S + \Delta S$  визуально отличны друг от друга, даже если  $\Delta S$  сам по себе не виден. В данном случае мы говорим, что стимул  $S$  облегчает детекцию стимула  $\Delta S$ , что получило название эффекта *фасилитации*.

Маскировка и фасилитация детекции — весьма распространенные вещи, более того — отсутствие маски или облегчения, как, скажем, в данных Кэмпбелла и Робсона — редкость.

Изображения на рис. 4.3.13 А демонстрируют феномен зрительной маскировки: решетка слева — тестовый контрастный стимул ( $\Delta S$ ), добавляемый к одной из маскирующих решеток в центральной колонке. Маскирующая решетка сверху схожа с тестовой по ориентации, но в три раза выше по пространственной частоте. Если тщательно взглянуть, то станет заметной разница между отдельно маской и маской в связке с тестовой решеткой. В частности, чуть левее центра решетки несколько прутьев выглядят темнее, а чуть правее — светлее.

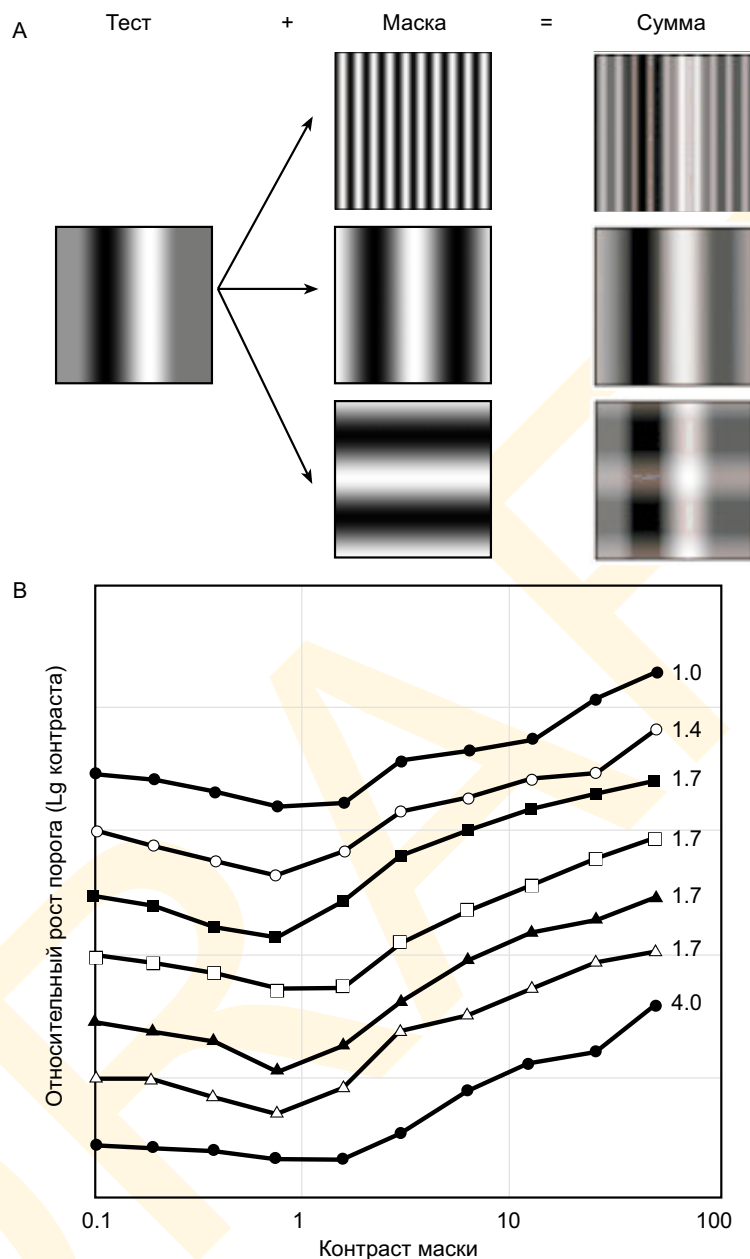
Вторая маска схожа с тестом по ориентации и пространственной частоте. В данном случае добавка тестового стимула видна отчетливее.

Третья маска аналогична тесту по пространственной частоте, но отлична по ориентации. В этом случае добавка тестового стимула выявляется очень легко.

На рис. 4.3.13 В показаны результаты измерений степени маскировки и фасилитации между решетками схожих пространственных частот и одной общей ориентации (Legge & Foley, 1980). Наблюдатели отличили маску  $S$  от маски с добавленной к ней синусоидальной тестовой решеткой ( $S + \Delta S$ ) с частотой 2 цикл/град. По оси ординат отложен пороговый контраст детекции тестового стимула, а по оси абсцисс — контраст маскирующего стимула  $S$ . Кривые демонстрируют результаты по маскам разных пространственных частот. В целом наличие стимула  $S$  облегчает детекцию теста на низких контрастах и детекцию маски на высоких. Когда пространственные частоты решеток  $S$  и  $\Delta S$  отличаются в два раза — мера фасилитации невелика, даже несмотря на явную маскировку.

Прочие экспериментальные измерения показали, что когда пространственные частоты тестовой решетки и маски отличаются в три раза, эффект маскировки понижен (DeValois, 1977a, b; Wilson et al., 1983). Позднее мы коснемся эффекта ориентации маски.

Итак, выводы из экспериментов по моделированию мультиразрешающих представлений можно разделить на две части. Во-первых, данные Кэмпбелла — Робсона свидетельствуют о том, что эффекты маскировки и фасилитации отсутствуют, когда контрасты теста и маски низки, а пространственные частоты сильно разнятся. Во-вторых, данные Легги и Фоли указывают на то, что по множеству пар стимулов, сходных по пространственной частоте, видность стимула  $S$  растет при добавке  $\Delta S$ . И первое, и второе приводит нас к мысли, что стимулы разных пространственных частот кодируются разными нейронными субизображениями.



**рис. 4.3.13** Маскировка и фасилитация. А — изображения, иллюстрирующие эффект зрительной маскировки. Тестовая решетка — слева; в средней колонке — маскирующие решетки; справа — сумма тестового и маскирующего контрастов. Когда пространственные частоты теста и маски отличаются с коэффициентом 3 (вверху), влияние тестовой решетки прослеживается очень четко. Когда пространственная частота теста и маски близки друг другу (в середине), довольно сложно углядеть присутствие теста. Когда ориентация теста и маски сильно разнятся (внизу), то до добавка теста легко читается. В — пороговый контраст детекции мишени с пространственной частотой в 2 цикл/град ( $\Delta S$ , ордината) зависит от контраста маскирующей решетки ( $S$ , абсцисса). Каждая из кривых представляет эффект решетки —  $S$  — той или иной пространственной частоты. Когда контраст  $S$  невысок, а пространственная частота и ориентация те же самые — это облегчает детекцию теста; когда контраст высок — затрудняет (маскирует). Для удобства чтения кривые разнесены по оси ординат.

По материалам Legge & Foley, 1980



#### *4.3.3.4 Концептуальные преимущества идеи мультиразрешающего представления*

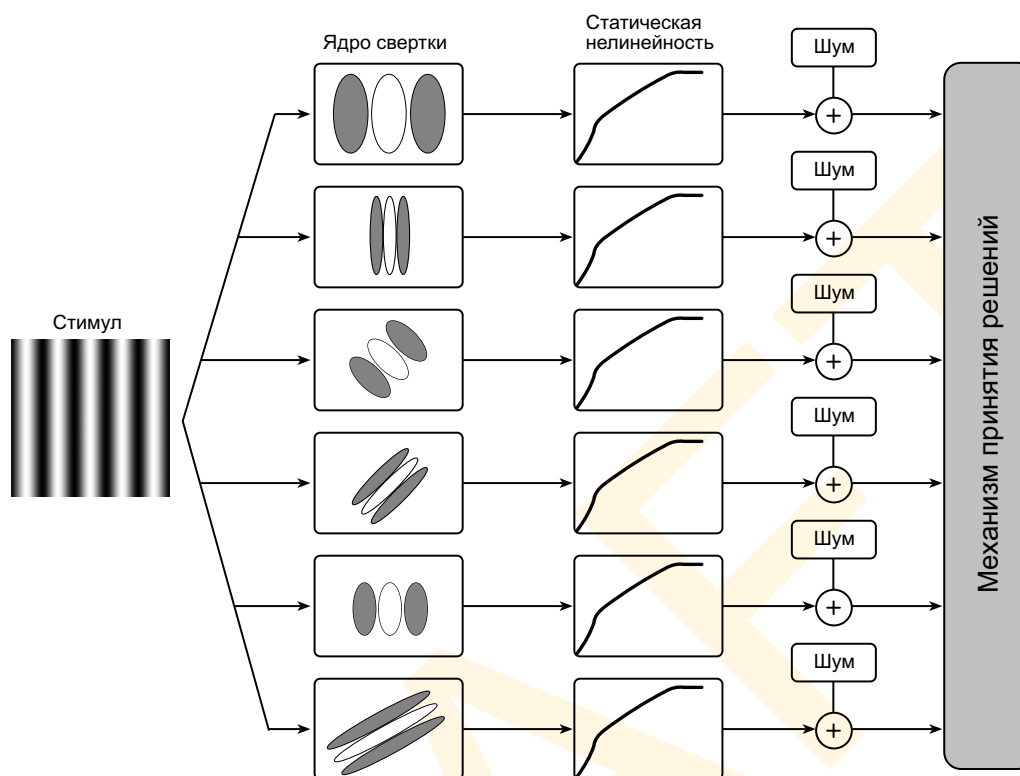
Сегодня во многих сферах работы с изображениями активно применяется мультиразрешающий формат: оригинальные данные превращаются набор субизображений, отличных друг от друга главным образом по их пиковой пространственно-частотной селективности. Иными словами, идея мультиразрешающих представлений открыла целую серию направлений исследования, о которых мы будем говорить позднее. Помимо того, что психофизические аргументы концепции мультиразрешения весьма интересны — они же вполне достаточны для объяснения того, почему гипотеза мультиразрешающего представления привела, без преувеличения, к революции в науке о зрении.

Теперь рассмотрим два вопроса, встающие перед исследователями, когда те отказываются от шифт-инвариантной гипотезы. Во-первых, множество возможных линейных функций кодирования огромно — какое преобразование выбрать? Во-вторых, без шифт-инвариантности исследователь теряет, как говорят, прогностическую способность, поскольку множество важнейших выводов мы делаем именно на базе шифт-инвариантности — как быть?

Отказ от шифт-инвариантности чреват тем, что способов кодирования оказывается бесчисленное множество, и исследователю нужен какой-то принцип выбора. Концепция мультиразрешения помогает в таком выборе, поскольку дает определенные правила спецификации свойств шифт-инвариантных преобразований, мотивируя исследователя и описывая подходящие виды линейного кодирования.

Даже несмотря на то, что теория мультиразрешающего представления предлагает лишь самую общую организационную схему линейного кодирования, она, тем не менее, дает гибкость и позиционирует свойства субизображений. Сия гибкость нужна в построении субизображений, чьи реактивные качества аналогичны свойствам рецептивных полей различных популяций корковых нейронов. Простая шифт-инвариантная теория, несущая лишь одно ядро свертки, не может представить чувствительность, например, к пространственной ориентации стимула, поскольку когда ядро свертки селективно по ориентации, прогноз оказывается таков, будто наблюдатель наиболее чувствителен к стимулам определенной ориентации. Коль скоро очевидно, что зрительная система не обладает повышенной чувствительностью к стимулам какой-то одной пространственной ориентации, ядра свертки, «профилированные» по ориентации, не подходят к модели моноразрешающего представления. Однако применение таких ядер возможно в логике мультиразрешающего представления пространственной визуальной информации. Поскольку множество ядер на каждом разрешении кодируют все ориентации, модель не дает таких абсурдных прогнозов как гиперчувствительность к определенным направлениям. По этой причине некоторые исследователи говорят о том, что существует тесная связь между идеей мультиразрешения и корковым представлением пространственной информации.

На рис. 4.3.14 дана простейшая мультиразрешающая модель, описанная **Вильсоном и Регом** (1984; а также Foley & Legge, 1981; Watson, 1983; Wilson & Gelb, 1984; Watt & Morgan, 1985). Исходное ретинальное изображение линейно преобразовано в нейронное, состоящее из нескольких субизображений. В сравнительно недавней версии этой модели Вильсон и Рего (1984) собрали нейронное изображение из 48 субизображений, разложив их на шесть пространственных групп по восьми вариантам ориентации (т.е. по шесть пространственных частот в восьми ориентациях). За каж-



**Рис. 4.3.14** Мультиразрешающая модель пространственной чувствительности. Стимул свернут с набором пространственных фильтров разной пространственной селективности. Выходной сигнал этих фильтров видоизменен нелинейной компрессией и шумом. Результат объединен в нейронное изображение.

По материалам Spillman & Werner, 1990.

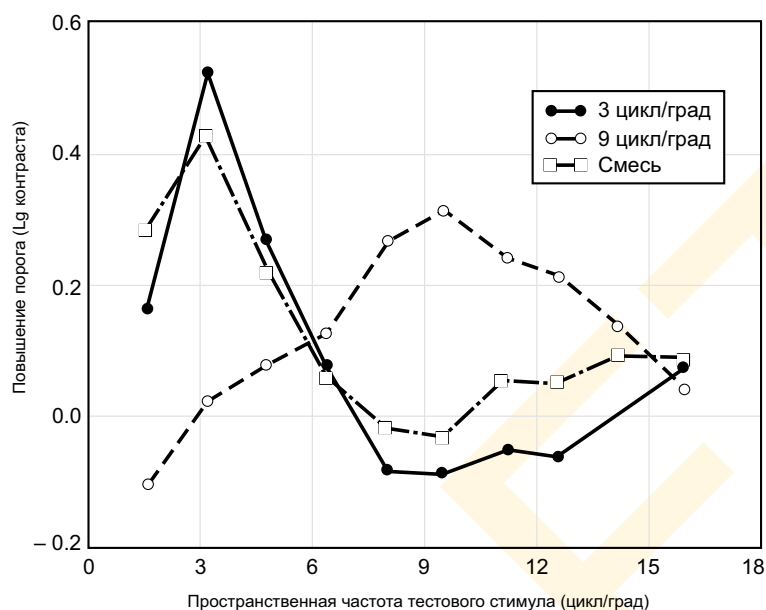
дым субизображением следовала статическая нелинейность, то есть модификация размерности вектора.

Впоследствии Вильсон и Реган разработали еще более мощную модель, применив очень сложные допущения, касаемые внутреннего физиологического шума зрительной системы, а также способности индивида к принятию разного рода решений.

Наконец отметим, что сложность количественных преобразований — это серьезная преграда на пути к разработке и применению мультиразрешающих моделей пространственной чувствительности зрительной системы человека, и потому сегодня в решении этой задачи активно используются компьютеры.

#### 4.3.3.5 Издержки концепции мультиразрешающего представления

Сегодня теория мультиразрешающего представления — это базовая концепция рассуждений о пространственной чувствительности зрения. Как было показано в предыдущем разделе у мультиразрешающих представлений множество замечательных свойств, которые можно с успехом применять при объяснении некоторых важных экспериментальных данных. Однако в теории мультиразрешения есть ряд эмпирически спорных моментов, и в данном разделе мы поговорим о нескольких опытах, иллюстрирующих эти моменты. Как мы увидим, многие из них исходят из общей точки, — из экспериментов со стимулирующими смесями.



**Рис. 4.3.15** Эксперимент по адаптации к смеси стимулов. Кривые демонстрируют повышение порога (в логарфмических величинах) при различных пространственных частотах тестового стимула после адаптации к синусоиде с пространственной частотой в 3 цикл/град (сплошная черная линия); к синусоиде с частотой 9 цикл/град (пунктирная линия); к их смеси (пунктирно-точечная линия). Рост порога после адаптации смесью явно меньше, чем при адаптации отдельными ее компонентами.

По материалам Nacmias et al., 1973

**Пространственная адаптация и смеси стимулов.** Если же для укрепления концепции мультиразрешающих представлений мы хотим использовать феномен пространственной адаптации, то должны уделить сколько-то внимания изучению общих принципов ее измерения. Возможно, первый шаг — это приложение данных, полученных при оценке адаптации по простым синусоидам, к более общим стимулам, представляющим собой смесь двух синусоид.

В 1973 г. Жакоб Начмиас провел эксперимент по пространственной адаптации, используя в роли адаптирующих как отдельные синусоидальные стимулы, так и их смеси. Вопрос, поставленный исследователем, был классическим вопросом всех смесевых экспериментов: «Можем ли мы, используя моноизмерения, прогнозировать психофизическую реакцию на сумму стимулов?».

Результаты опыта Начмиаса показаны на рис. 4.3.15: каждая кривая представляет повышение порога синусоидальной тестовой решетки на различных пространственных частотах. Сплошная линия характеризует повышение порога, при адаптации по стимулу в 3 цикл/град. Согласно Блэйкмо — Кэмпбеллу (1969) имеет место ощутимое повышение порога на пространственной частоте от 3 цикл/град и существенно меньшая степень адаптации на больших и меньших частотах.

Пунктирная линия характеризует рост порога при адаптации по стимулу в 9 цикл/град. Так получилось, что контраст этой решетки составил одну треть от контраста решетки базовой частоты. Но даже при столь сниженном контрасте решетка с частотой 9 цикл/град дает существенный рост порога по стимулам той же частоты.

Полые квадратики демонстрируют повышение порога после адаптации к смеси

двух адаптирующих стимулов. У данного наблюдателя адаптация к смеси не подняла порог восприятия тестовой решетки в 9 цикл/град. По всем наблюдателям в этом исследовании после адаптации по решетке в 9 цикл/град повышение порога было меньшим, чем после адаптации по решетке в 3 цикл/град. В логике подъема порога смесь двух стимулов с частотами в 9 и в 3 цикл/град обладают меньшим адаптирующим потенциалом, чем отдельный стимул в 9 цикл/град.

Полученный результат трудно согласовать с простой интерпретацией адаптации и пространственно частотных каналов (рис. 4.2.17 и 4.3.10): если адаптация к стимулу с частотой в 3 цикл/град стимулирует иной набор нейронов, нежели адаптация к стимулу в 9 цикл/град, то почему адаптация к их смеси повышает чувствительность к стимулу в 9 цикл/град? Однако нам до сих пор неизвестно какого-либо внятного объяснения этого феномена.

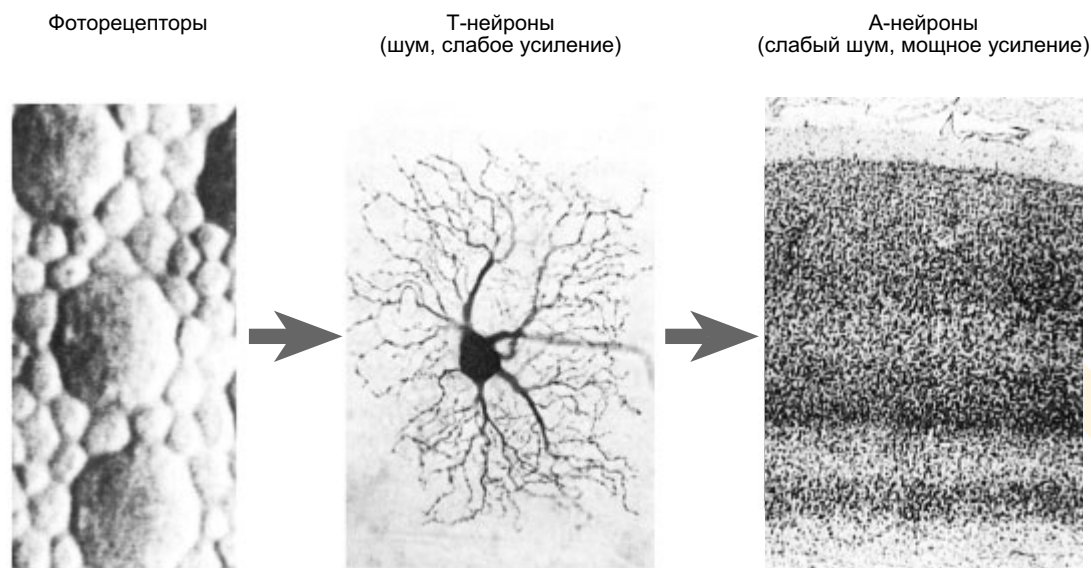
Данные смесевых экспериментов побуждают нас пересмотреть текущий взгляд на механизмы пространственной адаптации. рис. 4.3.16 иллюстрирует одну из сложностей, с которыми мы сталкиваемся, когда пытаемся увязать друг с другом результаты экспериментов по детекции и адаптации. Когда мы объединяем результаты обоих экспериментов, то мы косвенно предполагаем, что зрительные механизмы, ограничивающие детекцию, те же, что влияют на зрительную чувствительность после пространственной адаптации. Но психофизические измерения не дают прямого доказательства того, что нейроны, ограничивающие чувствительность — это те же самые нейроны, что отвечают за пространственную адаптацию.

Схема рис. 4.3.16 иллюстрирует то, почему такое предположение неверно. Допустим, что нейроны, отмеченные как «Группа Т», расположены в начальном участке зрительного пути, что нейроны эти сильно шумят и дают слабое контрастное усиление. Ежели существует хоть какая-то вероятность существования таких нейронов, то скорее всего именно они ограничат чувствительность системы. В этом случае детективность наблюдателя мы можем поднять, применив тестовые стимулы, хорошо согласующиеся с реактивными свойствами нейронов Т-группы.

Мы предположили в свое время, что эффекты пространственной адаптации — суть следствие нейронного утомления, вызванного мощной стимуляцией. Поскольку у нейронов Т-группы относительно низкое усиление, и они не дают мощного отклика на большинство стимулов, то существенного влияния на пространственную адаптацию они оказать не могут. Вполне вероятно, что, вместо того, другая группа нейронов — «Группа А» оказывает наибольшее влияние на пространственную адаптацию. На рис. 4.3.16 нейроны группы А располагаются дальше от сетчатки по зрительному пути, чем нейроны группы Т. Пространственные параметры эксперимента по паттерн-адаптации (к примеру, вариабельность чувствительности к тестовому стимулу в зависимости от пространственных параметров стимула адаптирующего) могут оказаться зависимыми от пространственных свойств рецептивных полей нейронов группы А. При этом последние могут сильно отличаться от рецептивных полей нейронов группы Т.<sup>1</sup>

Из сказанного ясно, что пространственные свойства адаптированного нейронного кодирования могут отличаться от пространственных свойств нейронного кодирования

<sup>1</sup> Мы также должны принимать во внимание и то, что базовый механизм нейронного утомления — это не главный источник пространственной адаптации. Сравнительно недавно Барлоу и Фолдиак (1989) дали принципиально иное объяснение пространственной адаптации, основанное не на принципах нейронного утомления.



**Рис. 4.3.16** Нейроны, ограничивающие детекцию и нейроны, отвечающие за пространственную адаптацию — это не одни и те же клетки. К примеру, нейроны одной группы (Т-нейроны) могут оказаться шумными и вяло усиливающими контраст. Из-за высокой зашумленности отклика эти нейроны будут повышать порог детекции. Нейроны второй группы (А-нейроны) способны интегрировать отклики нейронов первой группы и значительно усиливают контраст. Последнее ведет к их быстрой утомляемости, что вполне может стать физиологической основой пространственной адаптации. Если действительно за оба варианта отклика отвечают разные группы клеток, то получается, что рецептивные поля, выявленные в процессе экспериментов по детекции и пространственной адаптации — это поля разных групп нейронов.

По материалам Christine Curcio & Jennifer Lund.

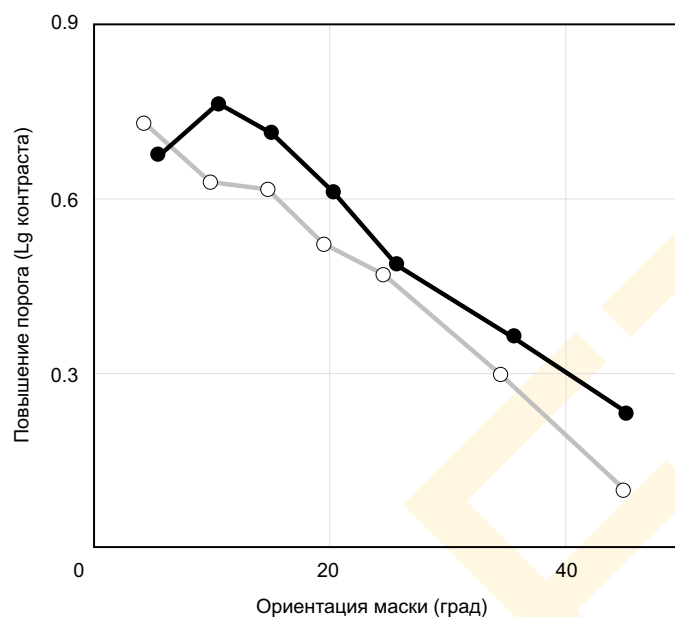
при детекции. Дабы показать, что механизмы, ограничивающие детекцию и провоцирующие пространственную адаптацию — это одни и те же механизмы, мы должны добыть психофизические экспериментальные данные, поддерживающие данную точку зрения. В этом случае нам удастся связать воедино результаты по детекции и пространственной адаптации, и мы сможем сделать вывод о внутренней организации моделей мультиразрешающего представления.

**Маскировка смесями.** В эксперименте Кэмпбелла и Робсона (1968) наблюдателей просили найти отличие между стимулом  $S$  и стимулом  $S + \Delta S$ , где  $S$  и  $\Delta S$  синусоидальные решетки. Исследователи обнаружили, что когда  $S$  и  $\Delta S$  — это синусоиды, далеко отстоящие друг от друга по своим пространственным частотам, две решетки начнут визуально отличаться друг от друга когда контраст  $\Delta S$  окажется на уровне собственного порога видности. Анализируя эксперименты Кэмпбелла — Робсона, мы рассмотрим то, насколько маскировка зависит от относительной пространственной частоты тестовой и маскирующей решеток (см. рис. 4.3.13).

Данные рис. 4.3.17 демонстрируют нам зависимость маскировки от относительной ориентации тестового стимула и стимула маскирующего (Phillips & Wilson, 1984). В этом исследовании пространственные частоты маскирующего стимула ( $S$ ) и тестового ( $\Delta S$ ) были одинаковыми.

В 1984 г. американцы Грегори Филипс и Хью Вильсон выявили уровень контраста, которым должен обладать тестовый стимул, чтобы наблюдатель начал отличать  $S + \Delta S$





**Рис. 4.3.17** Рост порога детекции стимула в 2 цикл/град как функция от ориентации маскирующего стимула той же пространственной частоты (по двум наблюдателям).

По материалам Phillips & Wilson, 1984.

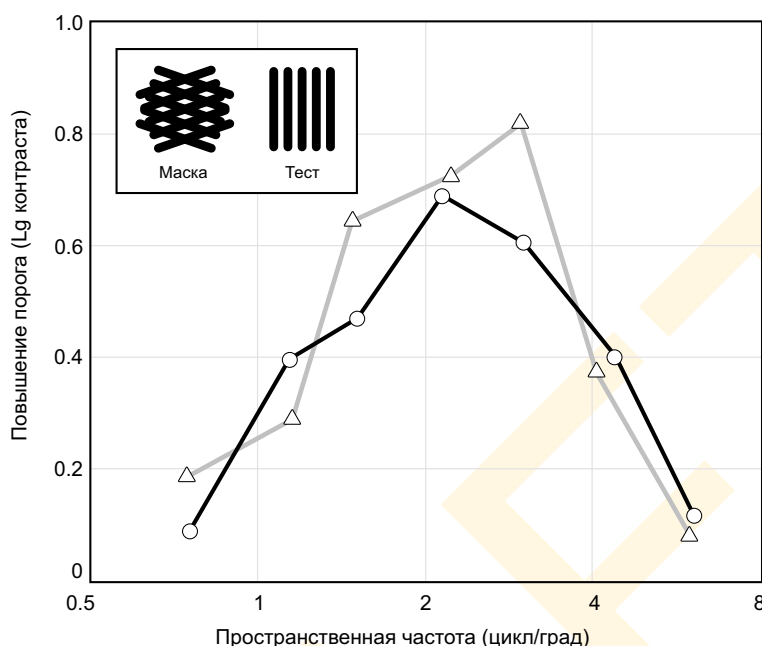
от  $S$  при той или иной ориентации маски (рис. 4.3.17). По оси абсцисс графика отложены ориентации маскирующего стимула  $S$ ; по ординатам — рост порога стимула тестового. Четко видно, что по мере роста разницы в ориентациях тестового стимула и стимула маскирующего, маскирующий эффект падает. Когда разница в ориентациях превышает  $40^\circ$ ,  $S + \Delta S$  начинает отличаться от  $S$ , если контраст  $\Delta S$  выходит на собственный порог.

График рис. 4.3.17 явно намекает на то, что одномерные контрастные паттерны, разделенные по ориентации более чем на  $40^\circ$ , кодируются разными группами нейронов. То есть, экспериментальные данные такого рода можно использовать в моделях мультиразрешающих представлений для построения ядер свертки, селективных по ориентации.

Результаты, полученные при экспериментах по визуальной маскировке смесями, свидетельствуют о допустимости такого рассуждения.

В 1989 г. британцы Эндрю Деррингтон и Брюс Хеннинг доложили о проведении ими смесевых экспериментов, при которых с помощью двух отдельных маскирующих паттернов и их смеси был выявлен рост порога. По отдельности маскирующие паттерны представляли собой синусоидальные решетки с частотой 3 цикл/град, при этом одну решетку сориентировали на  $+67.5^\circ$  относительно вертикали; другую — на  $-67.5^\circ$ . После чего исследователи измерили эффект этих масок на различные вертикально ориентированные синусоидальные решетки.

Итак, ежели два маскирующих стимула представлены сигналами нейронов, отличных от тех, что представляют вертикальные тестовые паттерны, то суперпозиция этих масок не должна влиять на видность тестового стимула. Однако данные рис. 4.3.18 свидетельствуют об обратном: смесь двух маскирующих паттернов тоже является по-



**рис. 4.3.18** Эксперимент по маскировке смесями. Когда тестовая и маскирующая решетки развернуты на  $67.5^\circ$ , маска не оказывает влияния на видность теста, но при этом комбинация двух маскирующих решеток, развернутых друг относительно друга на этот же угол, действует как очень мощная маска. Черная и серая линии — результат эксперимента по двум наблюдателям.

По материалам: Derrington & Henning, 1989.

тенциальной маской даже если каждый в отдельности не оказывает маскирующего эффекта.<sup>1</sup>

**Текущий статус теории мультиразрешающего представления.** Концепция мультиразрешающего представления — это очень мощный теоретический инструмент, помогающий нам рассуждать о проблемах пространственной чувствительности зрения и дающий базис построения количественных моделей пространственной чувствительности и прочих зрительных пространственных эффектов. Сегодня у нас имеются твердые данные за то, что мультиразрешающие представления — это важный функциональный элемент зрительных потоков. Однако же существует и внушительный массив экспериментальных данных (от детекции до пространственной адаптации и маскировки), результаты которых никак не стыкуются с главными постулатами теории мультиразрешающих представлений. Как уже было показано, результаты смесевых экспериментов, задействующих эффекты пространственной адаптации и маскировки, оказываются весьма сложны для истолкования, если мы полагаем, что составляющие ретинального изображения, ассоциированные с тем или иным диапазоном пространственных частот и/или ориентаций, кодируются независимыми друг от друга наборами нейронов.

Неоднозначная и противоречивая картина экспериментальных данных свидетель-

<sup>1</sup> Аналогичные сложности в интерпретации эффектов маскировки, но в отношении смесей синусоидальных решеток, рассматривались Жакобом Начмиасом и Бернисом Роговиц (1983), а также Марком Перкинсом и Майклом Лэнди (1991).

ствуем о том, что пока мы еще не достигли полного понимания базовых нейронных процессов, вызывающих адаптацию и маскировку. Также мы пока не понимаем того, как эти нейронные процессы соотносятся с нейронными процессами, ограничивающими пространственную чувствительность. Однако же разобраться в этом необходимо, бо эксперименты эти дают ключевые сведения, укрепляющие позиции концепции мультиразрешающего представления. Возможно, прочтя, наконец, скрытый текст этих экспериментов, мы станем лучше понимать роль мультиразрешающих представлений.

#### 4.3.4 Пространственная чувствительность и прочие зрительные параметры

Теперь поговорим о том, как и насколько пространственная чувствительность зрения зависима от прочих аспектов условий просмотра, таких как уровень общей освещенности сцены, временные параметры стимуляции, а также спектральные свойства стимула. При этом мы будем пользоваться контрастно-сенситивной функцией как краткой характеристикой психофизической реакции наблюдателя на стимуляцию.

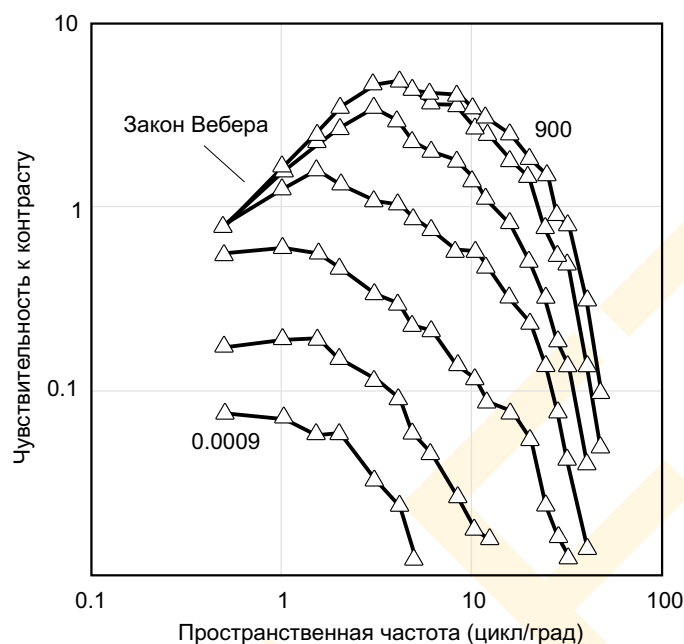
Сразу отметим, что в оставшейся части текущей главы контрастно-сенситивная функция будет выступать в несколько ином качестве, чем до сих пор. Здесь следует подчеркнуть особую роль контрастно-сенситивной функции теории линейных систем: если мы хорошо понимаем смысл экспериментальных данных, контрастно-сенситивную функцию можно использовать для прогнозирования чувствительности ко многим сложным и даже очень сложным стимулам. Простейший пример — шейдово применение контрастно-сенситивной функции: если зрительная чувствительность шифт-инвариантна нейронным изображениям, то мы можем использовать контрастно-сенситивную функцию для прогноза чувствительности к любому сложному стимулу.

Наша цель — выработать *общий принцип* диагностики зрительной чувствительности к контрастным стимулам, но, к сожалению, в нашем распоряжении до сих пор нет полноценной теории, позволяющей нам применять контрастно-сенситивную функцию в общей характеристике психофизики зрения. Проще говоря: мы не можем рассматривать контрастно-сенситивную функцию как полную характеристику пространственной чувствительности зрения, но скорее как описательный инструмент, помогающий нам изучать общую пространственную чувствительность зрительной системы человека.

Мотивом к стандартизации контрастно-сенситивной функции отчасти является то, что показатель этот широко используется как в физиологии, так и в психофизике. Следовательно, психофизические измерения контрастно-сенситивной функции дадут нам некий эталон сравнения с нейронным откликом в разных точках зрительного пути. Если нейроны какого-то специфического класса, скажем, ганглиозные клетки сетчатки, ограничивают зрительную чувствительность, мы ожидаем того, что кривая психофизической контрастно-сенситивной функции и нейронной контрастной чувствительности начнут меняться синхронно, когда мы начнем менять условия эксперимента.

##### 4.3.4.1 Световая адаптация

По рис. 4.3.19 хорошо видно, что контрастно-сенситивная функция меняется, когда диагностируется при разных уровнях средней интенсивности стимуляции: кривая внизу слева — это контрастно-сенситивная функция, построенная при низком уровне



**рис. 4.3.19** Контрастно-сенситивная функция зрительной системы человека меняется при изменении уровня средней фотометрической яркости стимула. Каждая кривая демонстрирует контрастно-сенситивную функцию при различных уровнях средней освещенности сетчатки, лежащих в диапазоне от 0.0009 троландов до 900. По каждой кривой освещенность сетчатки меняется в 10 раз. Стимул представляет собой монохроматический поток с длиной волны 525 нм. На самом нижнем уровне (скотопические условия) контрастно-сенситивная функция работает как низкочастотный фильтр с пиком на 1 цикл/град. На высоких уровнях фотопической яркости функции становятся полосовыми и дают пик примерно на 8 цикл/град. На еще больших яркостях контрастно-сенситивная функция остается неизменной.

По материалам Nes & Bouman, 1967.

средней яркости ( $9 \times 10^{-4}$  троландов сетчатки), т.е. практически в ситуации палочкового зрения.<sup>1</sup> В таких условиях контрастно-сенситивная функция дает пик на 1-3 цикл/град и график скорее низкочастотный, нежели полосовой; кривая справа вверх — это контрастно-сенситивная функция при 6-8 цикл/град пространственной частоты и форма кривой явно полосовая; на средних интенсивностях (порядка 1000 троландов) контрастно-сенситивная функция остается неизменной (Westheimer, 1960; van Nes & Bouman, 1967).

Изменение формы контрастно-сенситивной функции тесно связано с несколькими простыми принципами формирования изображения на сетчатке.

Первый касается важности получения адекватного сигнала в тех или иных условиях внешнего освещения. На очень низких уровнях освещенности наблюдателю для получения внятного отклика сетчатки необходимо собрать весь, упавший на нее свет. Добиться сего можно лишь пространственно, т.е. выполнив пространственное усред-

<sup>1</sup> Один троланд — это единица измерения освещенности сетчатки. Для измерения в троландах необходима поправка на диаметр зрачка. Освещенность сетчатки в 1 Т возникает тогда, когда наблюдатель смотрит на поверхность с фотометрической яркостью в 1 кд/м<sup>2</sup> и диаметром зрачка в 1 мм.

нение оптического сигнала. В этом случае речь не может идти о различении стимулов высоких пространственных частот. Следовательно, в тусклых, скотопических условиях чувствительность зрения к высоким пространственным частотам резко снижена.

При высокой интенсивности внешнего освещения, когда фотонов много, зрительная система может собрать требуемую ей информацию с меньших областей, и поэтому пространственное разрешение растет;

— второй принцип касается большей значимости контраста в обработке зрительной информации. На рис. 4.3.19 показано, что чувствительность к контрасту на низких пространственных частотах стимуляции (ниже 1 цикл/град) растет по мере роста средней яркости, а затем выходит на плато. Длина этого плато, т.е. диапазон неизменности контрастной чувствительности, именуется *веберовским участком*. По низким пространственным частотам веберовский закон работает очень устойчиво. На больших частотах контрастная чувствительность начинает плавно расти по мере роста средней яркости. При таком характере стимулов веберовский закон уже не является точным описанием психофизической чувствительности.

Однако даже несмотря на то, что закон Вебера не всегда точен, он по-прежнему ценен. Рассмотрим основные параметры нашего эксперимента: фоновые интенсивности меняются на шесть порядков (т.е. миллионкратно); чувствительность к абсолютному уровню яркости — на 4-5 порядков, но при этом чувствительность к контрасту меняется всего лишь в 20 раз. Результаты свидетельствуют о том, что зрительная система в первую очередь удерживает чувствительность к контрасту стимула (что следует из закона Вебера), нежели к абсолютной световой интенсивности.

Наконец отметим, что зрение уверенно сохраняет контрастную чувствительность на низких пространственных частотах и весьма неплохо на высоких, а также, что зрительный приоритет контраста над абсолютным уровнем яркости согласуется с общей точкой зрения нами принятой.

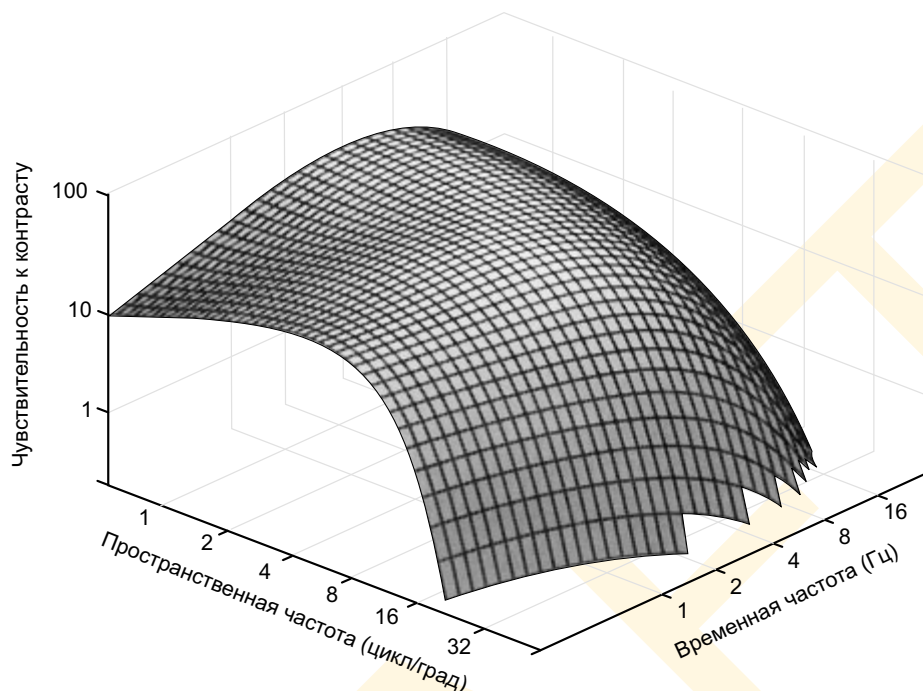
#### 4.3.4.2 Пространственно-временная контрастная чувствительность

На рис. 4.3.4 показаны несколько контрастно-сенситивных функций, построенных с помощью реверсивных синусоид. Полученные данные иллюстрируют функцию контрастной чувствительности, меняющуюся в зависимости от того, по какой временной частоте она построена. На рис. 4.3.20 дан трехмерный график, представляющий то, как пространственные контрастно-сенситивные функции меняются, когда мы строим их на многих временных частотах. По оси абсцисс графика дана пространственная частота тестового стимула; по оси ординат — частота временная; ось аппликат представляет чувствительность зрительной системы к контрасту стимула. При этом мы говорим, что сама поверхность графика являет собой пространственно-временную контрастно-сенситивную функцию наблюдателя.

Отметим, что показанная на рис. 4.3.20 поверхность — это совокупность множества пространственных и временных контрастно-сенситивных функций. Линии параллельные оси пространственной частоты, представляют пространственные контрастно-сенситивные функции; линии параллельные темпоральной оси — временные. Наконец стоит сказать, что из огромного числа психофизических данных американец Дональд Келли (1979 а, b) вывел аналогичную аналитическую кривую, по форме близкую к показанной на рис. 4.3.20.

Из формы поверхности контрастно-сенситивной функции видно, что пространственные контрастно-сенситивные кривые разнятся, когда их строят на разных времен-





**рис. 4.3.20** Пространственно-временная контрастно-сенситивная функция зрительной системы человека. Ось абсцисс и ось ординат представляют, соответственно, пространственную и временную частоты контрастно-реверсивной решетки; ось аппликат — чувствительность наблюдателя к контрасту. Данные, использованные при построении этого графика, были получены при среднем уровне освещенности сетчатки в 1 троланд. Кривые, идущие параллельно оси абсцисс, определяют множество пространственных контрастно-сенситивных функций, построенных на разных временных частотах (см. рис. 4.3.4). Кривые, идущие параллельно оси ординат, представляют темпоральную контрастную чувствительность, измеренную при различных пространственных частотах стимуляции. То есть, ясно видно, что пространственно-временная чувствительность зрительной системы человека неразделима в пространстве и времени.

По материалам Kelly, 1966, 1979 a,b.

ных частотах (рис. 4.3.4). Следовательно, контрастная чувствительность человеческого зрения не разделима в пространстве и времени (Kelly & Burbeck, 1984).

Однако же разделяемость в пространстве и времени — весьма удобное нам свойство.

Во-первых, мы уже говорили о том, что только разделяемые в пространстве и времени системы обладают уникальными пространственными и темпоральными функциями чувствительности. Но когда система несепарабельна, у нее нет уникальной контрастно-сенситивной функции; скорее у нее имеются различные функции по каждому из временных условий.

Во-вторых, пространственно-временная сепарабельность весьма желательна еще и потому, что упрощает вычисления и представления. К примеру, положим, мы хотим представить пространственно-временную контрастно-сенситивную функцию на пространственной частоте  $N_s = 60$  и временной —  $N_t = 60$ . Если контрастно-сенситивная функция не разделяема, нам понадобится записать  $N_s \times N_t = 3600$  значений функции чувствительности. Но если функция сепарабельна, нам нужно будет представить лишь одну пространственную контрастно-сенситивную функцию и одну — временную

( $N_s + N_t = 120$ ). Чувствительность к любому пространственно-временному паттерну можно было бы рассчитать из произведения этих двух функций.

Итак, несмотря на то, что психофизическая реакция наблюдателя неделима в пространстве и времени, нет необходимости, чтобы мы отказались от всех удобств идеи пространственно-временной сепарабельности: мы можем описать психофизическую реакцию наблюдателя как состоящую из комбинации нескольких разделяемых в пространстве и времени механизмов.<sup>1</sup> Мы говорили о таком подходе в контексте разговора о рецептивных полях ганглиозных клеток. И хотя их рецептивные поля не разделяемы в пространстве и времени, можно моделировать их как состоящих из двух пространственно и темпорально самостоятельных компонентов — центра и окружения.

Келли (1971, 1979 a, b; Kelly & Burbeck, 1984) смоделировал пространственно-временную контрастно-сенситивную функцию зрительной системы человека как если бы зрительная чувствительность ограничивалась вкладом лишь двух разделяемых в пространстве и времени компонент. Сие описание — это моноразрешающее описание, весьма похожее на шейдово: ядро свертки системы состоит из центрального и периферического участков, отделяемых друг от друга в пространстве и во времени. Когда оба компонента вместе, как, например, у ретинальных ганглиозных клеток, результирующее ядро свертки неразделяемо. Используя подходящие величины по гауссову и темпоральному параметрам, вполне можно аппроксимировать контрастно-сенситивную поверхность, определяя тем самым выходные параметры ядра свертки. Такое ядро, будучи моноразрешающим, дает удобный метод расчета поверхности, но как мы увидим далее — моноразрешающая система не самым лучшим образом прогнозирует чувствительность по множеству прочих пространственно-временных стимулов, образованных смесью гармонических функций.

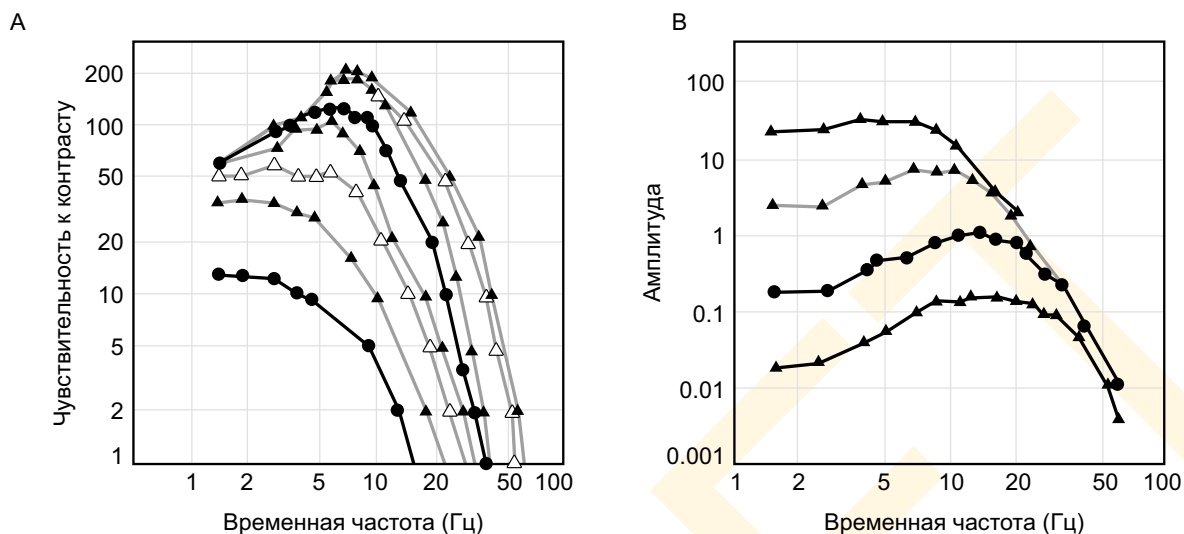
#### 4.3.4.3 Временная чувствительность и средняя фотометрическая яркость

Темпоральная контрастно-сенситивная функция характеризует чувствительность к временным синусоидальным вариациям в контрасте стимула (рис. 4.3.21 А).

Вначале, рассмотрим то, как чувствительность к контрасту меняется на низких временных частотах. При самых низких частотах контрастная чувствительность растет по мере роста средней яркости. Когда фоновая яркость начинает давать 5 троландов на сетчатке, чувствительность к контрасту на низких пространственных частотах меняется меньше чем в два раза, тогда как фоновая интенсивность меняется на два порядка (100 раз). На низких временных частотах чувствительность к контрасту остается более или менее постоянной при изменениях в фоновой яркости. Отметим, что сие одна из форм световой адаптации по Веберу.

Теперь, рассмотрим контрастную чувствительность на высоких временных частотах. Здесь чувствительность систематически растет на всех фоновых уровнях, существенно отклоняясь от веберовской зависимости. Характер девиации лучше виден, если перестроить график — рис. 4.3.21 В, где чувствительность дана как функция от *амплитуды* высокочастотных миганий, но не от контраста (который есть амплитуда, деленная на средний уровень яркости). Когда график дан как функция от амплитуды, кривые временной фликкерной чувствительности сходятся на высоких частотах. Конвергенция была выявлена на множестве разных уровней фоновой яркости и исходя из

<sup>1</sup> Сие действительно возможно, по крайней мере умозраительно, что следует из алгебраического представления, именуемого *разложением по сингулярным матричным числам*.



**рис. 4.3.21** Темпоральная (временная) контрастно-сенситивная функция зрительной системы человека, измеренная при различных уровнях средней яркости стимула. А — темпоральная чувствительность к контрасту. Стимул — диск размером в  $2^\circ$ , расположенный на крупном равномерном фоне. Каждая кривая представляет чувствительность к контрасту (ордината) как функцию от временной частоты (абсцисса) и на различных фонах. Снизу вверх: средняя освещенность сетчатки от собственно стимула составляла 0.375, 1, 3.75, 10, 37.5, 100, 1000 и 10000 троландов. Поскольку фон создавал освещенность примерно в 5 троландов, чувствительность к контрасту на низких временных частотах оставалась примерно постоянной (согласно закону Вебера). В — темпоральная амплитудная чувствительность. Стимул — диск размером в  $60^\circ$ . Каждая кривая представляет порог амплитуды (но не контраста) как функцию от временной частоты. Стимул давал среднюю освещенность сетчатки в 0.85, 7.1, 8.5 и 850 троландов. Отметим, что на высоких временных частотах порог амплитуды сливается в единую кривую, не зависящую от уровня средней яркости стимула. Сие объясняется сугубой линейностью отклика на высоких временных частотах стимуляции, исключающих световую адаптацию.

По материалам De Lange, 1958, Kelly, 1961.

предположения, что чувствительность к сигналам высокой временной частоты прогнозируется амплитудой сигнала, но не его контрастом. Сие реакция, ожидаемая от сугубо линейной системы, то есть вне световой адаптации (и в данном интервале временных частот веберовский закон работает плохо). То есть, данные указывают на то, что световая адаптация не играет существенной роли в видности стимулов высокой временной частоты.

#### 4.3.4.4 Паттерн-хроматическая чувствительность

Нетрудно догадаться, что спектральный состав светового потока радикально влияет на чувствительность зрения к нему.

Одним из важнейших факторов, соотносящих пространственную чувствительность зрения со спектральным составом стимула являются хроматические aberrации в оптической системе глаза, поскольку их последствия оказывают очень серьезное влияние на организацию зрительного тракта. К примеру, хроматические aberrации хрусталика в сочетании с высокой рассеянностью S-колбочек, вкупе ведут к тому, что сигнал, идущий от колбочек этого типа, может представить лишь те решетки, про-

пространственная частота которых не превышает 3-4 цикл/град. При этом по сигналам, инициированным смесью реакций L- и M-колбочек базовое ограничение составляет порядка 50 цикл/град. Последствия сих оптических и нейронных ограничений можно легко оценить по способности наблюдателя к детекции, различению и восприятию окрашенных паттернов: вспомним, что наша способность различать детали в коротковолновых стимулах очень слаба (Williams, 1986).

Действительно, рассуждать о восприятии окрашенных паттернов в логике оптики и колбочковой мозаики довольно просто, но все же вернее и полнее мыслить в логике нейронных путей, кодирующих хроматическую информацию. Важным физиологическим и психофизическим доказательством того, что цветовое восприятие строится на работе проводящих путей, объединяющих сигналы от колбочек трех типов, является тот факт, что один проводящий путь несет сумму колбочковых сигналов, а прочие являются, напомним, хроматически оппонентными проводящими путями, т.е. транслирующими сигналы, представляющие *отличия* между сигналами колбочковыми. Важно отметить при этом, что у сигналов, представленных этими путями, сильно разнятся показатели пространственных разрешений (Noorlander & Koenderink, 1983; Mullen, 1985; Poirson & Wandell, 1993; Sekiguchi et al., 1993 b).

Судя по всему сигналы высоких пространственных частот (20 — 60 цикл/град) возбуждают нейроны только того пути что передает суммарный колбочковый сигнал. Мы ощущаем такие стимулы как модуляции в темную и светлую стороны от средней яркости. Пространственно-частотные паттерны с частотой ниже 12 цикл/град могут оказывать возбуждающее влияние на нейроны пути, кодирующего отличия между L- и M-сигналами. И только стимулы самых низких частот возбуждают нейроны третьего пути — S-колбочкового.

В общих чертах суть описанных эффектов была понятна довольно давно. К примеру, система цветного телевидения организована по принципу яркостного и двух цветоразностных сигналов. При этом только яркостный сигнал несет в себе пространственно высокочастотную информацию — два цветоразностных канала представляют информацию лишь низкой пространственной частоты. Такое представление весьма удобно и экономично в передаче, поскольку высокочастотная информация исключена из двух каналов, позволяя радикально сэкономить полосу пропускания сигнала. Несмотря на отсутствие пространственно-частотной информации телевизионные изображения не воспринимаются размытыми. Причина в том, что пространственно высокочастотная хроматическая информация, исключенная из передачи, обычно и не воспринимается зрением.

Нейронные пути, проводящие хроматическую информацию, разнятся также по своей временной чувствительности. Экспериментальные наблюдения базируются на данных т.н. *фликкерной фотометрии* — опытах, в которых чередуются два тестовых потока: когда смена стимулов идет медленно — наблюдатель воспринимает ситуацию как смену окрашенных объектов; когда стимулы начинают быстро менять друг друга, наблюдатель перестает улавливать хроматические модуляции и все отличия воспринимает как модуляции яркости на равномерно окрашенном фоне. В процессе эксперимента выяснилось, что временное разрешение нашего зрения при различении сине-желтых миганий крайне низкое; среднее — при различении красно-зеленых и очень высокое — при различении светло-темных.

Взаимоотношения между пространственным и временным разрешениями стимуляции адресуют нас к гипотезе, которую мы обсуждали ранее, в частности, к тому, что пространственное и временное разрешения, как говорят, *ковариантны*, поскольку со-

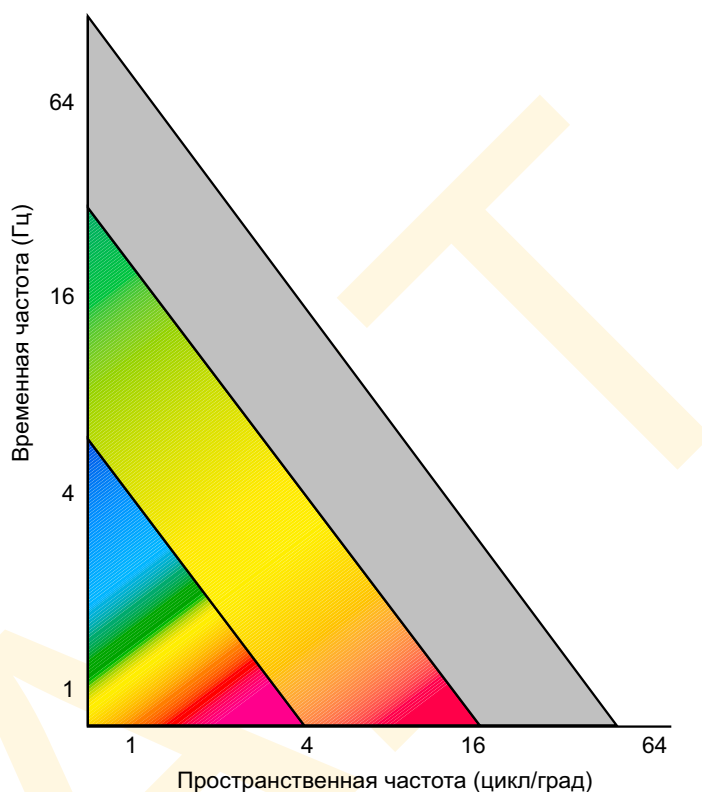


относятся через т.н. *жесткое движение объектов*. Следовательно, в тех диапазонах длин волн, в которых передаются лишь низкие пространственные частоты, зрительная система обычно не требует высокого временного разрешения.

Ковариантность хроматики, пространства и времени проиллюстрирована рис. 4.3.22 — мы видим, как цветовой перцепт меняется на разных пространственно-временных частотах: наше зрение трихроматично лишь в относительно малом диапазоне низких пространственных и временных частот (представленных возле начала координат); по мере роста этих частот мы теряем способность различать сине-желтые хроматические вариации и зрение становится дихроматическим; на высоких пространственных и временных частотах — монохромным, то есть мы воспринимаем лишь яркостные модуляции стимулов.

#### 4.3.4.5 Ретинальная эксцентрика

Все обсуждавшиеся ранее измерения контрастной чувствительности выполнялись по мелким синусоидальным образцам с угловым размером в несколько градусов и по центру поля зрения. Однако по мере приближения к его периферии чувствительность зрения падает.<sup>1</sup> Существует множество нейронных факторов совокупно понижающих как абсолютную чувствительность зрения, так и его пространственное разрешение: плотность колбочковой мозаики стремительно падает по мере удаления от центра сетчатки, и на ее периферии лишь отдельные сенсоры участвуют кодировании оптического сигнала; пространственная плотность ретинальных ганглиозных клеток к периферии тоже падает, из-за чего, собственно, и в коре пе-



**рис. 4.3.22** Наши цветовые ощущения зависят от пространственно-временной частоты стимуляции. Диаграмма делит пространственно-временную плоскость на три участка, соответствующих трем основным режимам цветового восприятия. На низких пространственно-временных частотах цветовое восприятие окрашенных предметов трихроматично: мы ощущаем сине-желтые, красно-зеленые и темно-светлые вариации контраста. На средних пространственно-временных частотах зрение дихроматично: мы теряем способность ощущать сине-желтые контрастные вариации. На высоких частотах зрение становится монохроматичным: любые модуляции цветности мы воспринимаем как модуляции по светлоте.

<sup>1</sup> Качество оптики глаза не оказывает существенного влияния на восприятие в первых 20° зрительного угла (Jennings & Charman, 1981).



риферия сетчатки представлена меньшим числом нейронов; примерно половина первичной зрительной коры представляет центральные  $10^\circ$  поля зрения (314 квадратных градусов), тогда как оставшаяся половина должна представлять оставшуюся часть сетчатки, охватывающую по радиусу примерно  $80^\circ$  (20000 квадратных градусов).

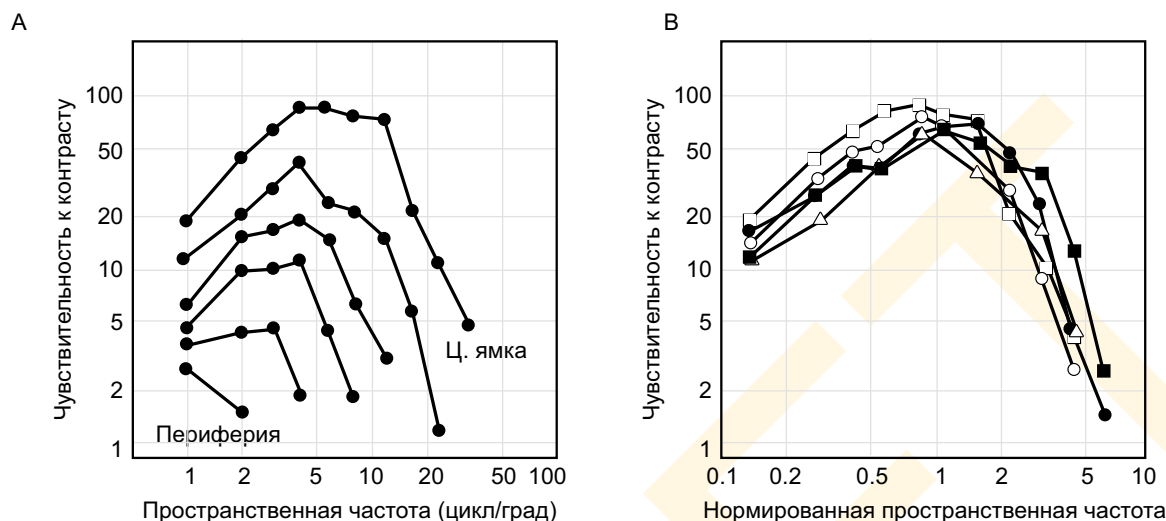
На рис. 4.3.23 А дан набор контрастно-сенситивных функций, измеренных по решетке малого размера и в нескольких эксцентричных позициях в поле зрения. Верхняя кривая — контрастная чувствительность в центральной ямке. Мы видим, что максимум чувствительности зрительной системы приходится на решетки с частотой 5-8 цикл/град, которые наблюдатель обычно распознает от 1% их фотометрического контраста. Фовеальное зрение человека способно распознать решетки с частотой 40-60 цикл/град, однако когда точно такой же стимул расположен на периферии поля зрения, чувствительность падает во всех отношениях: у стимулов, расположенных на  $30^\circ$  к периферии пик контрастной чувствительности приходится всего лишь на 2-3 цикл/град.

Падения контрастной чувствительности нашего зрения мы обычно не замечаем: большинство из нас уверены, что пространственное разрешение зрительной системы равномерно в весьма широком диапазоне размеров — вплоть до  $2^\circ$  визуального угла (большой палец на вытянутой руке). Однако же из кривых рис. 4.3.23А хорошо видно, что острота нашего зрения крайне слаба и ограничена  $7-10^\circ$  (кулак на вытянутой руке). Следовательно, впечатление о высокой остроте нашего зрения частично обязано нашей способности интегрировать пространственную информацию посредством глазных движений.

В 1978 г. итальянец Джирки Ровамо (1978; Rovamo & Virsu, 1979; Virsu & Rovamo, 1979) показал, что падение чувствительности к контрасту по мере роста эксцентричности стимуляции можно объяснить как количественно (снижение представления поля зрения в коре), так и качественно (огрубление нейронного представления периферии поля зрения).

Относительную согласованность между этими факторами продемонстрируют данные рис. 4.3.23 В: контрастно-сенситивные функции, подобные тем, что на рис. 4.3.23А, были получены с разной степенью эксцентричности стимуляции. С целью компенсации ущербности коркового представления размер и пространственная частота решетки шкалировались.

Итак, теперь мы знаем, что при решении любых известных задач на пространственное различение и пространственную локализацию зрительная реакция снижается по мере нарастания эксцентричности стимуляции. Однако физиологическая реакция, как функция от ретиальной эксцентричности, — изменчива и варьирует как между особями, так и между зрительными задачами. Сокращенное периферическое представление имеет место на всех нейронных этапах зрительного акта, начиная с фоторецепторов и продолжаясь в центральной нервной системе. Вариабельность наблюдательской реакции, связанная с большим числом нейронных представлений, для которых характерно аналогичное падение в представлении периферическом, почти исключает возможность возложить ответственность за такое падение на какую-то единичную анатомическую структуру зрительного тракта. Действительно, падение остроты зрения по мере нарастания эксцентричности стимуляции — это важнейшая особенность зрительной системы, но при этом невозможно отыскать анатомическую локаль в зрительном пути, реализующую эту особенность (Westheimer, 1979; Levi et al., 1985; Legge et al., 1987; Yap et al., 1987; Farrel & Desmarais, 1990).



**рис. 4.3.23** Контрастно-сенситивная функция, меняющаяся по мере нарастания ретиальной эксцентрики. А — контрастно-сенситивные функции, построенные с помощью решетки размером  $1 \times 2^\circ$ , стимулирующей сетчатку с величиной эксцентрики в  $0, 1.5, 4, 7.5, 14$  и  $30^\circ$ . Мы видим, что чувствительность к контрасту оказывается наибольшей, когда стимулируется центральная ямка и резко падает по мере удаления от последней. В — контрастно-сенситивные функции, построенные с помощью стимула, с целью приблизительной компенсации неравенства в корковом представлении различных зон сетчатки шкалированного по размеру и пространственной частоте.

По материалам Rovamo et al., 1978

#### 4.3.4.6 Связь теорий

Итак, мы разобрали ряд примеров, в которых вариации психофизических контрастно-сенситивных функций при смене характера стимуляции аналогичны вариациям в откликах ганглиозных клеток сетчатки. Сия корреляция указывает на существование причинно-следственных взаимосвязей между рецептивными полями ганглиозных клеток и психометрическими данными. Однако доказать существование таких связей труднее, чем того хотелось бы.

Психофизические и нейронные теории дополняют друг друга. Психофизики измеряют уровни психологической реакции на физическую стимуляцию и строят затем гипотезы о нейронных механизмах; свойства таких гипотетических механизмов обобщаются и строятся новые психофизические прогнозы. Нейрофизиологические измерения напрямую поставляют нам данные о нейронной активности, но при этом мы должны теоретически вывести то, как нейронная активность влияет на психологическую реакцию. Словом, каждая область вносит свой вклад в наше знание о зрительной функции.

В одной из ключевых глав своей монографии «Физиология сетчатки и зрительных путей», упомянутый уже ранее Джилс Бриндли (1970), озвучил гипотезу, соединяющую данные по обеим областям. Бриндли вывел правило, связывающее физиологию восприятия с его психологией:

«<...>всякий раз когда два стимула модулируют физически неразличимые сигналы, идущие от сенсорного органа в мозг, ощущения, вызванные этими стимулами

(о которых испытуемый сообщает нам словами, жестами или действиями), также должны быть неразличимы.»

Как видим, правило Бриндли само по себе звучит весьма внятно, но главное его достоинство — в междисциплинарности. Наша текущая задача — убедиться в том, что исследователь был последователен: предположение о связи психофизических данных с данными физиологическими рождает порой сомнения в полноте экспериментальной логики каждой из дисциплин.

Отметим, что наиболее прочную связь между психофизическими и физиологическими данными мы установили, когда тщательно и по отдельности изучили данные измерений, например: взаимоотношения между трихроматическими кривыми и чувствительностью фотопигментов оказались очень жесткими, о чем свидетельствуют количественные измерения, полученные во множестве разных условий. Эксперимент по трихроматическому уравниванию не принимает во внимание фотохимию наблюдателя, а диагностика колбочкового фотопигмента не принимает во внимание психофизику зрительного восприятия. Каждый комплект результатов остается сам по себе, но мы можем быть уверены, что их взаимоотношения — это как раз мотив к установке связи между двумя областями знания.

Однако же рассмотрим и иную точку зрения, при которой психофизик пользуется специфическими данными о нейронном отклике как логическим базисом сугубо психофизических экспериментов; или наоборот — когда физиолог толкует набор экспериментальных данных в логике потенциальных психофизических показателей. Такие идеи вполне могут оказаться кстати в постановке специфических экспериментов. Однако логика теорий и опытов, основанная на сетке связей психофизики с физиологией часто путает ход нашей мысли.

Приняв сию схему рассуждений, что же мы должны думать о связи между психофизикой контрастной чувствительности и нейронными рецептивными полями? В текущей главе мы показали, что существует полноценная теория, положенная в основу психофизических контрастно-сенситивных функций и четко согласуемая с логикой организации рецептивных полей. Однако психофизические данные, основанные на контрастно-сенситивной функции, не самодостаточны: мы по-прежнему не умеем применять контрастно-сенситивные функции в оценке чувствительности к другим видам стимулов. Следовательно, согласование между свойствами рецептивных полей и контрастно-сенситивными функциями весьма предварительно, нежели, например, связь между трихроматическими функциями и спектральной чувствительностью фотопигмента.

Но даже несмотря на серьезные издержки нашего сегодняшнего понимания зрения, не стоит отчаиваться: сходство между свойствами контрастно-сенситивных функций и свойствами нейронных рецептивных полей слишком велико, чтобы им пренебречь. Продолжая раздельно совершенствовать психофизические модели и модели рецептивных полей, мы наблюдаем то, что связь между первыми и вторыми крепнет, а количественные сравнения постепенно ведут нас к созданию единой модели, связывающей психофизическую пространственную чувствительность с рецептивными полями нейронов.

#### 4.3.5 Пространственная локализация

В данном разделе нам предстоит обсудить, насколько точно зрительная система человека способна локализовать пространственную позицию стимула. Немецкий физио-

лог XIX века Е. Вульфing (1892) показал, что человек может на удивление тонко различать позиции двух объектов: молодой наблюдатель в состоянии разглядеть зазор между двумя линиями, ширина которого в ретинальном изображении равна одной пятой ширины единичной колбочки. Более того, зазор остается видимым и при движении стимула (Westheimer, 1979).

Способность различать стимулы разных пространственных позиций — это один из аспектов пространственного зрения человека. Важно понимать, что способность локализовать стимул — это вид разрешения отличный от разрешения пространственного, которое мы оцениваем, когда просим наблюдателей разглядеть решетку на равномерном фоне.

Отличия между задачами проиллюстрирует рис. 4.3.24: слева дана ретинальная яркостная картина нескольких стимулов, которые в состоянии воспринимать наблюдатель при решении задачи на пространственное различение. В данном эксперименте испытуемый должен был поймать момент визуального раздвоения физически парного стимула по мере увеличения зазора между его компонентами.

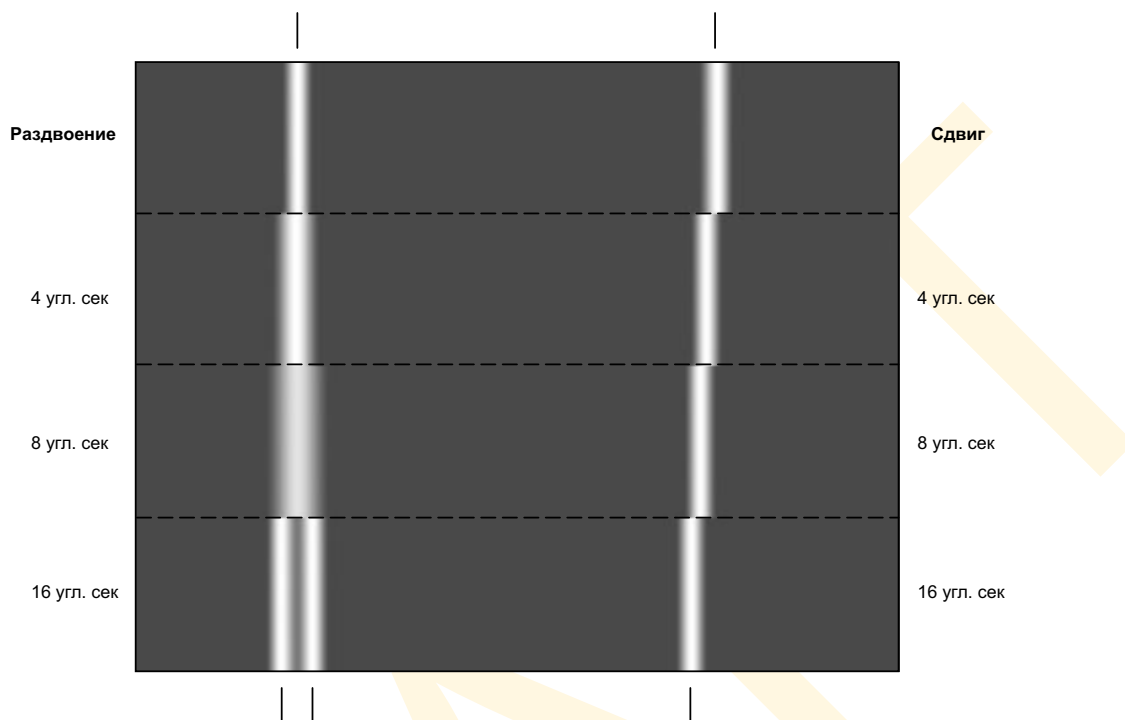
В правой половине рис. 4.3.25 показана световая картина стимулов, предъявленных наблюдателю при решении задачи пространственной локализации. В данном эксперименте испытуемый должен был отличить позицию эталонной линии (вверху справа) от позиции смещенной линии (внизу справа).

В 1977 г. американские нейрофизиологи Геральд Вестхаймер и Сюзанн Мак-Ки (Беркли) диагностировали способность зрительной системы к локализации линий (рис. 4.3.25). Испытуемые определяли, справа или слева от вершины шеврона расположена тестовая линия (см. вставку слева вверху). По ординатам графика отложена величина смещения линии от центра шеврона, необходимая для устойчивого различения этого смещения. Задачу повторяли по шевронам разных углов. На всех углах порог смещения составил порядка 5 дуговых секунд, то есть, грубо говоря, одна пятая от ширины одной колбочки. Реакция при этом не сильно менялась при смене стимула. Сие указывает на то, что детекция пространственной позиции стимула жестко связана с пространственным смещением относительно стимула тестового и относится к самым разнообразным видам стимуляции, включая вариабельность по относительным позициям пар вертикальных линий, движущихся линий и многих-многих других (Westheimer, 1979).

Итак, судя по всему, несколько неожиданно прозвучало то, что мы в состоянии локализовать стимулы с разрешением бóльшим, чем таковое у колбочковой мозаики: вспомним, что по критерию пространственного разрешения восприятие сложных стимулов ограничено пространственной частотой колбочковой мозаики. Почему же тогда характер этой мозаики никак не препятствует нашей способности точно оценивать пространственное положение стимулов?

Действительно, низкая пространственная частота залегания фоторецепторов никак не влияет на нашу способность локализовать стимул. Физический принцип, на который опирается наше зрение для достижения высокой точности в восприятии пространственной локализации при низкой частоте мозаики сенсоров показан на рис. 4.3.26: на большей части графика дана картина колбочковых поглощений, которую мы ожидаем увидеть в качестве отклика на эталонную линию, центрованную по колбочке, а также линию, смещенную вправо на 12 угловых секунд. Цена деления оси абсцисс графика составляет 30 угловых секунд — то есть угловой размер отдельной колбочки.

Поскольку пространственное смещение в сравнении с шагом сенсорной мозаики



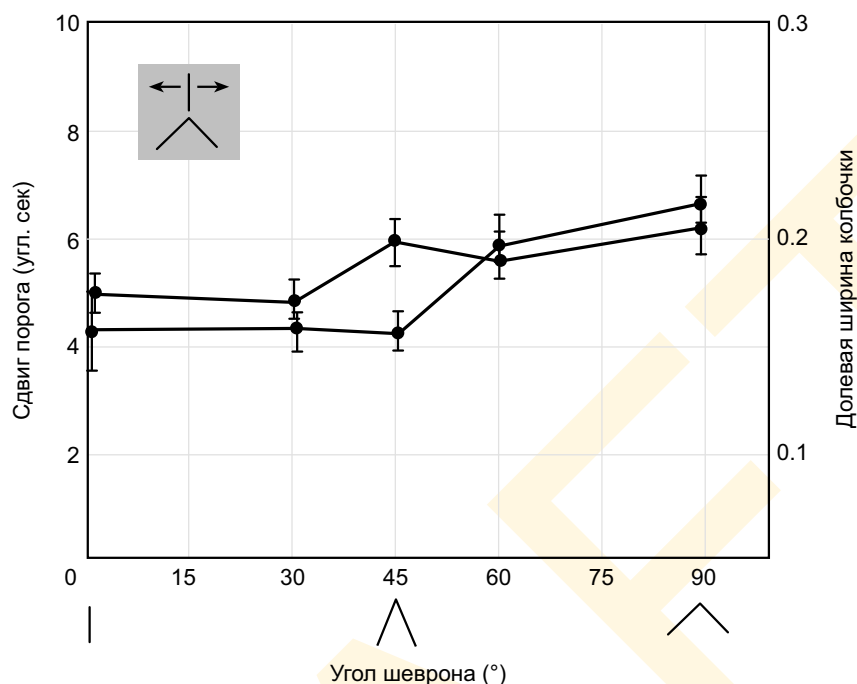
**рис. 4.3.24** Сравнение экспериментов по локализации и пространственному различению. В первом эксперименте (слева) наблюдатели ловили момент визуального различения парности физически парного стимула по мере увеличения углового расстояния между компонентами. Во втором эксперименте (справа) наблюдатели отслеживали смещение вертикальной полоски относительно ее исходного положения. Показано распределение света на сетчатке наблюдателя.

весьма невелико, одни и те же колбочки отвечают как на эталонную линию, так и на тестовую (т.е., сдвинутую). Следовательно, эти колбочки никак не могут послужить в оценке взаиморасположения двух линий. Однако несмотря на то, что одни и те же колбочки отвечают на обе линии, пространственная картина колбочковых поглощений, когда линии находятся в двух разных позициях, оказываются весьма различными. Вставка на рис. 4.3.26 показывает отношение колбочковых поглощений по двум разным линиям: крошечный пространственный сдвиг — всего лишь в 12 угловых секунд — вызывает 50%-е изменение в вероятности поглощения у одной колбочки. Следовательно пространственная картина вероятностей поглощения — это значимый сигнал, который можно использовать для выявления пространственной позиции стимуляции, причем с разрешением бóльшим, чем шаг колбочковой мозаики.

Отметим, что, если мы хотим локализовать линию с большей точностью, чем шаг мозаики сенсора — потребуется оптическое размытие. Там где такого размытия нет, изображение линии окажется внутри ширины фоторецептора и пространственные смещения, меньшие этой ширины, окажутся неразличимыми. Удивительно, но оптическое размытие, которое доставляет столько хлопот при оценке пространственного разрешения контрастных паттернов — помогает в оценке пространственного положения стимула.

Любопытно, что наша способность определять позицию краев и линий не снижается при манипуляциях со стимулом: ежели, к примеру, мы меняем контраст стимула,





**Рис. 4.3.25** Чувствительность к локализации стимула. В ходе эксперимента двух наблюдателей просили определить направление смещение вертикальной линии относительно верхушки шеврона при разных углах его раскрытия. Оба наблюдателя смогли выявить смещение всего лишь при 6 угловых секундах, что в пять раз меньше диаметра единичной колбочки.

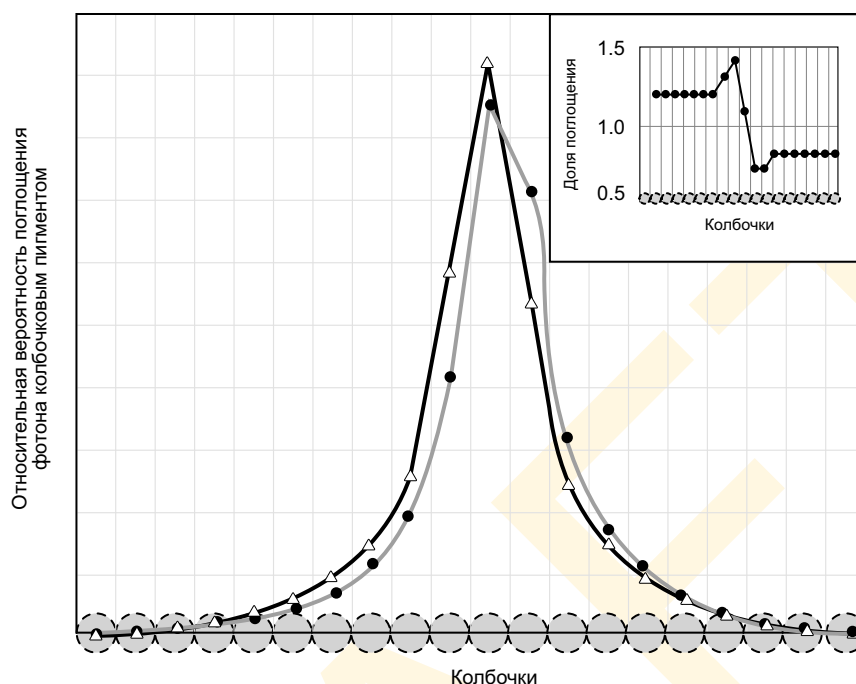
По материалам Westheimer & McKee.

двигаем его или заставляем мигать — реакция все равно остается четкой и устойчивой. Но зрительная система при этом использует невероятно сложные методы «расчета» пространственной позиции стимулов.

Из-за непрерывных глазных движений мы не в силах сравнить выходные сигналы единичных колбочек (или их мелких групп), но мы предполагаем, что зрительная система должна быть способной устанавливать позицию стимуляции с высокой точностью, даже если точность идентификации колбочкового сигнала все время меняется. Несмотря на то, что мы кое-что уже знаем о пространственном зрении, то, как именно зрительная система определяет относительную позицию движущихся объектов (используя при этом активные глазные движения) — остается предметом дальнейшего углубленного исследования.

#### 4.3.5.1 Заключение

Теории пространственной чувствительности зрительной системы человека строятся на нескольких базовых принципах. По началу видность тестовых паттернов различных типов объяснялась свойствами единичной шифт-инвариантной системы: теория проста как для расчетов так и для нашего понимания исходного кодирования световых воздействий нервными клетками сетчатки и их распределения по разным зрительным потокам. Ядро свертки шифт-инвариантной линейной системы и рецептивное поле нейрона играют аналогичную роль и дают добротный базис сравнения психофизических и физиологических данных. Используя общие экспериментальные пока-



**Рис. 4.3.26** Физический базис решения задачи зрительного распознавания пространственной локализации стимула. Точки на главном графике демонстрируют выявленную вероятность поглощения света пигментом колбочек центральной ямки в ответ на стимуляцию тонкой светящейся линией. Треугольнички демонстрируют вероятность поглощения в ответ на эталонную линию, а черные кружки — на линию, сдвинутую на 12 угловых секунд. Серые вертикальные линии делают ось абсцисс на отрезки в диаметр одной колбочки. На вставке показано отношение колбочковых поглощений в ответ на эталонную линию и на смещенную линию в каждой колбочковой позиции. График показывает, что смещение на 12 угловых секунд (примерно одна треть от диаметра колбочки), меняет вероятность поглощения примерно на 50%. Эту информацию зрительная система может использовать для локализации позиции линии с разрешением ощутимо большим, чем колбочковый шаг.

затели, такие, в частности, как контрастно-сенситивные функции, свойства нейронных механизмов и психофизические построения можно сравнивать непосредственно.

Несмотря на то, что концепция моноразрешающего представления дает достаточно серьезное описание пространственной чувствительности зрения человека, она сбивает на многих прямых эмпирических тестах. Поэтому исследователи пытались объединить ее с новым подходом, в котором пространственная информация базируется на совокупности шифт-инвариантных представлений, а не на каком-то одном. Идея сия параллельна физиологической точке зрения о поточности функциональной структуры зрительной системы. Более сложные современные теории описывают большее число ядер свертки (рецептивных полей), и чтобы сохранить данную организацию и увязать ее с некоторыми свойствами корковых рецептивных полей, исследователи обычно выбирают т.н. *профилированные ядра свертки*, то есть «настроенные» на узкий диапазон пространственных частот и ориентаций стимулов. Такие теории способны прогнозировать многие экспериментальные результаты, но при этом все еще остается множество трудностей, которые предстоит преодолеть, прежде чем мы получим полную и удовлетворяющую нас теоретическую схему пространственной чувствительности зрения.

Поскольку зрительная система постоянно адаптируется к новым условиям просмотра, пространственную чувствительность зрения нельзя описать единой функцией. Чувствительность эта меняется по мере изменения временных свойств тестового стимула, средней фоновой яркости, а также спектрального состава стимуляции. Понятно, что общее описание пространственной чувствительности зрительной системы человека должно учитывать все эти факторы.

Психофизические эксперименты показали, что зрение человека исключительно чувствительно к пространственному положению стимула: наблюдатели могут локализовать тестовый стимул с точностью, ощутимо превышающей шаг колбочковой мозаики. Способность к локализации очень сильна и устойчива в очень широком диапазоне стимульных вариаций. Механизмы, позволяющие нам локализовать стимул с высочайшей точностью в целом понятны, но то, как зрительная система пользуется проводящими путями для получения необходимой информации — еще предстоит выяснить.

#### 4.4 КОМПРЕССИЯ ПРОСТРАНСТВЕННОЙ ИНФОРМАЦИИ

Данный выше обзор нейрофизиологии и психофизики пространственного зрения подвел нас к нескольким специфическим гипотезам о том, как зрительная система представляет пространственную информацию. Нейронные данные указывают на то, что зрительная информация делится на зрительные потоки, специализированные на решении разных визуальных задач; данные психофизические — на то, что внутри этих потоков, отвечающих за пространственное восприятие, информация подвергается дальнейшей организации по локальной ориентации, пространственной шкале и хроматике. При этом все имеющиеся у нас по этому вопросу сведения, говорят о том, что контраст ретинального изображения (в отличие от интегральной яркости) — это ключевая переменная, представленная в зрительных путях.

Итак, какие же вопросы мы должны обсудить?

Визуальных задач в природе великое множество, но зрительная система человека способна решать лишь некоторые из них. Вспомним, что один из подходов к разговору о зрительных алгоритмах — это общий подход к зрению, иногда называемый *количественным* или *вычислительным*, при котором мы не ограничиваем ход нашей мысли теми задачами, что важны лишь зрению человека: мы позволяем себе мыслить шире и принимать во внимание те зрительные алгоритмы, которые могут и не иметь биологического эквивалента.

Однако же в нашей книге мы, ограничимся анализом подмножества алгоритмов, непосредственно относимых к зрению человека. Таковых не так уж много, но они очень важны, бо существует множество потенциальных практических приложений, в которых учитывается или имитируется зрительная реакция человека. К примеру, положим, что некто хочет отыскать в электронной библиотеке все изображения зеленых воздушных шаров. Соответствующая программа в компьютере должна запускать в работу алгоритм, логика которого должна быть схожей с логикой человеческого зрения, ведь, в конце концов, прилагательное «зеленый» и словосочетание «воздушный шар» определены нашим восприятием. И это лишь крошечный пример из множества компьютерных алгоритмов, способных помочь человеку в анализе, поиске, создании изображений и манипулировании ими.

Существует и вторая причина тому, чтобы детально изучать зрение человека. Множество исследователей убедительно показали, что исследование зрительных путей и

зрительных реакций — это эффективный метод разработки новых алгоритмов компьютерного зрения. Идея в том, что, изучая специфические свойства биологической зрительной системы, мы приходим к пониманию общих принципов построения зрения компьютерного. Сей процесс аналогичен идее т.н. *реверсивной технологии*, т.е. исследования работы одной системы с целью улучшения эксплуатационных характеристик при разработке другой. Такая точка зрения проповедуется многими авторами, но припоминаемый Дэвид Марр (1982) с особой наглядностью и убедительностью показал то, что биология — это великолепный источник инженерных идей. Однако наша точка зрения на реверсивные методы несколько иная: очевидно, что методы эти больше подходят тем, кто пытается, так скажем, постичь фундамент бытия: очень трудно анализировать работу системы, исходя лишь из ее конструкции, и, если при этом нету знания и понимания основных принципов ее работы. То есть, перетекание идей из одной дисциплины в другую — например, из нейронауки в инженерию и обратно — приносит пользу обеим. И само наличие этого потока — это достаточное основание для того, чтобы исследователь зрения и разработчик систем визуализации как следует разобрались в биологии, психофизике и математике.

В контексте разговоров о зрении человека мы еще не раз вернемся к обсуждению количественных вопросов, а в данном разделе опишем идеи и алгоритмы, связанные с мультиразрешающими представлениями ретинального изображения и компрессией визуальной информации.

#### 4.4.1 Эффективные представления ретинального изображения

Идея мультиразрешающих представлений заложена в различные зрительные алгоритмы — от сегментации ретинального изображения до стереопсиса и движения (к примеру Burt, 1988; Vetterli & Meten, 1992). Однако, дабы обобщить материал, мы рассмотрим его на примере решения отдельных инженерных задач, т.е. задач эффективного сохранения информации, заложенной в ретинальное изображение.

Эффективные и экономичные представления нужны системам с ограниченным ресурсом, которыми являются и все возможные системы зрения: неважно сколько в нашем распоряжении компьютерной памяти или сколько нейронов — мы всегда можем выполнить вычисления точнее, передать информации больше или сохранить изображения с большим качеством, если задействуем лучшие алгоритмы передачи и хранения.

Алгоритмы компрессии преобразуют данные изображения из одного представления в новое представление, требующее ощутимо меньше места для хранения. Чтобы повысить эффективность этих алгоритмов нам нужно каким-то образом описать требуемый объем пространства хранения, и наиболее частый вариант — это его оценка в байтах.<sup>1</sup> Цветные изображения, полученные с фотокамер или сканеров, или каким-то иным способом и предназначенные к показу на экране мониторов или цифровой печати, представлены в показателях световых интенсивностей по каждой из элементарных составляющих — пикселей. Хроматическая информация при этом представлена тремя каналами, как правило поименованными как «RGB».

Мы можем легко посчитать количество байт информации, представленной в изо-

<sup>1</sup> Один бит — это одна бинарная единица информации, которая может быть равной 0 или 1. Один байт — единица информации, состоящая из 8 битов и представляющая 256 уровней сигнала ( $2^8$ ). Мегабит —  $10^6$  байтов. Гигабит —  $10^9$  байтов.

бражении. Например, имеется небольшое изображение, состоящее из 512 рядов и 512 колонок и в каждом из его трихроматических каналов одним байтом информации представлена световая интенсивность. Для трехканального представления потребуется  $512 \times 512 \times 3$  байт информации, или, приблизительно 0.75 мегабайта.

Если в изображении 1024 ряда и столько же колонок, нам понадобится 3 мегабайта.

Видеоролик, в котором изображение обновляется 60 раз в секунду за одну минуту показа затребует 10 гигабайт информации, а за один час — 600 гигабайт.

Легко догадаться, что кодирование хроматической информации активно задействует механизмы *компрессии* изображений, которые в принципе своем возможны лишь благодаря функциональной специфике зрительной системы человека.

Оптические сигналы сцены полноспектральны; изображение этой сцены, однако, не несет в себе полноспектральной информации, но лишь трихроматическую — трехканальную, т.е. весьма и весьма сжатое представление оптической информации. Трихроматическое уравнивание — яркий пример компрессии оптической информации (и характер этой компрессии настолько хорошо изучен, что почти не обсуждается в явном виде).

Помимо трихромазии еще два важных фактора позволяют нам выполнять компрессию изображений без потери их качества:

— во-первых, соседние пикселы в изображениях естественных сцен обычно сходны по интенсивности. Мы говорим при этом, что в этих изображениях имеет место определенная *пространственная избыточность*. Сия избыточность — это составляющая сигнала и дабы получить более эффективное представление составляющую эту можно удалить без всякого риска потерять визуально значимую информацию.

— во-вторых, мы знаем, что пространственное разрешение зрения человека в отношении определенного рода паттернов очень низкое, в частности, в отношении коротковолновых стимулов, да и в целом в отношении стимулов выраженной хроматики. В представлении такой информации в сохраняемом изображении нет необходимости, поскольку приемник (зрительная система) не в состоянии воспринять ее. Понятно, что, избавившись от такой информации, мы повышаем эффективность представления.

Итак, на примере монохромных (яркостных) изображений нам предстоит разобрать эффективные алгоритмы кодирования пространственной яркостной информации, заложенной в ретинальное изображение, посвятив при этом большую часть времени вопросам пространственной избыточности и избыточности по интенсивности.

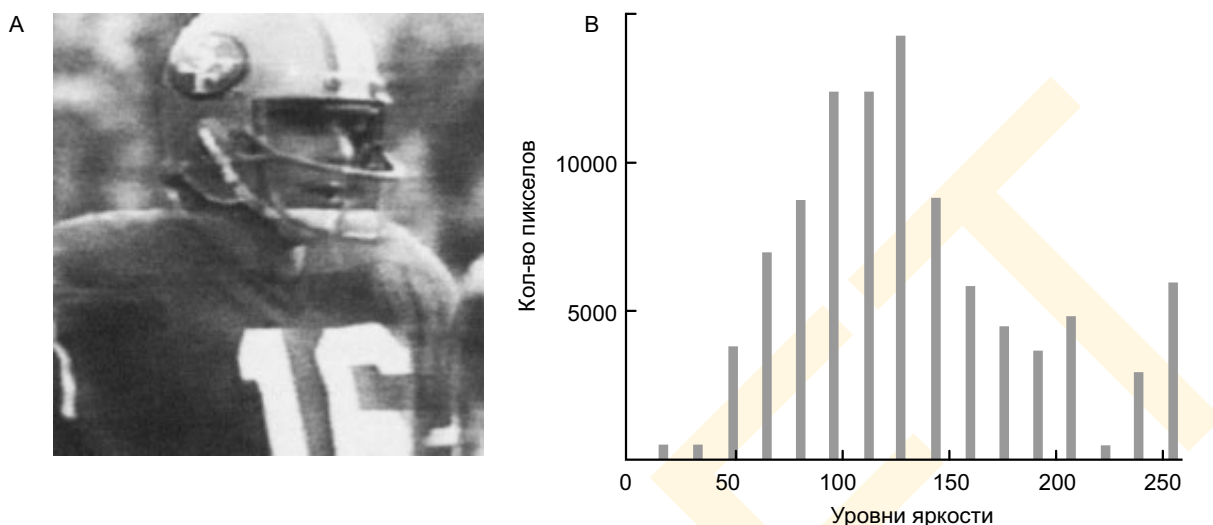
#### 4.4.1.1 Избыточность по интенсивности

Предположим, что у нас имеется изображение (скажем, такое, как на рис. 4.4.1 А), которое мы хотим кодировать с максимальной эффективностью. При этом камера способна зафиксировать 256 уровней яркости (8 бит).

На рис. 4.4.1 В дана *яркостная гистограмма*, в которой по оси абсцисс отложены 256 уровней возможной яркости, а по оси ординат — количество пикселов в изображении. Хорошо видно, что на уровень 128 приходится наибольшее число пикселов.

Несмотря на то, что фотокамера в состоянии представить все 256 уровней интенсивности, в самом изображении такого количества яркостей нет. Следовательно, у нас нет необходимости представлять данные с полным тональным разрешением устройства, но только с таковым у изображения, которое очевидно меньше. Поскольку изображение на рис. 4.4.1 А несет в себе всего лишь 16 уровней яркости (а не 256) мы можем





**Рис. 4.4.1** Распределение яркостей в изображении — весьма важный фактор достижения эффективного представления этого изображения. А — изображение, полученное с помощью фотокамеры, способной зарегистрировать 256 уровней серого, но при этом само изображение несет в себе всего лишь 16 уровней. В — гистограмма, использованная при кодировании этого изображения.

представить их всего лишь 4 битами на точку изображения (а не 8-ю, как способна фотокамера). В итоге мы сэкономим ровно половину объема памяти.

Таким образом, первый принцип экономии места очевиден: нам нет нужды выделять место в памяти под яркости, которых нет в изображении. Мы можем даже модифицировать этот принцип, воспользовавшись тем фактом, что распределение яркостных уровней в изображениях всегда неравномерно: одни встречаются чаще, другие — реже. Например, присвоим однобитовую последовательность (1) для кодирования 128-го уровня; прочие уровни закодируем с применением 5-битовой последовательности, начинающейся с 0, скажем 0xxxx, где xxxx — оригинальный 4-битный код. То есть, к примеру, пятый яркостный уровень кодируется как 00101. Однозначно декодировать входные данные мы можем следующим образом: если первый бит — это 1, то текущее значение равно 128; если первый бит — это 0, то читаем следующие четыре значения для определения уровня интенсивности по данному пикселу.

Теперь представим, что на 128-уровень приходится 60% пикселей. Принятая нами схема кодирования сокращает потребность в объеме памяти, понижением битрейта по этим пикселям с 4 до 1 бита, то есть экономя 3 бита по 60% пикселей. Общая экономия составляет  $0.6 \times 3 - 0.4 \times 1 = 1.2$  бита на пиксел. С помощью сего незатейливого правила мы можем сократить объем до 2.8 бит на пиксел.

Приведенный пример предельно прост, но существует множество намного более сложных алгоритмов, эффективно управляющихся с избытком информации в изображении. В целом, чем больше мы знаем о входном распределении яркостей, тем эффективнее мы экономим код.

Наконец отметим, что были потрачены огромные усилия на разработку общей тактики кодирования как отдельных изображений, так и различных их классов, в частности таких, как технические изображения, изображения естественных сцен и пр.

### 4.4.1.2 Пространственная избыточность

Обычно яркостная гистограмма изображений естественных сцен (естественных изображений) выглядит не так примитивно как на рис. 4.4.1 — в естественных изображениях набор яркостей гораздо богаче, и поэтому алгоритмы, нацеленные лишь на устранение яркостной избыточности, не дают ощутимой экономии места.

Однако в естественных изображениях имеет место т.н. *пространственная избыточность*, и мы можем применить к ней те же принципы кодирования, добившись тем самым определенной экономии места хранения и полосы пропускания сигнала (в частности, в отдельных группах пикселей их яркости могут оказаться весьма близкими друг к другу).

Существуют несколько формальных и неформальных способов диагностики пространственной избыточности в изображении.

Во-первых, рассмотрим опыт американского психолога Дэниеля Керстена (1987), в котором нескольким пикселям изображения (рис. 4.4.2) было присвоено случайное значение яркости, резко отличное от исходного. Испытуемых просили вернуть этим пикселям начальную яркость. При этом оказалось, что когда в изображении искажен 1% пикселей, наблюдатель полностью восстанавливает их в 80% случаев. Но даже когда в изображении перепутана яркость 40% пикселей, наблюдателям удается восстановить исходную картину в половине случаев.

Во-вторых, попытаемся оценить пространственную избыточность в естественных изображениях, сравнив яркости соседних пикселей. На рис. 4.4.3А даны яркости пикселей изображения, показанного в центре иллюстрации. Измеренные по одному за раз, яркости эти распределились по множеству значений и не проявляют явной избыточности ни в одном из участков диапазона. На рис. 4.4.3В показана т.н. *кросскоррелограмма* изображения, на которой по оси абсцисс отложена яркость пиксела —  $p(x, y)$  — а по оси ординат яркость соседнего пиксела —  $p(x, y + 1)$ . Поскольку у соседних пикселей обычно один и тот же уровень яркости, они группируются вокруг одной общей линии. Поскольку интенсивность того или иного пиксела дает нам информацию о *вероятной* яркости его соседа, мы говорим, что яркостная информация избыточна.

Итак, мы можем поднять эффективность представления, избавившись от пространственной избыточности. Один из способов ее ликвидации — это преобразование



Рис. 4.4.2 Экспериментальная оценка пространственной избыточности в изображении, к которому добавлен случайный шум: испытуемых просили изменить яркость шумовых пикселей до подходящего, по их мнению, уровня. Наблюдатели легко и с очень высокой точностью справлялись с этой задачей, опираясь на информацию о яркости соседних пикселей. Сие эффектная экспериментальная демонстрация того, как зрительная система может весьма эффективно воспользоваться пространственной избыточностью ретинального изображения.

На снимке — профессор Кембриджа, нейрофизиолог Хорейс Барлоу.

По материалам Kersten, 1987.

представления. К примеру, вместо независимого кодирования интенсивностей соседних пикселей, мы можем закодировать какой-то один яркостный уровень —  $p(x, y)$  — а также *отличие* по яркости между ними:  $p(x, y+1) - p(x, y)$ . Пара этих значений полностью сохраняет информацию об изображении, поскольку мы можем восстановить оригинальные значения из  $p(x, y)$  и  $p(x, y+1) - p(x, y)$  простым вычитанием. Такой вид сжатия данных именуется *компрессия без потерь*.

После преобразования данных количество бит, требуемое для кодирования  $p(x, y)$  останется неизменным. Но разность  $p(x, y+1) - p(x, y)$  может оказаться в большем диапазоне, вплоть от  $-255$  до  $+255$ , в результате чего нам может понадобиться до 9 бит, чтобы сохранить ее значение. Теоретически, необходимость в дополнительном бите — это существенный недочет алгоритма, но на практике разность эта у соседних пикселей крайне невелика.

Сказанное проиллюстрирует кросскоррелограмма преобразованных значений на рис. 4.4.3 С: по абсциссам отложены величины  $p(x, y)$ , а по ординатам отличие в яркостях соседних пикселей —  $p(x, y+1) - p(x, y)$ . Отметим при этом, во-первых, что большинство значений кучкуются вокруг нулевого значения; во-вторых, что не существует виртуальной корреляции между преобразованными значениями, т.к. знание величины  $p(x, y)$  никак не может помочь нам узнать величину отличия от яркости соседа.

Чтобы создать эффективное представление, воспользуемся тактикой, описанной ранее: для кодирования небольших, но часто возникающих отличий, применим короткий код (скажем 5 бит); для кодирования реже возникающих больших отличий — более длинный (скажем, 10 бит). Поскольку большинство яркостных отличий между пикселями невелико, представление окажется более эффективным.<sup>1</sup>

#### 4.4.1.3 Декоррелирующие преобразования

Тактику компрессии изображения можно разделить на две составляющих.

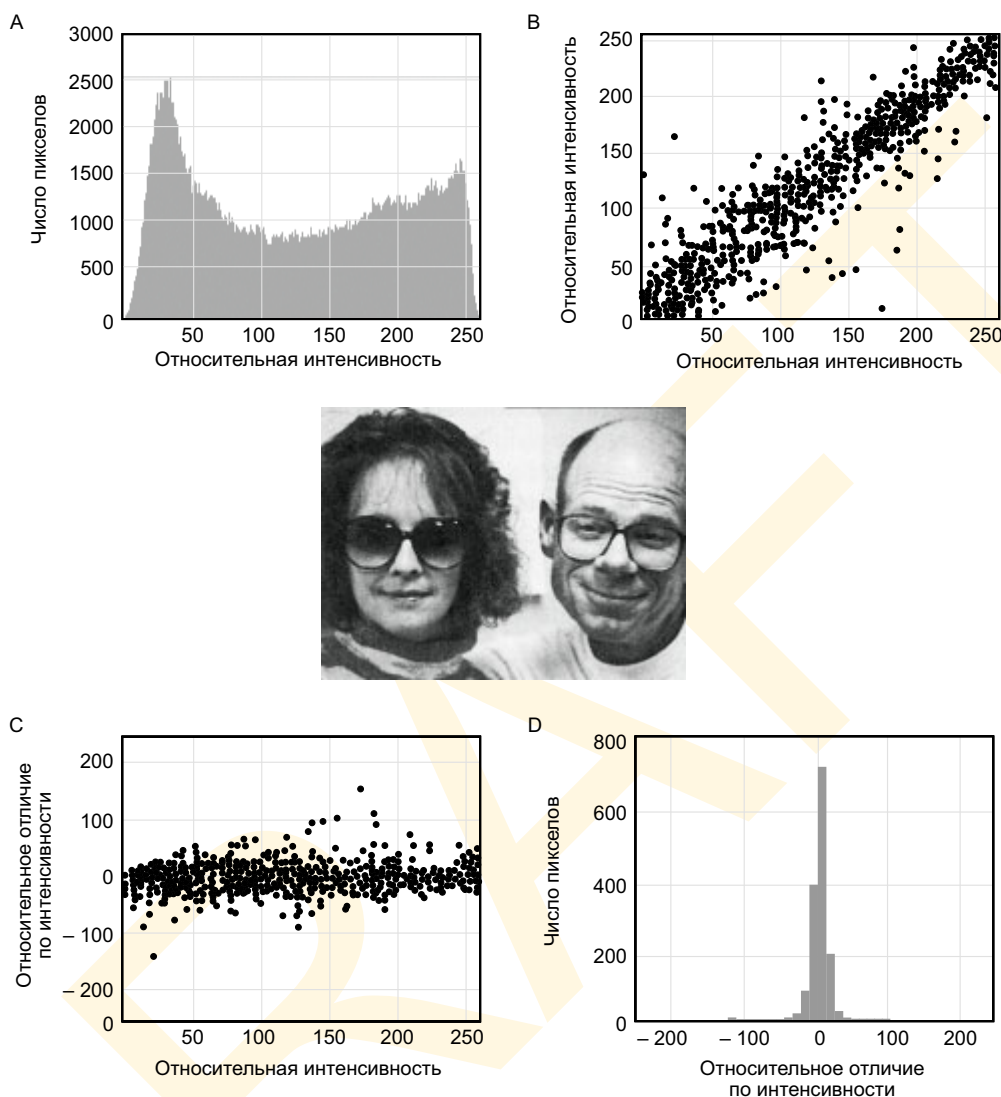
Первая — это линейное преобразование яркостей в новое представление, расчет  $p(x, y)$  и  $p(x, y) - p(x, y+1)$  из  $p(x, y)$  и  $p(x, y+1)$ . Матрица такого преобразования предельно проста:

$$\begin{pmatrix} p(x, y) \\ p(x, y+1) - p(x, y) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} p(x, y) \\ p(x, y+1) \end{pmatrix}$$

Отметим, что мы прибегаем к линейному преобразованию, потому что корреляция преобразованных значений выше, чем в оригинальном представлении.

Суть второй составляющей в том, что новое представление оказывается более эффективным: поскольку мы удалили корреляцию, в изображениях естественных сцен вариация преобразованных значений станет существенно ниже, чем вариация исход-

<sup>1</sup> Даже такую тактику кодирования можно усовершенствовать, например: после первой пары пикселей нам никогда не нужно кодировать величину абсолютной яркости пиксела, — мы всегда можем кодировать лишь отличие по яркости. Такой принцип именуется *ДИКМ* — *дифференциальная импульсно-кодовая модуляция*. Мы можем также рассматривать пару пикселей как вектор, рассчитать частоту распределения всех возможных векторов и построить эффективный код для согласования значений этих векторов. Сие именуется *векторным квантованием* — *ВК*. Все эти методы строятся на том факте, что изображения естественных сцен вероятнее несут в себе пространственные паттерны, чем изображения искусственные.



**Рис. 4.4.3** Количественная оценка пространственной избыточности естественных изображений. Изображение, использованное в данном измерении показано в центре. А — гистограмма, демонстрирующая распределение яркостей в изображении. В — коррелограмма яркостей пикселей в позиции  $(x, y)$  и в позиции  $(x, y + 1)$ . С — коррелограмма отличий по яркости между пикселями в позиции  $(x, y)$  и соседними в позиции  $(x, y + 1)$ . D — гистограмма, демонстрирующая то, что яркостные отличия кучкуются возле нулевого уровня.

ных яркостей пикселей. Следовательно, преобразованные значения кодируются экономичнее, нежели оригинальные.

Из нашего примера можно выделить ключевое свойство линейного преобразования, критично важное в достижении эффективности кодирования: новое преобразование конвертирует данные в *декоррелированные значения*. Значения полагают декоррелированными тогда, когда знание одной величины не дает ничего в плане прогнозирования другой. Интуитивно понятно, что декорреляция — важное слагаемое эффективности: ежели мы можем вывести одно значение из другого, то нет нужды кодировать их обоих.

Разовьем эту идею. Если мы можем примерно вывести одно значение из другого, значит, мы еще не достигли эффективности нашего представления. Мы видим, что в нашем примере величина  $p(x, y)$  — это добротный предиктор величины  $p(x, y + 1)$ . Следовательно, выгоднее исходить из того, что  $p(x, y) = p(x, y + 1)$  и кодировать лишь погрешность такого прогноза. Если предиктор добротен (т.е. корреляция высока) погрешность прогноза займет меньший диапазон, чем исходное яркостное значение. Следовательно, погрешность кодируется несколькими битами, что позволяет нам экономить пространство хранения и/или полосу пропускания сигнала.

Приведенное выше преобразование дает пару примерно декоррелированных значений. Но пример этот слишком прост и возникает необходимость сформулировать декоррелирующее линейное преобразование в общем виде.

Когда множество изображений, которое мы собираемся закодировать, точно известно, то лучшим линейным преобразованием, т.е. компрессией без потерь, явится матричная декомпозиция, именуемая *декомпозицией по сингулярным числам матрицы*. Такая декомпозиция определяет линейное преобразование из исходных яркостных данных в новое представление со статистически независимыми значениями, концентрирующимся в ощутимо меньшем диапазоне. Сие и будет тем самым представлением, которое мы давеча искали для эффективного кодирования.

Отметим, что декомпозиция по сингулярным числам матрицы — это сердце математического анализа главных компонент, но часто звучащее под разными названиями, например: *преобразование Кархунена-Лёве*, *преобразование Хотелинга* и пр.

На практике однако о популяции изображений нет точных данных — ни яркостных, ни статистических. В результате, нет возможности применить метод декомпозиции по сингулярным числам к компрессии изображений, то есть: практический выбор наилучшего алгоритма линейного преобразования достигается только опытным путем.

#### 4.4.1.4 Компрессия с потерями

До сего момента мы рассматривали методы компрессии, не ведущие к каким-либо потерям исходной визуальной информации. Методы эти обычно дают двух-трехкратную компрессию, и мера ее целиком зависит от контекста изображения.

Если же мы готовы допустить некоторое отличие оригинала (а в контексте зрения — ретинального изображения, транслируемого по проводящим путям) от его сохраняемой копии (нейронного изображения), то можем прибегнуть к схемам, позволяющим сэкономить радикально больше места. Преобразования, в определенной мере теряющие исходную информацию, именуются методами *компрессии с потерями*. То есть, мы не можем восстановить исходное волновое представление из закодированного сигнала, но все же мы пользуемся сим ущербным представлением, поскольку, например из фундаментальных экспериментов по трихроматическому уравниванию нам известно, что при достижении колориметрического равенства нет визуальной разницы между воспринятым изображением и оригиналом (см. раздел 3.2.1.1).

Компрессия с потерями порой неприменима, например при сохранении данных на архивное хранение, но в большинстве ситуаций вполне допустимо некоторое искажение оригинала.

Итак, мы можем построить алгоритм компрессии, который хоть и дает некоторую потерю исходных данных, но при этом отличие между оригиналом и его сжатой копией оказывается едва заметным, а экономия при этом составляет от 5 до 10 крат.

Наблюдатель часто оценивает эффективность компрессии по степени визуальных



искажений. В случаях, когда эти искажения незаметны иногда говорят о сжатии без *перцепционных потерь*.

Как было показано ранее, зрительная система человека чувствительнее к одним видам пространственных паттернов и спектров, чем к другим. Методы кодирования без перцепционных потерь активно эксплуатируют это свойство зрения, выделяя больше места для сохранения паттернов высокой видности и меньше — для низкой (Watson & Ahumada, 1989; Watson, 1990).

Алгоритмы сжатия без перцепционных потерь следуют той логике, что и обычные алгоритмы компрессии без потерь. Во-первых, набор данных линейно преобразуется в набор новых значений. При этом дочерние значения нацелены на представление *перцепционно декоррелированных составляющих* изображения (см. ниже). Во-вторых, алгоритм допускает различную степень точности преобразований. Уровень точности в нашем случае зависит от визуальной значимости компонента изображения, представленного данным значением. Следовательно, значимые компоненты занимают больше места, чем едва заметные.

Теперь мы видим, чем и как алгоритмы сжатия с потерями отличаются от таковых без потерь: сжатие без потерь забирает все возможное место и поэтому все участвующие в преобразовании величины представлены с идеальной точностью (несмотря на то, что достигаемая декорреляция экономит место). Алгоритмы компрессии с потерями не резервируют места для сохранения всей оптической информации и поэтому, в частности, изображение невозможно реконструировать во всей его исходной полноте. Однако при этом теряющие алгоритмы построены так, что потеря оптической информации перцепционно незаметна, то есть — сжатое изображение выглядит так же как и оригинал, но при этом требуется радикально меньше места для хранения этого изображения или его передачи по проводящим путям.<sup>1</sup>

#### 4.4.1.5 Перцепционно декоррелированные компоненты

В разговоре об алгоритмах компрессии без перцепционных потерь мы употребили (но не определили) словосочетание «перцепционно декоррелированные компоненты». Акцент на «перцепционно значимых компонентах» весьма широко используется для описания визуально важных параметров изображения, но при этом нет единой точки зрения на то, какие именно параметры определяют перцепционно значимые компоненты изображения. Однако же в контексте разговора о компрессии крайне важно иметь рабочее определение перцепционно значимого компонента. И мы можем дать такое определение в логике линейного преобразования, применяемого для декорреляции данных изображения.

Предположим, что мы представили оригинальное изображение в виде списка интенсивностей ( $i_x$ ) — по одному значению на пиксел. Затем к этим данным, с целью получения нового вектора коэффициентов преобразования ( $t$ ), мы применяем линейное преобразование:

$$\begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_{N-1} \\ t_N \end{pmatrix} = \begin{pmatrix} \text{Декоррелирующее} \\ \text{преобразование} \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \\ \dots \\ i_{N-1} \\ i_N \end{pmatrix}$$

<sup>1</sup> На практике оба варианта компрессии объединены: после применения алгоритма без потерь, начинает работу алгоритм теряющего сжатия.

Коль скоро каждый из коэффициентов преобразования представляет что-то из содержимого входного изображения, мы можем говорить об информации, представленной каждым из коэффициентов преобразования как о *компоненте изображения*.

Согласно такому определению мы можем рассчитать компоненты изображения из векторов коэффициентов преобразования равных нулю за исключением первого элемента. Например, мы можем вычислить компонент изображения, связанный с первым коэффициентом преобразования, решив линейное уравнение:

$$\begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix} = \left( \begin{array}{c} \text{Декоррелирующее} \\ \text{преобразование} \end{array} \right) \begin{pmatrix} i_1 \\ i_2 \\ \dots \\ i_{N-1} \\ i_N \end{pmatrix}$$

Мы говорим при этом, что вектор, решающий данное линейное уравнение — это компонент изображения, представленный первым коэффициентом преобразования. Мы можем найти компоненты изображения, ассоциированные с каждым из коэффициентов преобразования и таким образом найти все компоненты, ассоциированные с каждым из специфических декоррелирующих преобразований.<sup>1</sup>

Теперь нам предстоит определить значение словосочетания «перцепционно декоррелированные» компоненты изображения, для чего мы воспользуемся данным эксперимента Керстена (1987). В этом эксперименте испытуемые управляли яркостями определенных пикселей в попытке угадать уровень интенсивности оригинального изображения. Керстен обнаружил, что наблюдатели диагностировали уровень исходной яркости отдельных пикселей весьма успешно и что при сравнении отдельных пикселей испытуемые четко выявляли корреляцию. Следовательно, отдельные пиксели не подходят на роль декоррелированного компонента.

Вообразим теперь, что мы проводим керстенов эксперимент в одной из его возможных вариаций и предположим, что мы работаем не со случайными значениями яркостей пикселей, а со значениями коэффициентов преобразования. Предположим также, что мы просим испытуемых так менять эти величины, чтобы получить оригинальное изображение. Сей опыт будет идентичен керстеновому за исключением того, что работаем мы с коэффициентами преобразования, а не с отдельными пикселями.

Давеча мы пришли к выводу, что отдельные пиксели не представляют перцепционно декоррелированных компонентов, поскольку наблюдатели справляются с поставленной задачей поиска коррелятов очень эффективно. Мы можем утверждать, что множество коэффициентов преобразования представляет декоррелированные компоненты только в том случае, если наблюдатели плохо справляются с поставленной задачей. Когда известны все коэффициенты преобразования, но при этом мы не помогаем установить значение неизвестного коэффициента, мы скажем, что компоненты, представленные преобразованием, перцепционно независимы. Сегодня нет данных о том, что кто-то проводил перцепционные исследования аналогичные керстеновым по тесту на перцепционную независимость компонент изображения, но, в принципе, такие эксперименты предлагают метод оценки различных алгоритмов компрессии в явном виде.

<sup>1</sup> Отметим, что, собранные вместе, эти компоненты образуют колонки справа, обратные декоррелирующему преобразованию.

Важный этап в компрессии наступает, когда мы используем разную разрядность для представления коэффициентов преобразования. Чтобы решить, какое число бит отдать тому или иному коэффициенту, рассмотрим зрительную чувствительность к тому компоненту изображения, что представлен этим коэффициентом. Поскольку чувствительность зрения к некоторым компонентам крайне невелика, для представления этих компонентов мы можем использовать малое число бит и при этом с крайне незначительными потерями в восприятии этого изображения. Сохраняя информацию на уровне компонентов изображения, мы добиваемся того, что перцепционные искажения оказываются небольшими, а эффективность — огромной.

Описанная тактика компрессии зависима от степени перцепционной самостоятельности компонентов изображения: если компоненты не самостоятельны, то искажение, которое мы привносим в один компонент, может оказать паразитное стороннее влияние на второй компонент: если наблюдатель окажется чувствителен ко второму компоненту, мы испортим изображение. Следовательно, анализ множества перцепционно самостоятельных компонент изображения — это наиважнейшая составная часть разработки алгоритма представления без перцепционных потерь. Если искажения в некоторых компонентах оказывают паразитное влияние на перцепты других компонент, то есть, если представление пары компонентов перцепционно коррелировано, то мы говорим, что линейное преобразование не справилось со своей задачей.

#### 4.4.1.6 Блочное преобразование: JPEG-DCT

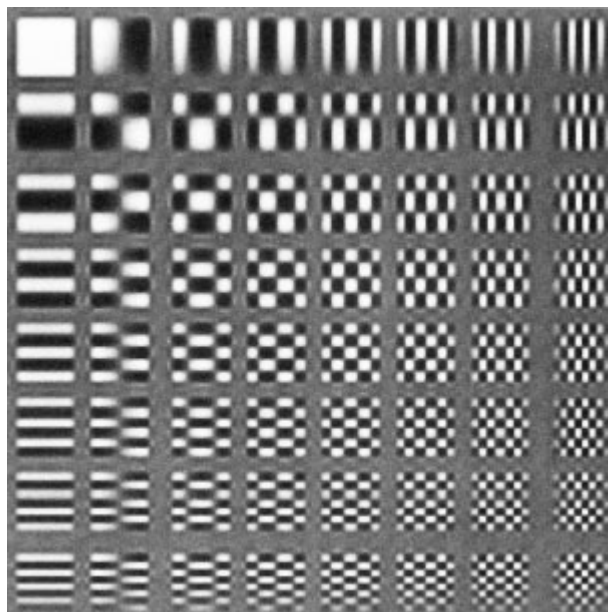
Объединенная группа фотоэкспертов (JPEG — Joint Photographic Experts Group) комитета Международной организации по стандартизации (ISO — International Standards Organization) позиционировала алгоритм компрессии изображений, основанный на линейном преобразовании, именуемом *дискретным косинусным преобразованием* (DCT — Discrete Cosine Transform). JPEG-DCT-алгоритм носит мультиразрешающий характер и очень схож с тем, что происходит при зрительном акте.<sup>1</sup>

При конверсии исходных данных в множество перцепционно независимых компонентов JPEG-DCT-алгоритм использует дискретное косинусное преобразование. Компоненты изображения, ассоциированные с DCT, показаны на рис. 4.4.4. Все они — суть произведение косинусоид различных пространственных частот в двух ориентациях. Следовательно, перцепционно независимые компоненты, подразумеваемые в DCT, — это, в какой-то степени, аналогия набору ориентированных пространственно-частотных каналов. Правда, компоненты не те же самые, что используются моделями зрения, поскольку DCT-компоненты изображения несут в себе сигналы как высоких и низких пространственных частот, так и сигналы взаимноперпендикулярной ориентации. Плюс к тому, имеет место довольно грубое подобие этих компонент ориентированной пространственно-частотной организации моделей мультиразрешающего представления зрения человека. В частности, сие касается компонент, расположенных по краям и по диагоналям рис. 4.4.4, где компоненты эти организованы вдоль линий растущей пространственной частоты и в одной общей ориентации.

Основные этапы работы JPEG-DCT-алгоритма даны на рис. 4.4.5:

— **на первом этапе** оригинальное изображение долбится на блоки. Количественные шаги алгоритма отдельно прилагаются к каждому блоку, то есть, мы говорим, что ал-

<sup>1</sup> Дискретное косинусное преобразование аналогично вычислению рядов Фурье, описанному ранее.



**Рис. 4.4.4** Перцепционные особенности DCT как произведение гармонических функций  $\cos(2\pi f_1 j) \cos(2\pi f_2 k)$ , где  $j$  и  $k$  — это позиции по горизонтальному и вертикальному направлениям. Функции могут иметь как положительные, так и отрицательные значения и показаны в виде контрастных паттернов, отклоняющихся от некоей средней величины (фона).

мы говорим, что коэффициенты эти весьма подходящи для компрессии без потерь. Действительно JPEG-DCT-алгоритм несет в себе субалгоритм компрессии без потерь, приложенный к квантованным коэффициентам. Сие представление используется для хранения или передачи данных изображения.

Итак, для реконструкции и аппроксимации оригинального изображения, нам нужно лишь применить обратный DCT к квантованным коэффициентам. Поскольку данные дискретны, то результат реконструкции, безусловно, будет чем-то отличен от оригинального изображения. Коль скоро в основном мы избавились от информации о пространственно-высокочастотных компонентах изображения, отличие между оригиналом и его реконструктом явит собой некое изображение, состоящее в основном из тонких текстур (отмечено как «Погрешность» на рис. 4.4.5).

Отметим, что одним из базовых ограничений алгоритма является деление оригинала на отдельные блоки (что требуется для независимой обработки). Пикселы, расположенные на краях этих блоков реконструируются независимо от яркости соседних пикселов в следующем блоке. То есть, наиболее значимым артефактом реконструкции является видность искажений на краях этих блоков, что часто называют *блок-артефактами*.

Итогом сказанного выше является то, что с пространственным зрением человека JPEG-DCT-алгоритм связывают два момента :

— во-первых, используется грубое мультиразрешающее представление. Вообразим

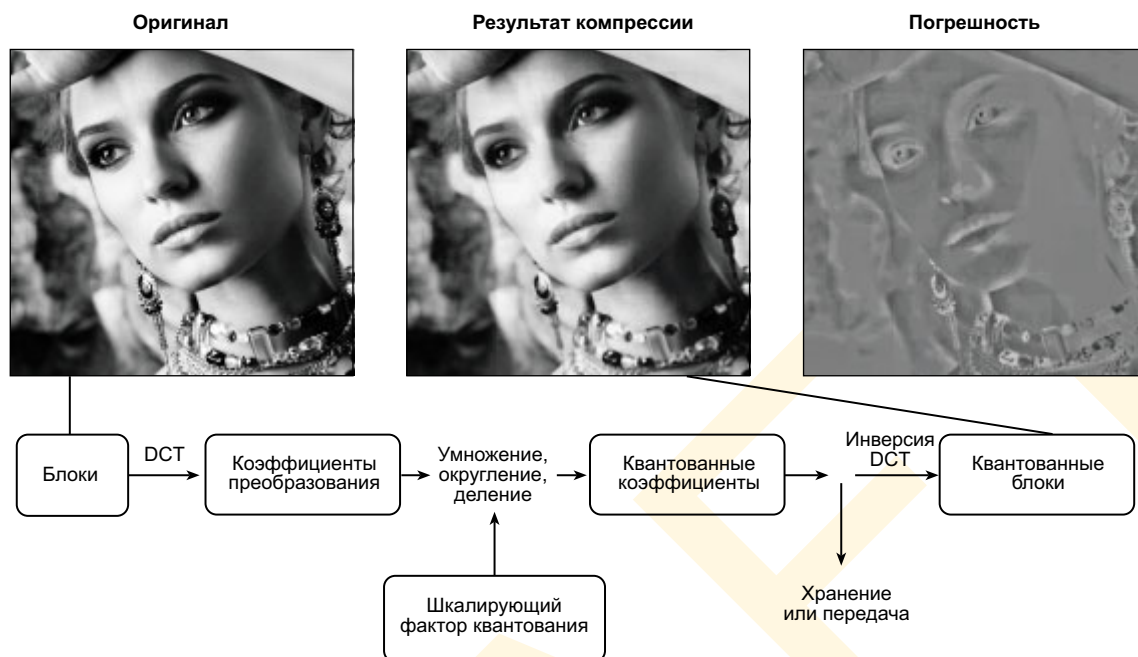
горитм пригоден к параллельному исполнению. Размер блока обычно составляет  $8 \times 8$  пикселов, но может быть и бóльшим. То есть, мы говорим об алгоритме *поблочного кодирования*;

— **на втором этапе** данные в каждом блоке линейно преобразуются по DCT;

— **на третьем этапе** коэффициенты преобразования квантуются. И это первый этап, на котором происходит компрессия. Квантование реализуется умножением каждого из коэффициентов на шкалирующий фактор, лежащий между 0 и 1, с последующим округлением результата до ближайшего целого и делением на шкалирующий фактор. Если шкалирующий фактор близок к нулю, то операция округления дает сильный эффект и значение коэффициента квантования оказывается небольшим. В данном примере выбраны факторы близкие к 1 по низким пространственным частотам (слева вверху на рис. 4.4.4) и близкие к нулю по высоким (справа внизу).

Отметим, что множество квантованных коэффициентов равны нулю, и





**Рис. 4.4.5** Схема работы JPEG-алгоритма, основанного на DCT. Оригинальное изображение делится на серию наклаывающихся друг на друга квадратных блоков, обычно в размер  $8 \times 8$  пикселей. Информация в изображении (с применением DCT) преобразована в новое множество коэффициентов. Коэффициенты преобразования квантованы с применением простых операций умножения, округления и деления. Квантованные коэффициенты, обнуленные этой операцией, делают изображение весьма подходящим к эффективной компрессии без потерь, применяемой перед сохранением файла изображения или его передачи. Для реконструкции оригинала квантованные коэффициенты конвертируются с помощью обратного DCT, приближающего сжатое изображение к оригиналу. Погрешность реконструкции (т.е. отличие между оригиналом и реконструктом) касается в основном пространственно-высокочастотных текстур, зрительная чувствительность к которым крайне мала.

группирующиеся вместе коэффициенты, полученные из отдельных блоков. Внутри каждого блока находятся 64 DCT-коэффициента, соответствующих 64-м компонентам изображения. Собрав с каждого блока соответствующие коэффициенты преобразования, получим количественную меру каждого компонента изображения внутри каждого блока. В неявном виде DCT-коэффициенты определяют 64 изображения и каждое описывает вклад одного из 64 компонент. Эти внутренние изображения аналогичны набору нейронных изображений, формирующих мультиразрешающую модель пространственного зрения.<sup>1</sup>

— во-вторых, JPEG-DCT-алгоритм базируется на той гипотезе, что квантование по

<sup>1</sup> Отметим, что каждый блок для представления пространственно-низкочастотной информации несет в себе то же количество коэффициентов, что и для представления пространственно-высокочастотной. Согласно правилу Найквиста (см. гл....) мы можем представить пространственно-низкочастотную информацию с помощью много меньшего числа отсчетов, чем то необходимо для представления высокочастотной информации. Почему разница в частоте дискретизации не учитывается JPEG-алгоритмом? Частично причина в блочном кодировании, частично — в свойствах компонентов изображения.



коэффициентам высоких пространственных частот не влияет на качество компонентов изображения, кодированных коэффициентами низких. Если пониженное разрешение высоких пространственных частот влияет на четко заметные компоненты изображения — алгоритм не работает. Следовательно, условие перцепционной независимости компонент — очень важное условие успешной работы алгоритма.

Независимые компоненты в JPEG-DCT-алгоритме в целом согласуются с мультиразрешающей организацией моделей пространственного зрения человека, но неидеально: высоко- и низкочастотные элементы иногда смешиваются внутри одного компонента; плюс к тому, элементы сильно различных ориентаций также могут объединяться в единый компонент.

В следующем разделе мы рассмотрим мультиразрешающие представления, которые несколько лучше, чем JPEG-DCT отражают характер передачи данных в зрительной системе человека.

#### 4.4.1.7 Пирамиды

Пирамиды — это мультиразрешающие представления, формат которых отличен от JPEG-DCT по ряду параметров. Два наиболее важных отличия, вероятно, состоят в том, что:

- алгоритмы пирамид не дробят изображение на блоки;
- пирамидное мультиразрешающее представление ближе к представлению визуальной информации в зрительной системе человека, нежели представление JPEG-DCT.

В 1975 г. Тео Павлидис и Стивен Танимото одними из первых начали заниматься данным вопросом. В частности, они пишут:

«Наша точка зрения такова, что ключ к эффективному анализу ретинальных изображений лежит в способности зрительной системы, во-первых, находить значимые фрагменты этих изображений очень быстро и, во-вторых, моментально игнорировать бесполезные детали. Сетчатка человеческого глаза “структурирована” так, чтобы в широком угле воспринимать пространственно-низкочастотную (“высокоуровневую”) информацию за счет периферического зрения, и, одновременно посредством центральной ямки воспринимать информацию пространственно-высокочастотную, детализированную.»

Линейные преобразования, заложенные в алгоритмы пирамид, оперируют компонентами изображения, представляющими собой периодические паттерны различных пространственных ориентаций, что весьма и весьма сходно с мультиразрешающими моделями зрения человека. Поскольку коэффициенты в пирамидах представляют данные, распределенные по разным пространственно-частотным каналам, то для представления этих каналов можно использовать различные количества коэффициентов преобразования: для представления пространственно-низкочастотной информации пирамиды задействуют небольшое число коэффициентов и много большее — для высокочастотной. Уменьшение числа представляющих коэффициентов по мере снижения пространственной частоты и привело к появлению термина «пирамида».

#### 4.4.1.8 Операции с пирамидами.

Пирамиды зиждутся на двух базовых операциях, которые в целом взаимнообратны: первая операция — это размывание и квантование входных данных; вторая — ин-

терполяция размытого и квантованного изображения до оригинала. Обе операции линейны.

**Свертка и квантование.** Для простоты мы опишем манипуляции с пирамидами как с одномерными сигналами, заранее оговорившись, что при работе с двухмерными изображениями не меняется ровным счетом ничего. В конце раздела мы обсудим то, как распространить одномерный анализ на двухмерные изображения.

Итак, предположим, что мы начали с одномерного входного вектора ( $g_0$ ), несущего  $n$  значений.

Первая базовая пирамидная операция — это свертка с размывающим ядром и последующее квантование результата. Размытие и квантование идут вместе, поскольку результатом размытия является производное оригинального изображения с меньшим количеством компонентов высокой пространственной частоты. Поскольку размытие ликвидирует высокочастотную информацию, то, согласно Найквисту, мы можем представить размытые данные, используя меньшее число отсчетов, чем того потребовал бы оригинал. Прделаем сие с помощью квантования размытого изображения по каждой следующей величине.

Как было сказано ранее, свертка и квантование — линейные операции. Следовательно, мы можем выразить каждую матричным умножением: ядро запишем как матричное умножение  $B_0 g_0$ , где ряды  $B_0$  несут ядро свертки; квантование выразим в виде прямоугольной матрицы  $S_0$ , чьи коэффициенты равны нулям и единицам. Объединив операцию размытия и квантования получим базовую матрицу пирамиды:  $P_0 = S_0 B_0$ .

Умножение входных данных на  $P_0$  дает редуцированную версию оригинального изображения —  $g_1 = P_0 g_0$  — содержащую лишь половину коэффициентов.

В матричном виде операторы размытия и квантования выглядят так:

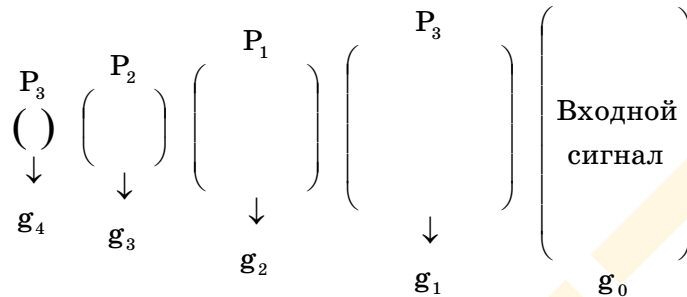
$$P_0 = \begin{pmatrix} S_0 & B_0 \end{pmatrix}$$

Для создания собственно пирамиды изображения повторим свертку и квантование по каждому из результирующих изображений: первая операция строит из оригинала  $g_1$  редуцированное изображение; для создания следующего уровня пирамиды ( $g_2$ ) мы размываем и квантуем  $g_1$ , получая  $g_2$ ; затем размываем и квантуем  $g_2$ , получая  $g_3$  и т.д. Когда входной сигнал одномерен, то каждый последующий уровень несет половину дискретных значений предыдущего.

Когда изображение двухмерно, квантование применимо как к рядам, так и к колонкам, и каждый следующий уровень разрешения несет только одну четверть дискретных значений оригинала. Сказанное в матричной форме проиллюстрирует рис. 4.4.6.

**Интерполяция.** Вторая базовая пирамидная операция — интерполяция — это операция обратная размытию и квантованию, и, несмотря на то, что точная инверсия невозможна, мы можем использовать  $g_1$  как основание нашей догадки о  $g_0$ . То есть, ежели имеется определенная избыточность пространственного сигнала, наша догадка об оригинальном изображении будет недалеко от истины, т.е.:

⇒ Интерполяция — это процесс т.н. *обоснованного предположения* об оригинале из его редуцированной копии.



**Рис. 4.4.6** Матричное представление одномерных пирамидных операций. Базовая операция состоит из размытия и последующего квантования сигнала. Для создания серии изображений понижающегося пространственного разрешения мы несколько раз повторяем размытие и квантование. После каждого повтора получаем новое редуцированное изображение —  $g_i$ .

Для расчета входного сигнала выберем *матрицу интерполяции*, скажем  $E_0$ . Следовательно:  $E_0 g_1 \approx g_0$ .

Теперь расположим обе базовых пирамидных операции в виде т.н. *конструктивной последовательности* (что мы неоднократно еще сделаем в текущей главе): во-первых, мы преобразуем входные данные путем свертки и квантования —  $g_1 = P_0 g_0$ ; во-вторых, выдвинем нашу лучшую версию об оригинале с использованием матрицы интерполяции  $g'_0 = E_0 g_1 = E_0 P_0 g_0$ .

Отметим, что у *интерполята*, т.е. у расчетного изображения  $g'_0$ , будет тот же самый размер, что и у оригинала.

Наконец, для сохранения всей информации, создадим отдельное изображение, несущее *погрешности*. Погрешность — это разность между подлинным сигналом и интерполированным —  $e_0 = g_0 - g'_0$ . Сие завершает расчет первого уровня пирамиды.

Для завершения строительства всех уровней пирамиды применим ту же последовательность операций, но начнем с первого уровня —  $g_1$ : построим новую матрицу свертки —  $B_1$ ; квантование проведем с помощью  $S_1$ ; построим  $g_2 = S_1 B_1 g_1$ ; интерполируем  $g_2$  с помощью матрицы  $E_1$ ; и, наконец сформируем новое *изображение погрешностей* —  $g_1 - g'_1$ , где  $g'_1 = E_1 g_2$ .

Чтобы построить полную пирамиду, повторим процесс, сокращая на каждом шаге число элементов, и остановимся тогда, когда решим, что редуцированное изображение —  $g_n$  — достаточно мало и что дальнейшего размытия и квантования не требуется.

Конструкция пирамиды определяет три последовательности сигналов: серию размытых и квантованных сигналов (размер которых был плавно уменьшен), серию интерполированных сигналов и сигналов погрешности. Их взаимоотношения мы можем подытожить несколькими простейшими уравнениями.

Во-первых, редуцированное изображение на  $i$ -уровне создается приложением базовой пирамидной операции к предыдущему уровню:

$$g_i = P_i - 1 g_{i-1}$$

Во-вторых, реконструкцию изображения  $g'_i$  мы выполняем из представления более низкого пространственного разрешения:

$$g'_i = E_i g_i + 1$$

Наконец, отличие между оригиналом и реконструктом явит собой изображение погрешностей:

$$e_i = g_i - g'_i = g_i - E_i P_i g_i$$

В итоге, информацию об оригинале хранят два разных множества сигналов.

Первое — это последовательность из входных данных оригинала и последовательность *редуцированных сигналов* —  $g_1, g_2, g_3 \dots g_n$ , в которой дается описание сигналов оригинала все с меньшим и меньшим пространственным разрешением (но при этом вся информация об оригинале сохраняется, поскольку сам оригинал — это часть последовательности. Описанное множество отвечает за представление пространственно-низкочастотных версий изображения.

Второе множество — это последовательность состоящая из *сигналов погрешности* —  $e_0, e_1 \dots e_{n-1}, g_n$  (особо подчеркнем, что  $g_n$  — это тоже часть последовательности).

Может показаться неожиданным то, что и эта последовательность несет в себе всю информацию об оригинальном (т.е. ретинальном) изображении. Чтобы осознать сей факт, отметим, что мы можем построить последовательность изображений —  $g_i$  — из сигналов погрешности. Показателей  $g_n$  и  $e_{n-1}$  достаточно для того, чтобы получить  $g_{n-1}$ ;  $g_{n-1}$  и  $e_{n-2}$  позволяют восстановить  $g_{n-2}$  и т.д. В конечном итоге, мы используем  $e_0$  и  $g_1$  для реконструкции оригинала —  $g_0$ .

Итак, у нас есть все основания полагать, что нечто подобное происходит с информацией о ретинальном изображении при его передаче по проводящим путям в высшие отделы мозга. Рассмотрим теперь математическую модель наиболее вероятного, на наш взгляд, сценария нейронных событий.

#### 4.4.1.9 Примеры пирамид

На рис. 4.4.7 показан процесс построения пирамиды по алгоритму, в 1983 г. предложенному американцами Петером Бартом и Эдвардом Адельсоном (Массачусетский технологический институт), которые, возможно, оказались первыми, кто применил понятие «пирамида» в технике кодирования изображений.

Иллюстрация начинается с одномерного квадратно-волнового входного сигнала —  $g_0$ , размытого при помощи гауссова ядра свертки, после чего квантованного по каждой локали; редуцированный сигнал —  $g_1$  — дан ниже. Когда ядро свертки являет собой гауссову функцию, последовательность редуцированных сигналов называют *гауссовой пирамидой*.

Для интерполяции редуцированного сигнала к более высокому разрешению Барт и Адельсон предложили следующую специальную процедуру: данные в  $g_1$  располагаются через элемент с тем же числом элементов, что в  $g_0$ . Алгоритм назвали *восходящим квантованием*, что эквивалентно умножению вектора  $g_0$  на транспонированную матрицу квантования  $S_0^T$ .

На следующем этапе квантованный вектор сворачивается с *примерно* тем же гауссовым ядром, что использовалось для редуцирования изображения. «Примерно» пото-

му, что ядро это все же несколько отлично гауссова ядра, использованного для размытия сигналов. То есть, в данном примере матрица интерполяции — это *удвоенная* транспонированная матрица свертки-квантования:

$$\mathbf{E}_0 = 2\mathbf{B}_0^T \mathbf{S}_0^T = 2\mathbf{P}_0^T$$

Интерполированный сигнал, то есть — результат расчета сигнала высокого разрешения, показан в центральной колонке рис. 4.4.7.

Следующим этапом мы вычисляем сигнал погрешности — разность между расчетным сигналом и оригинальным. Сигналы погрешности показаны в правой колонке рис. 4.4.7. Последовательность сигналов погрешности образует *пирамиду погрешностей*. Как было показано выше, мы можем идеально реконструировать оригинальное изображение  $\mathbf{g}_0$  из сигналов  $\mathbf{e}_0$  и  $\mathbf{g}_1$ . Барт и Адельсон назвали сигналы погрешности, полученные комбинацией гауссова размытия с функциями интерполяции — *пирамидой Лапласа*.

На рис. 4.4.8 дан результат приложения пирамидного процесса к двумерному сигналу, то есть — к изображению. Последовательность редуцированных изображений, формирующих гауссову пирамиду, дана вверху (с оригинальным изображением слева). Изображения эти получены размытием оригинала с последующим представлением новых данных с вдвое меньшей частотой квантования по рядам и колонкам. Таким образом, в двумерном варианте каждое редуцированное изображение содержит лишь четвертую часть от числа коэффициентов своего предшественника.<sup>1</sup>

Последовательность изображений погрешности, образующая лапласову пирамиду, показана на рис. 4.4.8 В. Поскольку для расчета изображений низкого разрешения интерполяция задействует гладкую гауссову функцию, наибольшие погрешности возникают возле краев изображения. Но поскольку изображения обычно плавны (соседние пиксели коррелируют) большая часть погрешности (не считая краев) невелика.<sup>2</sup>

#### 4.4.1.10 Компрессия с помощью пирамиды погрешностей

С точки зрения компрессии последовательность изображений в гауссовой пирамиде не столь интересна, поскольку несет в себе оригинал. А вот пирамида Лапласа привлекает внимание минимум по двум причинам.

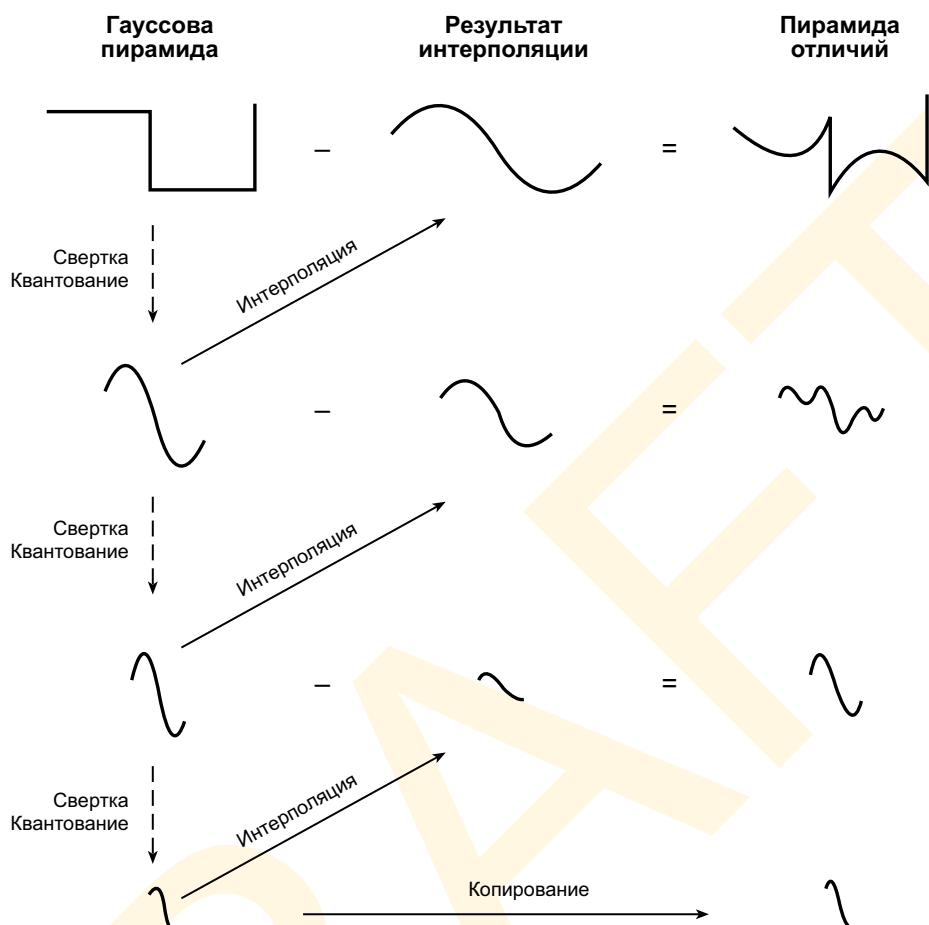
Во-первых, информация, представленная лапласовой пирамидой систематически варьирует по мере снижения пространственного разрешения: на высших уровнях, несущих наибольшее число коэффициентов преобразования, пирамида представляет мелкие детали изображения; на низших уровнях, несущих наименьшее число коэффициентов, «Лаплас» передает пространственно-низкочастотную информацию. Очевидно, что сие именно так, потому что изображение погрешностей — это отличие между оригиналом, несущим все детали, в том числе и самые мелкие, и расчетным оригиналом, базированным на слегка размытой копии.

Разница между оригиналом и оригиналом, высчитанным из размытой копии, представляет информацию в полосе пространственных разрешений, лежащую между

<sup>1</sup> Следовательно, мы умножаем матрицу интерполяции не на 2, а на 4 (т.е.  $\mathbf{E}_0 = 4\mathbf{P}_0^T$ ).<sup>1111</sup>

<sup>2</sup> С целью демонстрации изображений погрешностей, которые несут в себе отрицательные коэффициенты, яркости шкалированы так, что тени имеют отрицательные значения, средние тона — возле нуля, а света — положительные.





**Рис. 4.4.7** Одномерная пирамидная конструкция. Входной сигнал (слева сверху) сворачивается с гауссовым ядром, а результат дискретизируется. Результатом является размытая копия сигнала на наименьшем пространственном разрешении. Реконструкция оригинала — плод интерполяции. Отличия между оригиналом и реконструктом сохраняются в виде пирамиды погрешности. Процесс повторяется, начинаясь с размытой копии, благодаря чему образуется серия копий все меньшего пространственного разрешения (слева) и серия изображений погрешности (справа). Сигнал наименьшего пространственного разрешения сохраняется как как макушка пирамиды погрешности.

двумя уровнями. Таким образом, лапласова пирамида — это мультиразрешающее представление оригинального (т.е. ретинального) изображения.

Во-вторых, значения коэффициентов преобразования в изображениях погрешностей занимают много меньший диапазон, чем яркости пикселей в оригинале. На рис. 4.4.9 А дана диаграмма интенсивностей пикселей в первых трех элементах гауссовой пирамиды: мы видим, что она весьма широка и поэтому не очень подходит к решению задач компрессии. На рис. 4.4.9 В даны гистограммы величин лапласовой пирамиды: четко видно, как коэффициенты преобразования кучкуются возле нулевой отметки, что радикально повышает эффективность представления. Отметим, что (благодаря пространственной корреляции в естественных изображениях, которая позволяет нам успешно работать с расчетными изображениями, полученными методом сглаживающей интерполяции) диапазон коэффициентов преобразования в лапласо-



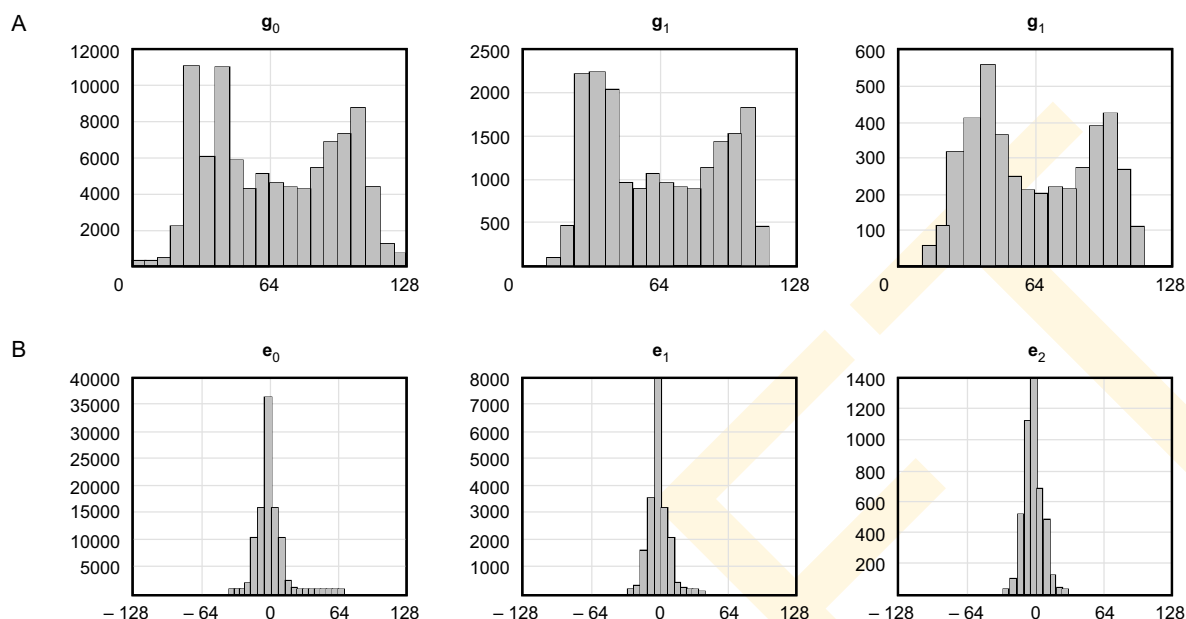
**Рис. 4.4.8** Гауссовы и лапласовы пирамиды. А — серия редуцированных изображений, образующих гауссову пирамиду, начинающуюся оригиналом слева. Оригинал размыт с применением гауссова ядра свертки, а затем квантован. Операция дала изображение меньшего пространственного разрешения и меньшего пространственного размера. В — каждое из редуцированных изображений гауссовой пирамиды можно использовать для реконструкции изображения большего пространственного разрешения и большего пространственного размера. Отличие между реконструктом и изображением большего разрешения образует изображение погрешностей, которое в случае гауссова размытия именуется «лапласовой пирамидой». Изображение погрешностей несет в себе как положительные так и отрицательные значения.

По материалам Burt & Adelson, 1983.

вой пирамиде невелик. Когда аппроксимация плотная — погрешности невелики и кодируются весьма и весьма эффективно.

В применении лапласовой пирамиды имеет место одна очевидная проблема: коэффициентов в пирамиде погрешностей оказывается больше, чем пикселей оригинала. Когда пирамида погрешностей строится из двумерных изображений, то, к примеру, мы квантуем через ряд и через колонку. Образуется последовательность из изображений погрешности в 1, 1/4, 1/16 размера оригинала. Следовательно в пирамиде в 1.33 раза больше коэффициентов, чем в оригинале (см. рис. 4.4.8).

Из-за избытка коэффициентов представление погрешностей именуют *переполнен-*



**Рис. 4.4.9** Гистограммы гауссовой и лапласовой пирамид. А — отдельные графики, представляющие собой яркостные гистограммы на каждом из уровней гауссовой пирамиды. Значения яркостей распределены в широком диапазоне значений, что затрудняет эффективное кодирование. В — коэффициенты лапласовой пирамиды, занимающие существенно меньший диапазон значений и кучкующиеся в основном возле нуля (что существенно повышает эффективность кодирования).

ным. Однако же нас интересует компрессия, и поэтому очевидно, что переполненное представление — шаг назад.

Барт и Адельсон показали однако (1983), что в контексте зрения имеет место некий факт, компенсирующий переполненность: дело в том, что подавляющее большинство коэффициентов преобразования представляет информацию о высших пространственных частотах, где острота зрения человека весьма низка. Следовательно, мы можем квантовать эти элементы очень жестко без видимой потери качества изображения. Отметим, что квантование — ключевой этап в компрессии, поэтому чем больше в нашем распоряжении коэффициентов, представляющих информацию которую можно жестко квантовать, — тем лучше.

Возможность жесткого квантования многих коэффициентов с минимальными перцепционными потерями (связанная с их пониженной вариабельностью) — это базис практической применимости лапласовой пирамиды при компрессиях изображений. Отметим, наконец, что расчет пирамиды может оказаться намного сложнее расчетов по DCT.

#### 4.4.1.11 QMF и ортогональные вейвлеты

У пирамидных представлений, построенных на гауссовом ядре свертки масса замечательных свойств, но есть и ряд недостатков. Углубившись в разбор издержек гауссовых и лапласовых пирамид, мы обнаружим, что при построении пирамиды имеет смысл пользоваться иным ядром свертки.

Первой малоприятной особенностью пирамид Гаусса и Лапласа является несовмес-

тимостью операций размытия и квантования. Предположим, мы начали анализ не с оригинала ( $g_0$ ), а с расчетного изображения —  $g'_0$ . С позиции создания пирамиды наш расчет должен быть эквивалентен оригинальному изображению. Сие логично, следовательно, редуцированное изображение, полученное из  $g'_0$  должно быть таким же, как и полученное из  $g_0$ . Условие это можно выразить уравнением:

$$g_1 = P_0 (2P_0^T) g_1 = P_0 g'_0,$$

из которого ясно, что квадратная матрица  $P_0 (2P_0^T)$  должна быть единичной, то есть колонки матрицы  $P_0$  должны быть *ортгональны* друг другу.<sup>1</sup> Однако ни гауссова, ни лапласова пирамиды не удовлетворяют этому требованию.

Вторым малоприятным свойством пирамид Гаусса и Лапласа является то, что представление переизбыточно (т.е. коэффициентов больше, чем пикселей в оригинале). Рост числа коэффициентов отслеживается по тому факту, что мы представляем изображение  $g_i$  с  $N_i$ -числом пикселей посредством редуцированного сигнала и сигнала погрешности, несущего больше, чем  $N_i$  коэффициентов. Например, мы представляем информацию на  $i$ -уровне пирамиды, используя редуцированное изображение  $g_{i+1}$  и изображение погрешностей  $e_i$ :

$$g'_i = (2P_i^T) g_{i+1} + e_i$$

В одномерном варианте изображение погрешностей ( $e_i$ ) несет в себе  $N_i$  коэффициентов преобразования. Редуцированный сигнал ( $g_{i+1}$ ) — вдвое меньше. Для создания эффективного представления мы должны представить  $g_i$ , используя  $N_i$  коэффициентов, а не  $1.5N_i$  как в гауссовой пирамиде.

Сигнал погрешности и интерполированный сигнал нацелены на кодирование разных компонент оригинального сигнала: интерполированный вектор —  $g'_i = (2P_i^T) g_{i+1}$  — кодирует пространственно-низкочастотную версию оригинала, а  $e_i$  — пространственно-высокочастотные компоненты, не вошедшие в низкочастотную версию. С целью повышения эффективности представления нам может потребоваться, чтобы оба показателя кодировали полностью различные типы информации о ретинальном изображении. Одним из вариантов толкования словосочетания «полностью различные» является требование ортогональности векторов, то есть:

$$0 = e_i^T g'_i$$

Ежели мы удовлетворим требованию ортогональности  $g'_i$  и  $e_i$ , мы можем получить высокую эффективность представления.

По определению нам известно, что интерполированное изображение  $g'_i$  — это взвешенная сумма колонок  $P_i^T$ . Если изображение погрешностей —  $e_i$  — ортогонально интерполированному изображению оригинала, то изображение погрешностей должно быть взвешенной суммой множества вектор-столбцов ортогональных колонкам  $P_i^T$ .

В одномерной гауссовой пирамиде число колонок в  $P_i^T$  равно  $N_i/2$ . Из основ линейной алгебры нам известно, что существуют  $(1/2)N_i$  векторов перпендикулярных ко-

<sup>1</sup> Два вектора  $a$  и  $b$  ортогональны друг другу, когда  $a^T b = 0$ .<sup>1111</sup>

лонкам  $\mathbf{P}_i^T$ . Следовательно, если  $\mathbf{g}'_i$  ортогональна  $\mathbf{e}_i$ , мы можем описать оба изображения, используя в совокупности только  $N_i$  коэффициентов, и представление не окажется переполненным.

Какие условия должны быть соблюдены, чтобы  $\mathbf{g}'_i$  и  $\mathbf{e}_i$  гарантированно оказались ортогональны? Подставив в последнее уравнение предыдущие, получаем:

$$0 = \mathbf{e}_i^T \mathbf{g}'_i = \left[ \mathbf{g}_i^T (2\mathbf{P}_i^T) \mathbf{P}_i \right] \left[ (2\mathbf{P}_i^T) \mathbf{P}_i \mathbf{g}_i - \mathbf{g}_i \right] = \left\{ \mathbf{g}_i^T (2\mathbf{P}_i^T) \left[ \mathbf{P}_i (2\mathbf{P}_i^T) \right] \mathbf{P}_i \mathbf{g}_i \right\} - \left[ \mathbf{g}_i^T (2\mathbf{P}_i^T) \mathbf{P}_i \mathbf{g}_i \right]$$

Если ряды  $\mathbf{P}_i$  — это ортогональное множество, то путем соответствующего шкалирования мы можем добиться того, что  $\mathbf{P}_i (2\mathbf{P}_i^T)$  окажется эквивалентно единичной матрице. В этом случае наше уравнение упрощается до:

$$\mathbf{e}_i^T \mathbf{g}'_i = \left[ \mathbf{g}_i^T (2\mathbf{P}_i^T) \mathbf{P}_i \mathbf{g}_i \right] - \left[ \mathbf{g}_i^T (2\mathbf{P}_i^T) \mathbf{P}_i \mathbf{g}_i \right] = 0,$$

гарантируя тем самым, что сигнал погрешностей и интерполированное изображение будут взаимно ортогональны. Итак, мы обнаруживаем, что ортогональность рядов пирамидной матрицы — весьма полезное ее свойство.

Подытожим сказанное. У базовой пирамидной операции имеется ряд весьма желательных свойств:

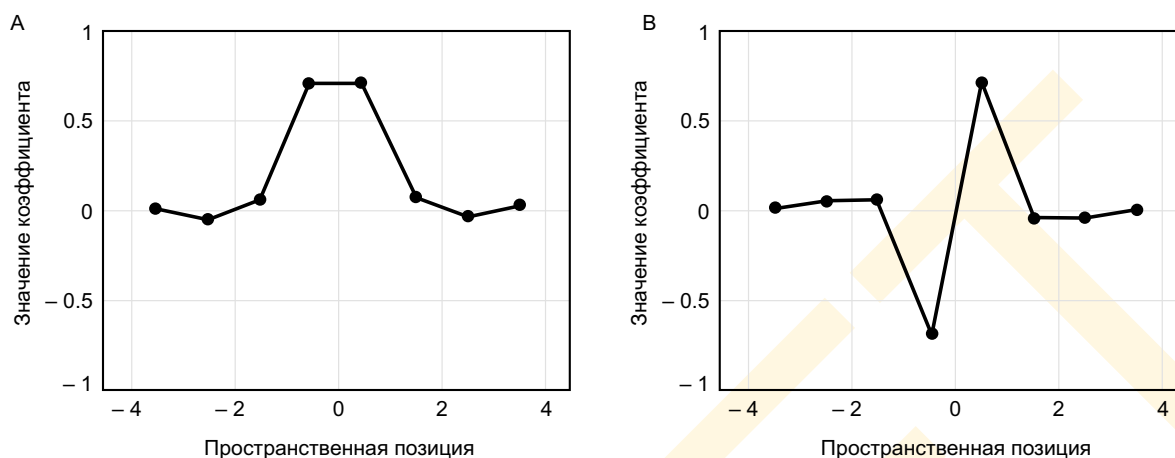
- ряды внутри каждого уровня пирамидной матрицы — суть сдвинутые копии друг друга, что упрощает вычисления при свертке;
- пирамидная операция представляет информацию разных пространственных разрешений, что очень сходно с происходящим в зрительной системе в контексте мультиразрешающих представлений;
- ряды пирамидной матрицы локализованы в пространстве как и рецептивные поля (но при этом не имеют резких границ подобно блокам JPEG-DCT-алгоритма).
- через пирамиду погрешностей, мы обрели новое свойство — ряды каждой пирамидной матрицы ортогональны.

Итак, мы полагаем матрицу  $\mathbf{P}_i$  удовлетворяющей нашим основным условиям. Теперь мы можем составить полный список требований, но нет гарантии того, что существуют функции, удовлетворяющие всему этому списку (и наибольшую сложность представит объединение ортогональности и пространственной локализации, к примеру: если мы рассматриваем операцию свертки отдельно, не существует сверток, которые были бы одновременно и ортогональны, и локализованы в пространстве).

#### 4.4.1.12 Квадратурно-зеркальные фильтры

Любопытно, что существует класс дискретных функций, именуемых *квадратурно-зеркальными фильтрами* — *КЗФ (QMF)*. Функции эти удовлетворяют всем требованиям из нашего списка (Esteban & Galand, 1977; Vetterli, 1988; Simoncelli & Adelson, 1990). Пара квадратурно-зеркальных фильтров делит входящий сигнал на два ортогональных компонента: первый фильтр определяет ядро свертки, используемое для размытия оригинала и получения редуцированного изображения; второй (ортогональный первому) может использоваться для вычисления эффективного представления сигнала погрешности. Следовательно, пара квадратурных зеркальных фильтров делит оригинальный сигнал на коэффициенты, определяющие два ортогональных показателя —  $\mathbf{g}'_i$  и  $\mathbf{e}_i$ . Каждый набор показателей несет только  $n/2$  показателей, так что новое представление — это эффективное пирамидное представление.





**Рис. 4.4.10** Пара квадратурно-зеркальных фильтров. При построении пирамид обе функции можно использовать в качестве ядер свертки. За сверткой с первым ядром (А) следует квантование, дающее набор редуцированных сигналов. Квантование, следующее за сверткой со вторым ядром (В) дает пирамиду погрешностей.

По материалам Simoncelli, 1988.

На рис. 4.4.10 показан пример работы пары квадратурно-зеркальных фильтров: функция на рис. 4.4.10 А — это ядро свертки, использованное для создания редуцированных изображений  $g_i$ ; функция на рис. 4.4.10 В — это ядро свертки, необходимое для вычисления коэффициентов преобразования непосредственно в пирамиде погрешностей. Когда фильтры строятся для непрерывного ряда значений (не дискретного), функцию ядра свертки называют *ортгональным вейвлетом* (Daubechies, 1992).

Открытие зеркальных фильтров и ортогональных вейвлетов оказалось несколько неожиданным. Известно, что не существует нетривиальных ортогональных ядер свертки, то есть не существует матрицы свертки —  $\mathbf{B}_i$  — удовлетворяющей критерию  $\mathbf{B}_i \mathbf{B}_i^T = \mathbf{I}$ . Тем более было неожиданно обнаружить, что существуют ядра свертки реализующие пирамидную операцию и одновременно удовлетворяющие требованию  $\mathbf{B}_i \mathbf{B}_i^T = \mathbf{I}$  — существуют.

У квадратурных зеркальных фильтров и ортогональных вейвлетов множество замечательных свойств, которые можно эффективно эксплуатировать при компрессиях, поскольку они ликвидируют проблему переполненности представления. Однако по сей день создание и апробация квадратурных фильтров остается областью активного исследования в сфере пирамидного кодирования изображений.

## 4.5 ДЕТЕКЦИЯ КРАЕВ И ОПОРНЫЕ СКЕТЧИ

В отличие от физиологов зрения, для которых открытие клеток — детекторов линий и краев оказалось явной неожиданностью, исследователи зрения компьютерного изначально знали, что отыскание краев и линий в поле зрения является критично важной целью оптического этапа зрительного акта.<sup>1</sup>

В контексте компьютерного зрения первое, что следует сделать с реальным изображением сцены или даже страницей текста — это найти в нем линии и края. Эти эле-

менты, которые, явно заметны при рассматривании контонного монохромного изображения (скажем чашки на рис. 2.4.2) полностью сокрыты, когда изображение рассматривается как матрица яркостных величин, зафиксированных видеокамерой (рис. 2.4.3 В). Даже в первых компьютерных программах Робертса (1965) первый этап заключался в детекции линий и краев.

В количественном подходе к оптическому этапу пространственного зрения много лет доминировала точка зрения Дэвида Марра (1982), предложившего идею опорных скетчей. Несмотря на то, что Марр был первым, кто осознал важность оптического этапа зрительного восприятия, ученый озвучил общую теорию вопроса на удивление цельно и четко.

Фактически Марр показал, что существуют два опорных скетча: *простой* и *полный*. Маррово специфическое толкование природы оптических представлений в логике опорных скетчей не получило широкого распространения и мы рассмотрим альтернативную точку зрения в конце данного раздела. И все-таки маррова идея играет важную роль в развитии количественных теорий низкоуровневого пространственного зрения и представляет как историческую, так и теоретическую ценность.

**Простой опорный скетч.** Марров простой опорный скетч завершает первый шаг перехода от полностью аналогового яркостного изображения к символьному представлению элементов этого изображения. Марр показал, что зрение должно моментально и полностью «символизироваться», и чтобы решить эту задачу он предложил четыре качественно отличных друг от друга типа элементов изображения в опорном скетче: *края, линии (полосы), пятна и оконцовки*. Сии примитивы он окрестил «символьными», поскольку они формируют дискретные категории оптического представления.

У каждого символьного примитива (края, полосы, пятна или оконцовки) имеются свои позиция, размер, ориентация и контраст.

О движении и о хроматике как о параметрах символов Марр в явном виде ничего не говорил, но мы полагаем, что направление движения, скорость этого движения и хроматический контраст — тоже можно занести в список аналоговых показателей, определяющих каждый символичный элемент изображения.

**Полный опорный скетч** — это более организованная версия простого скетча. Включает в себя результаты операций, таких как соединение коротких сегментов линий и краев в длинные, группировку сходных друг с другом элементов и возможную характеристику текстуры участков поля зрения. Данному вопросу посвящено много меньшее число исследований, но мы все-таки обсудим некоторые из них в гл. 6.

#### 4.5.1 Детекция краев

Итак, пространственные примитивы марровой первой стадии оптической обработки — это края, линии, точки и оконцовки. Больше всего работы выпадает на долю зрительных механизмов, отвечающих за детекцию яркостных краев. Мы рассмотрим ее детально как пример извлечения низкоуровневых компонентов из ретинального изображения. Как мы увидим далее, механизмы эти необычайно сложны.

Когда объекты трехмерной сцены отражают упавший на них свет на двухмерную

<sup>1</sup> Исследователи компьютерного зрения строят электронные зрительные системы, что называется «с нуля» и поэтому хорошо знают, как должен работать тот или иной механизм. Тогда как у физиологов принципиально иная задача — вывести из экспериментальных данных то, как именно работает существующая система.

поверхность, скажем сетчатку или матрицу сенсоров фотокамеры, то изменение в фотометрической яркости по всей протяженности некоей равномерно окрашенной поверхности — градуально, то есть является результатом легкого изменения освещенности. Изменения, возникающие при переходе от одной поверхности к другой всегда намного более резкие и, чаще всего, образуют вещественные края. Эти края играют огромную роль в восприятии, потому что указывают либо на изменение в отражательной способности поверхности (например при смене материала), изменении количества света, падающего на эту поверхность (например, при образовании тени), либо об изменении ориентации поверхности относительно источника света (поскольку большинство поверхностей под разными углами к источнику отражают разное количество света). По этим причинам (и ряду других), яркостные края признаются всеми исследователями как важнейший оптический компонент ретинального изображения.

#### 4.5.1.1 Край-операторы и свертка

Первые попытки автоматического распознавания яркостных краев, т.е. с помощью компьютерных алгоритмов, были по большей части бессистемными: исследователи применяли в основном эмпирические методы, пытаясь подобрать к различным изображениям наиболее подходящий к ним т.н. *край-оператор*.

⇒ *Край-оператор* — это количественная схема интеграции яркостных величин по группе соседствующих элементов изображения в единичное значение, более или менее достоверно представляющее то, что в определенном участке изображения имеется яркостный край.

На рис. 4.5.1 показаны простейшие операторы вертикальных и горизонтальных краев.

Говоря количественным языком: детекция краев реализуется зрительной системой путем *свертки*, проделываемой с изображением край-оператором. Как мы увидим далее «свертка» — это лишь математический термин, означающий то, что именно ретинотопически организованный пласт корковых клеток проделывает с ретинальным изображением (точнее, с информацией о ретинальном изображении), если рецептивные поля этих клеток имеют структуру край-операторов. Каждая клетка «вычисляет» сумму возбуждающих сигналов на своем входе и добавляет к ней сумму сигналов тормозящих (которые, разумеется, имеют отрицательный знак), после чего дает финальный выходной сигнал. Выходной сигнал всего клеточного пласта как отклик на полное ретинальное изображение — это и есть собственно свертка (т.е. край-операция), изображения с рецептивными полями этих клеток.

Чтобы глубже понять механизм работы край-операторов, рассмотрим рис. 4.5.2, на котором показано простое бесшумовое изображение креста в яркостной (А) и числовой (В) формах, а также его свертка с вертикальным (С) и горизонтальным (D) край-операторами. Чтобы рассчитать свертку числового изображения (В) с оператором вертикального края (С), начнем с левого верхнего пиксела этого изображения и приложим весовую схему вертикального край-оператора как к самому левому верхнему пикселу, так и к пикселу сразу справа от него. Чтобы выполнить сие, нужно просто умножить значение левого верхнего пиксела (10) на минус 1 ( $= -10$ ) и значение пиксела справа от него (10) на +1 ( $= +10$ ), а затем сложить оба этих значения ( $-10 + 10 = 0$ ). Результат окажется равным нулю, поскольку яркостное значение у обоих пикселей одно и то же. Полученное значение введем в левый верхний пиксел на выходе (Е).

## КРАЙ-ОПЕРАТОРЫ ПЕРВОГО ПОРЯДКА

+ 1	- 1
-----	-----

+ 1
- 1

А. Вертикальный

В. Горизонтальный

## КРАЙ-ОПЕРАТОРЫ ВТОРОГО ПОРЯДКА

- 1	+ 2	- 1
-----	-----	-----

- 1
+ 2
- 1

- 1	- 1	- 1
- 1	+ 8	- 1
- 1	- 1	- 1

С. Вертикальный

D. Горизонтальный

Е. Разнонаправленный

**Рис. 4.5.1** Пять простейших край-операторов как взвешивающие схемы интеграции соседних пикселей. Край-операторы первого порядка (вертикальные и горизонтальные) попросту «вычисляют» яркостную разницу между соседними пикселями. Операторы второго порядка эффективно «вычисляют» отличия между отличиями. Особо отметим, что во всех случаях весовые суммы равны нулю.

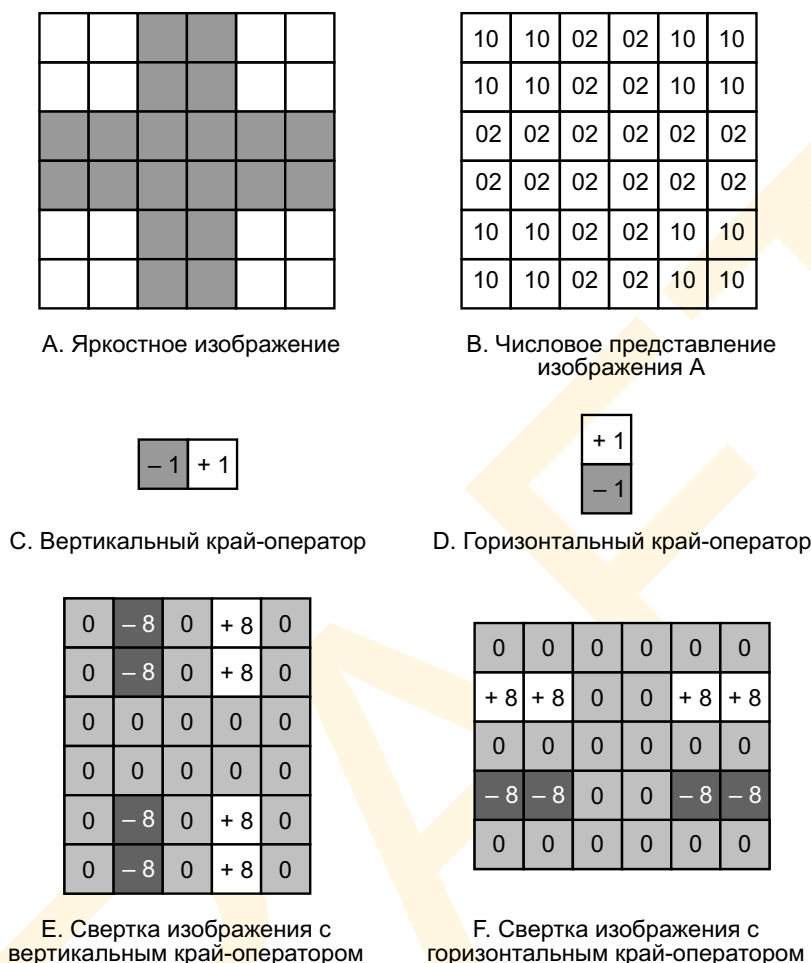
Отметим, что всякий раз, когда значения соседних пикселей оказываются одинаковы или близки друг к другу, выходной сигнал оператора будет равен нулю (или около нуля). Следовательно, оператор не откликается на участки изображения равномерные по фотометрической яркости, слегка откликается на плавные ее изменения и выдает наибольший сигнал на резкие изменения фотометрической яркости, возникающие на краях. Все последующие шаги в свертке изображения с край-оператором аналогичны описанным выше и выполняются последовательно пиксел за пикселем.

Теперь сдвинемся дальше на один пиксел и увидим, что его яркость (10) существенно выше таковой (2) у третьего пиксела в ряду, и поэтому выходной сигнал край-оператора окажется резко отрицательным ( $-10 + 2 = -8$ ), указывающим на наличие в данной позиции *светло-темного края*. И так далее по всем оставшимся пикселям в ряду.

Когда оператор дойдет до черной вертикальной полосы, на выходе образуется серия нулей, а затем на границе темного участка и светлого — резко положительное значение. То есть, положительный сигнал на выходе край-оператора указывает на наличие *темно-светлого края*.

Итак, выполнив данную процедуру по каждой паре пикселей в каждом ряду изображения, мы осуществим свертку этого изображения с край-оператором. В числовом представлении результат показан на рис. 4.5.2 Е, но его можно представить и в яркостном виде, присвоив 0 нейтрально-серым пикселям, положительные значения пикселям ярче серого, а отрицательные тем, что темнее. Отметим, что вертикальный край-оператор реагирует только на вертикальные края изображения и ни на какие другие.

Процедура расчета свертки изображения с край-оператором — это только один из этапов «вычислений», параллельно выполняемых теми нейронами, чьи рецептивные

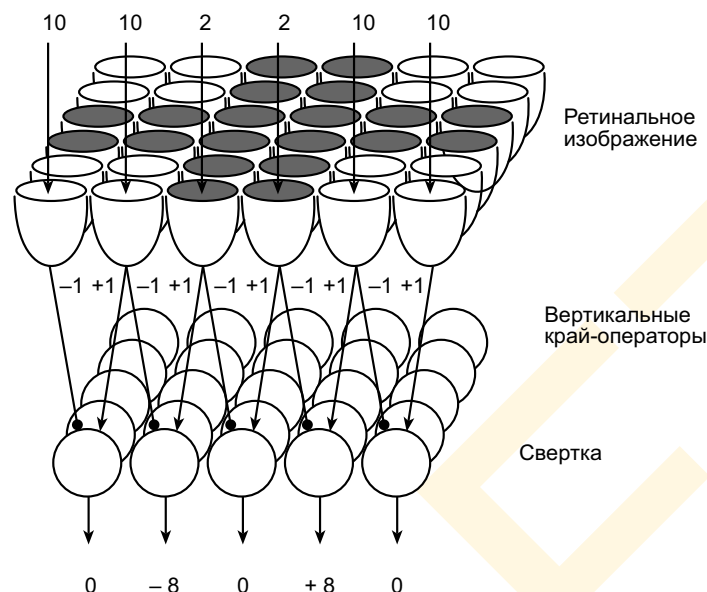


**Рис. 4.5.2** Свертка изображения с край-операторами. Монохромное яркое изображение (A) можно свернуть с край-операторами различных видов (например, C и D) умножением их весовых коэффициентов на значения яркостей (B). Темные пиксели в E и F представляют резко отрицательные величины, светлые — положительные.

поля соответствуют весовой схеме край-оператора. На рис. 4.5.3 показана схема монтажа вертикального край-оператора в нейронную сеть: взвешивание от входного слоя к выходному указывает на тормозящую связь слева ( $-1$ ) и возбуждающую справа ( $+1$ ), которые на выходе суммируются. Таким образом, выходной сигнал данной сети — это свертка изображения (представленного участками активации входного слоя) с оператором, представленным схемой весовых связей.

Отметим, что данную свертку нельзя назвать собственно последовательной (как можно было бы предположить на основании вышеописанной шаговой процедуры), но скорее серийной. Зрительная система выполняет свертку намного быстрее, выполняя все операции параллельно, используя при этом связи между слоями нейронов для «вычислений» сигналов свертки по каждому участку изображения. Везде, где в изображении имеются вертикальные края, значение свертки будет либо резко положительным, либо резко отрицательным. То есть, если край-оператор выдал мощный сигнал — это свидетельство наличия края.





**Рис. 4.5.3** Нейронная сеть, организующая свертку ретинального оптического изображения с вертикальным край-оператором. Простейший вертикальный край-оператор первого порядка (см. рис. 4.5.2 С) реализован в нейронной сети за счет возбуждающих и тормозящих связей фоторецепторов с клетками, представляющими результат свертки ретинального изображения с этим оператором. Выходной сигнал сети тот же, что и на рис. 4.5.2 Е.

С тем же успехом мы можем выполнить свертку этого же изображения с оператором горизонтальных краев (рис. 4.5.2 D). Результат будет аналогичным (рис. 4.5.2 F).

Пример работы край-операторов в реальных изображениях дан на рис. 2.2.2 В, где показаны результаты первой робертсовской программы поиска вертикальных, горизонтальных, право- и леводиагональных краев (1965).

Специализированные край-операторы именуются *дифференциальными операторами первого порядка*, поскольку они ловят простейшие отличия между соседними пикселями и вычисляют исключительно наличие яркостной функции вдоль определенного пространственного направления. Наклон этой функции — это лишь отличие между соседними значениями.

Более сложный дифференциальный оператор — *оператор второго порядка*, можно получить, вычислив разность между соседними операторами первого порядка (рис. 4.5.1 С, 4.5.1 D, 4.5.1 Е). Работа операторов второго порядка отлична от работы операторов первого порядка тем, что края диагностируются не экстремальными величинами, а нулями с пограничными им экстремальными величинами. На рис. 4.5.1 Е показан край-оператор второго порядка, одновременно детектирующий вертикальные, горизонтальные, леводиагональные и праводиагональные края. Сие возможно, поскольку позитивно-негативно-позитивный поперечный профиль, представленный во всех четырех ориентациях, проходит через центр край-оператора.

Итак, теперь нам становится понятной роль количественной теории в краевой детекции и как эта теория соотносится с функцией различных структур в зрительной системе.

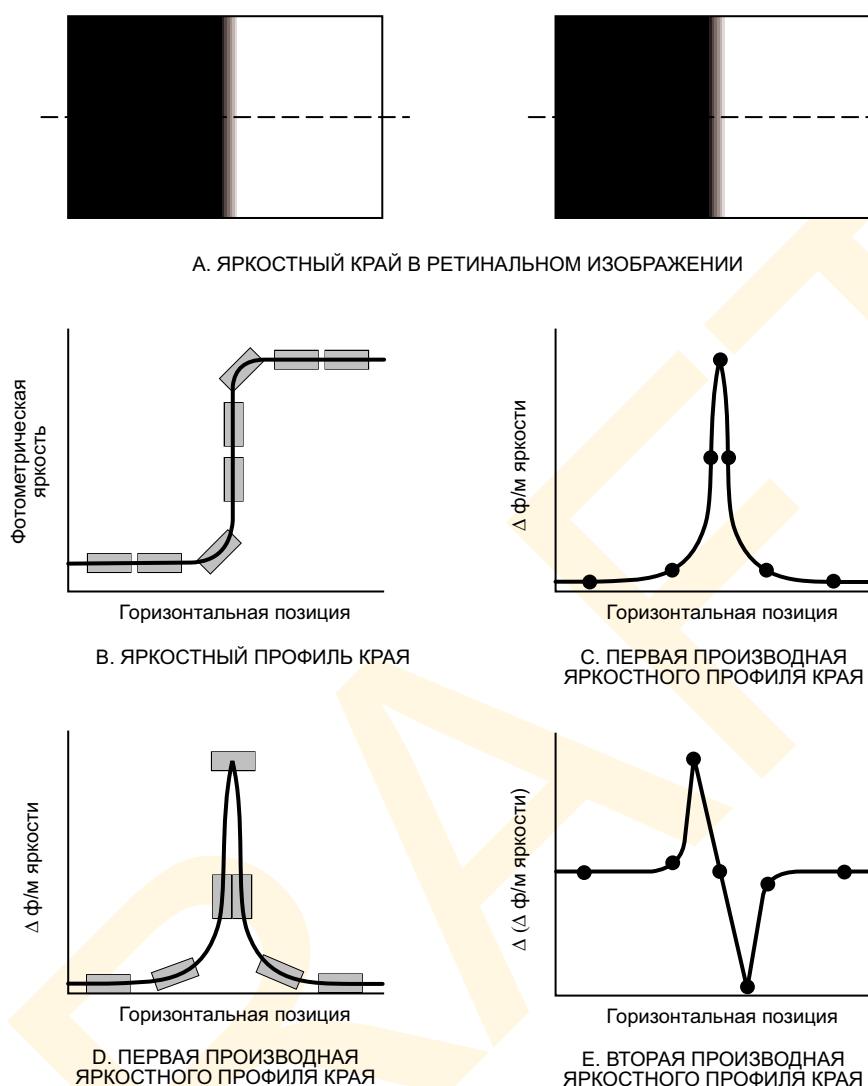
### 4.5.1.2 Нуль-детектор Марра — Хилдрет

В конце 70-х годов прошлого века Дэвид Марр и Элен Хилдрет (1980) в рамках марровой концепции трехуровневой информационной обработки сформулировали свою собственную теорию краевой детекции. Сразу оговоримся, что каких-либо серьезных альтернатив современным алгоритмам краевой детекции, предложенным еще в 1986 г. Джоном Канни (Беркли) и его последователями (например, Deriche, 1987; Spasek, 1985) не существует, но алгоритм Марра — Хилдрет представляет, в первую очередь, исторический интерес как одна из самых первых схем краевой детекции, в которой предпринята серьезная попытка согласовать количественный подход к проблеме с физиологией зрительной системы млекопитающих. Поэтому рассмотрим ее достаточно подробно.

Количественное описание Марр и Хилдрет (1980) начали с характеристики идеализированного одномерного яркостного края, который должна была отыскивать их система (рис. 4.5.4 А). Исследователи рассудили так, что такой край можно локализовать, найдя позицию максимального наклона его пространственно-яркостной функции (фотометрическая яркость как функция от пространственного положения). Таким образом яркостный профиль края можно изобразить в виде кривой в координатной системе, у которой по оси абсцисс отложена пространственная позиция, а по оси ординат — фотометрическая яркость (рис. 4.5.4 В).

Дифференциальные операторы первого порядка «вычисляют» наклон пространственно-яркостных функций. На рис. 4.5.4 В наклон яркостной функции представлен семью серыми прямоугольниками. На рис. 4.5.4 С дана т.н. *первая производная* этой функции, характеризующая ее наклон, т.е. *изменение* световой интенсивности как функцию от пространственной позиции. Таким образом, местоположение яркостного края можно определить детекцией максимума (наивысшего значения) первой производной, который соответствует максимуму оператора первого порядка. Однако Марр и Хилдрет раскритиковали эту схему, поскольку операторы первого порядка должны были работать во всех возможных ориентациях, что крайне неэкономично. При этом исследователи высказали предположение, что должен существовать более экономичный алгоритм поиска краев, распознающий момент перехода через нулевое значение *второй производной* яркостной функции. Вторая производная яркостной функции — это наклон первой производной. Чтобы получить вторую производную, нужно попросту повторить давеча проделанную процедуру, но используя при этом первую производную в качестве базовой функции. Проще говоря, вторая производная рассчитывает «наклон наклона» исходной яркостной функции (рис. 4.5.4 D и 4.3.4 E). Отметим, что во второй производной положение яркостного края соответствует нулевому значению, находящемуся между наивысшим положительным и наивысшим отрицательным значениями. Локальные операторы, рассчитывающие вторую производную — это дифференциальные операторы второго порядка, упомянутые ранее (рис. 4.5.1 С, 4.5.1 В и 4.5.1 E).

Пересечение нуля второй производной (рис. 4.5.4 E) — формально эквивалентно максимуму первой производной, но обладающей потенциальным количественным преимуществом, обсуждавшимся выше: поскольку операторы второго порядка симметричны относительно своей срединной точки (см. рис. 4.5.1 С и 4.5.1 D), вторую производную двумерного изображения можно рассчитать по всем направлениям одновременно с помощью единичного двумерного оператора, показанного на рис. 4.5.1 E и 4.5.5 А. Этот оператор представляет собой строго одновременную комбинацию опе-



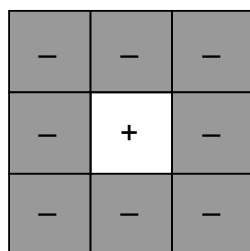
**Рис. 4.5.4** Первая и вторая производные яркостных краев. Яркостный край (А) можно представить в виде графика (часто именуемого «яркостным профилем края»), на котором фотометрическая яркость дана как функция от пространственной позиции (В). Первая производная (или, как еще говорят — наклон) этой функции имеет пик в центре яркостного края. Вторая производная (Е), которая представляет уже наклон первой производной, пересекает нулевую отметку при проходе из области положительных значений в область отрицательных.

раторов второго порядка вертикальной, горизонтальной, леводиagonalной и праводиagonalной ориентаций.

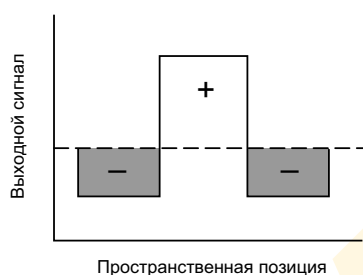
В недискретном пространстве тот же самый двухмерный оператор второго порядка показан на рис. 4.5.5 В, а профиль его кросс-нулевого отклика там же, но ниже. Тот факт, что вторую производную по всем направлениям можно сжать в единый оператор свидетельствует о том, что данная схема количественно эффективна. Структура рецептивного поля у такого оператора окажется очень близка к «сомбреро» и строго та же, что у рецептивных полей ретиальных ганглиозных клеток и клеток колленчатого тела (см. раздел 4.1.1).

На рис. 4.5.6 показана работа описанного алгоритма на конкретном примере —

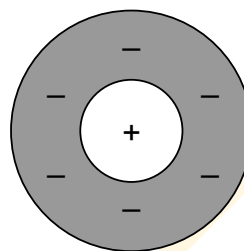
## А. ДИСКРЕТНЫЙ КРАЙ-ОПЕРАТОР



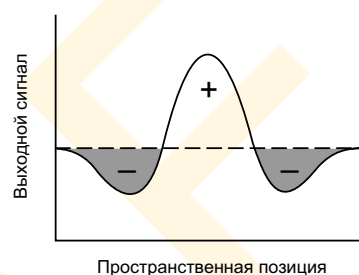
Рецептивное поле



## В. НЕПРЕРЫВНЫЙ КРАЙ-ОПЕРАТОР



Рецептивное поле



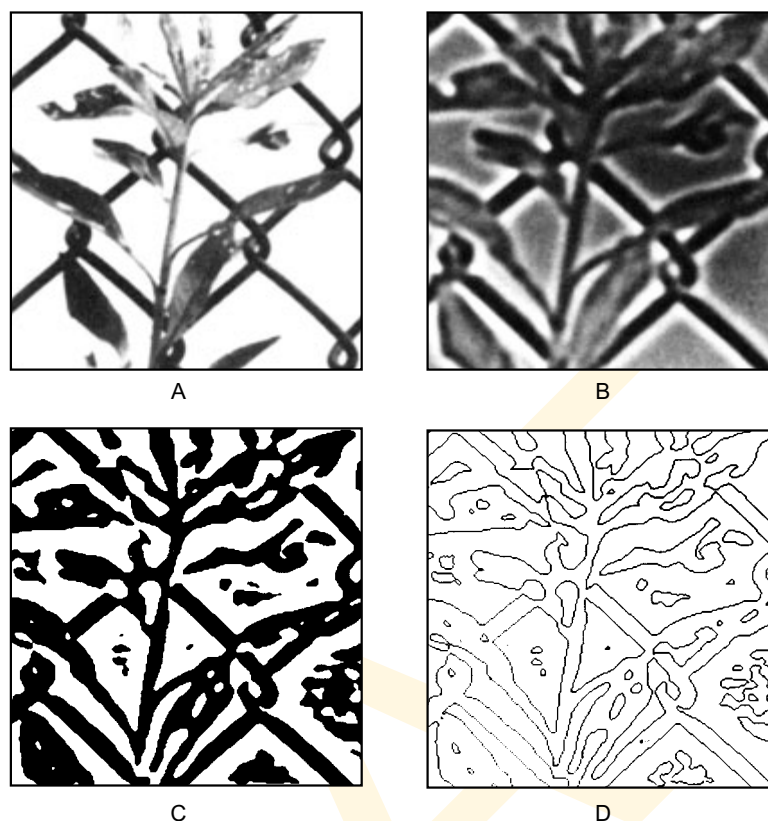
## ПРОФИЛЬ ОТКЛИКА

**Рис. 4.5.5** Дискретная (А) и непрерывная (В) версии разнонаправленного край-оператора второго порядка. Верхнее изображение — рецептивные поля; нижнее — профили их отклика по горизонтальному направлению.

изображении растения на фоне проволочного забора; на рис. 4.5.6 А дано серое контонное изображение; на рис. 4.5.6 В — результат свертки изображения А с оператором второго порядка (нулевые значения соответствуют средне-серому, положительные значения — светло-серым, отрицательные — темно-серым); на рис. 4.5.6 С — то же в бинарной форме, где переход через ноль представлен черно-белыми краями; на рис. 4.5.6 D даны фактические проходы нуля, идентифицированные кросс-нулевым алгоритмом Марра — Хилдрет.

Отметим, что сложность количественных теорий краевой детекции в том, что яркие края в ретинальном изображении возникают на разных пространственных размерах: одни являются следствием медленных изменений в световой интенсивности больших участков сцены, другие — быстрых, но на мелких ее участках. Марр и Хилдрет работали над вопросом, применяя край-операторы второго порядка трех размеров: большие для детекции краев низкого разрешения (грубые края); средние — для детекции краев среднего разрешения и мелкие — для детекции краев высокого разрешения.

Чтобы оценить, как выглядят края трех уровней пространственного разрешения, посмотрим на дискретную цифровую аппроксимацию портрета Авраама Линкольна — рис. 4.5.7 (Harmon & Julesz, 1973). Проходы нуля, зафиксированные край-операторами Марра — Хилдрет на уровнях высокого разрешения (часть В), среднего (часть С) и низкого (часть D) демонстрируют то, что края, позволяющие наблюдателю различать черты лица Линкольна, локализируются, в первую очередь, на краях низкого



**Рис. 4.5.6** Поиск краев путем нуль-детекции. Серое яркостное изображение (А) вначале свертывается с разнонаправленным край-оператором второго порядка (например, показанном на рис. 4.3.5 В), результатом чего является изображение В, в котором положительные значения — это белые участки, отрицательные — черные, а нули — серые. В изображении С положительные значения — белые участки; отрицательные — черные. Проход через ноль в В показан в изображении D, что и является демонстрацией работы алгоритма краевой детекции.

По материалам Marr, 1982.

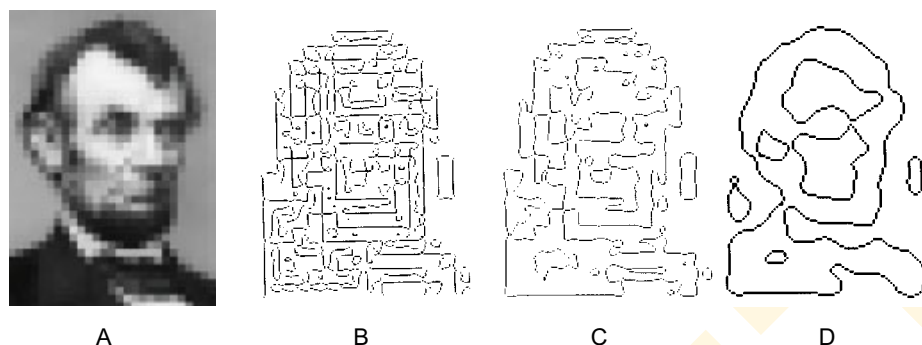
разрешения. Мы можем отделить эти края, размыв оригинальное изображение и отнеся его на такое расстояние, при котором края высокого разрешения станут незаметны (или же просто сняв очки, ежели мы близоруки).

В данном, скорее искусственном, случае некоторые края на низких и высоких уровнях разрешения оказываются одними и теми же. В изображениях естественных сцен, однако, большинство краев, выявляемых на низких уровнях разрешения, обнаруживаются обычно и на высоких в той же пространственной позиции. Более того, обнаружение краев на разных уровнях пространственного разрешения повышает вероятность того, что представлен реальный край, а не шум.

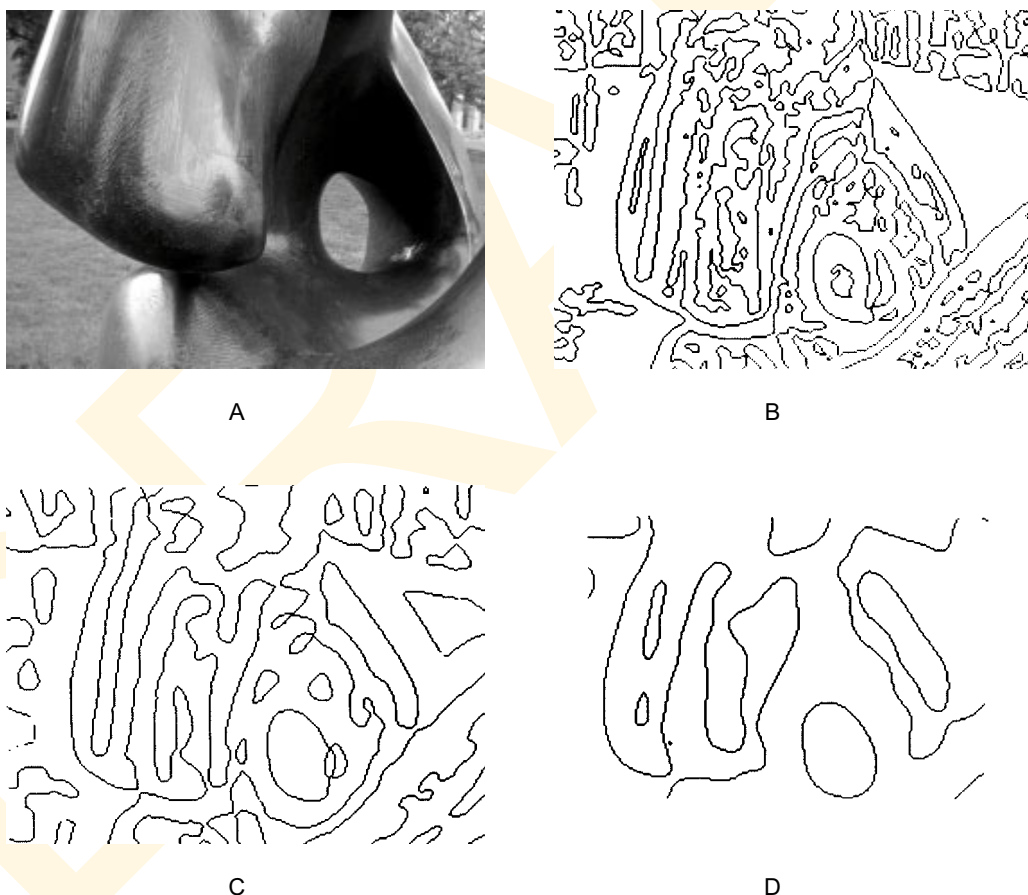
На рис. 4.5.8 показан проход нулевой отметки, зафиксированный по изображению А на трех разных пространственных разрешениях. Отметим, что на высоких пространственных частотах краев много больше, но самые визуально значимые представлены на всех.

Наконец, стоит отметить, что Марр и Хилдрет вывели свою концепцию не на ровном месте: каждый ее компонент (свертка с оператором краев и линий, нуль-детекция дифференциальными операторами второго порядка, а также работа на разных уров-





**Рис. 4.5.7** Краевая детекция на разных пространственных частотах. Картина краев в дискретном портрете Авраама Линкольна (А) анализировалась на высоких (В), средних (С) и низких (В) пространственных частотах по алгоритму Марра — Хилдрет. Отметим, что контуры лица Линкольна оказались наиболее узнаваемыми в картине краев, отработанной оператором на низких частотах.



**Рис. 4.5.8** Нуль-детекция на различных пространственных частотах: края в яркостном изображении фрагмента скульптуры Генри Мура (А) отыскивались по алгоритму Марра — Хилдрет на высоких (В), средних (С) и низких (D) пространственных частотах. Столь разные результаты должны каким-то образом интегрироваться зрительной системой в единое представление.

По материалам Marr & Hildreth, 1980.

нях пространственного разрешения) был описан в предшествующих исследованиях машинного зрения (Duda & Hart, 1973; Horn, 1971; Rosenfeld, Thurston & Lee, 1972). Однако заслуга Марра и Хилдрет в том, что они собрали эти компоненты в единый, добротный и работоспособный алгоритм.

#### 4.5.1.3 Нейронное воплощение

Марр и Хилдрет выполнили еще одну не менее важную работу: они тщательнейшим образом и путем сопоставления с функцией известных к тому времени клеток нервной системы млекопитающих, исследовали то, существует ли у алгоритма физиологическая реализация.

Легко догадаться, что миссию свертки изображения с край-оператором второго порядка исследователи возложили на ганглиозные клетки сетчатки и нейроны коленчатого тела, рецептивные поля которых организованы по типу «сомбреро». Результат нейронной обработки (предположительно на выходе клеток коленчатого тела) в этом случае окажется аналогичен тому, что показано на рис. 4.5.6 В. В данном представлении имеет место нуль-детекция дифференциальным оператором второго порядка, но еще в неявном виде. Следовательно, детекция эта должна продолжиться за счет какого-то иного процесса.

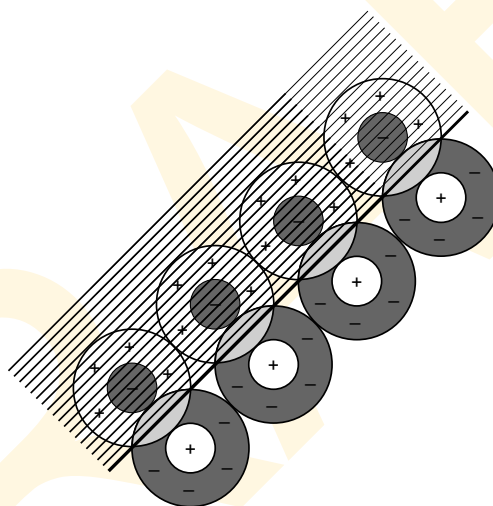
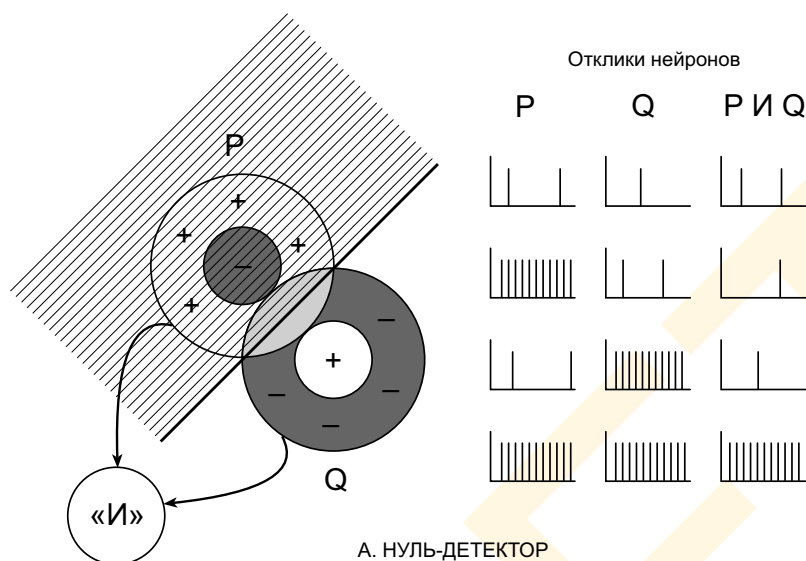
Чтобы понять, как нуль-детекция реализуется нейронными процессами, вспомним, что на выходе оператора второго порядка она имеет место между наивысшим положительным и наивысшим отрицательными значениями. Марр и Хилдрет предположили, что нуль-детекция осуществляется по положительно-отрицательным, либо отрицательно-положительным изменениям на выходе рецептивных полей, организованных, разумеется, по схеме центр/окружение. Сие возможно силами оператора, выполняющего логическую операцию «И» в отношении соседствующих on- и off-центр-клеток коленчатого тела (рис. 4.5.9). При этом краевой детектор выдает явный и устойчивый сигнал только в том случае, когда активны гангикулярные клетки (клетки коленчатого тела) обоих типов (рис. 4.5.9 А).

Краевые детекторы, тонко настроенные на специфическую пространственную ориентацию, могут строиться объединением сигналов от множества пространственно-упорядоченных нуль-детекторов (рис. 4.5.9 В).

Сия точка зрения весьма близка к догадке Хьюбела и Визеля о том, каким образом простые клетки стриарной коры, получая и суммируя сигналы от клеток коленчатого тела, исполняют функцию край-операторов.

Итак, Марр и Хилдрет пришли к тому, что простые стриарные клетки — детекторы краев, фактически являются нуль-детекторами. То есть, их гипотеза гласит, что эти клетки действительно являются детекторами краев, но при этом не такими как изначально предполагали Хьюбел и Визель: эти клетки скорее ловят работу краевого алгоритма, «вычисляющего» проход через ноль выходного сигнала дифференциального оператора второго порядка.

Несколько позднее, Канни (1986) придумал «более эффективный» компьютерный алгоритм краевой детекции, использующий сложный анализ, оптимизирующий детекцию краев в определенных условиях. Канни отошел от алгоритма Марра — Хилдрет, применяя дифференциальные операторы не второго, а первого порядка. Путем ряда сложных математических построений Канни разработал оптимальный детектор двухмерных «степ-краев», представлявших собой весьма плотную аппроксимацию идеализированного хьюбело-визелевского детектора (рис. 4.5.10). Великолепная работа этого оператора продемонстрирована рис. 4.5.11.



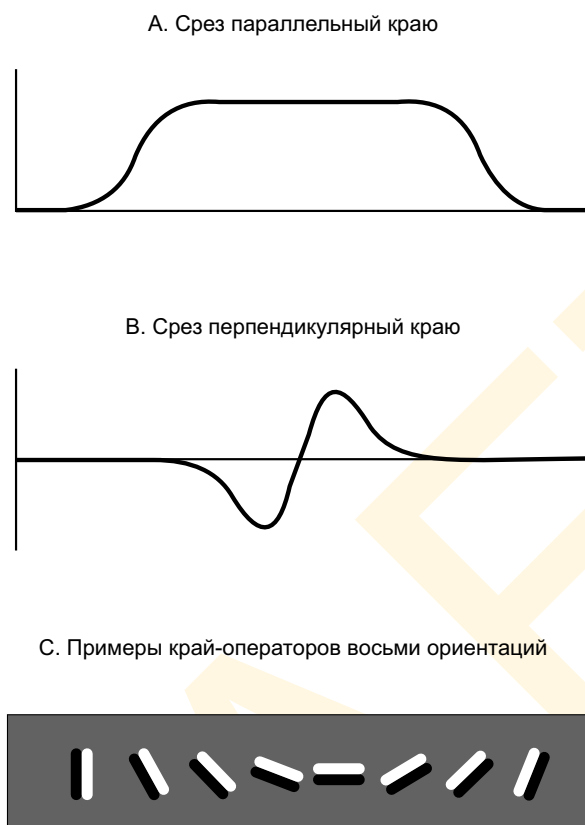
**Рис. 4.5.9** Нейронная модель алгоритма Марра — Хилдрет. Ноль-детектор (А) можно построить, выполнив логическое «И» в отношении выходных сигналов on- и off-центр рецептивных полей. Серия таких полей, выстроенная по определенному направлению, оказывается высокочувствительной к краям этого направления.

По материалам Marr & Hildreth, 1980.

Отметим, что разрабатывались и другие весьма эффективные алгоритмы краевой детекции, в частности **Дерич** (1987) и **Спэйсеком** (1985), использовавшими методы максимизации, сходные с методами Канни.

#### 4.5.1.4 Интеграция шкал

На первый взгляд ноль-детекция по алгоритму Марра — Хилдрет самодостаточна и, собственно, и есть факт детекции края. Однако же очевидно, что зрительной системе приходится обобщать, объединять информацию от краев разных размеров или, как



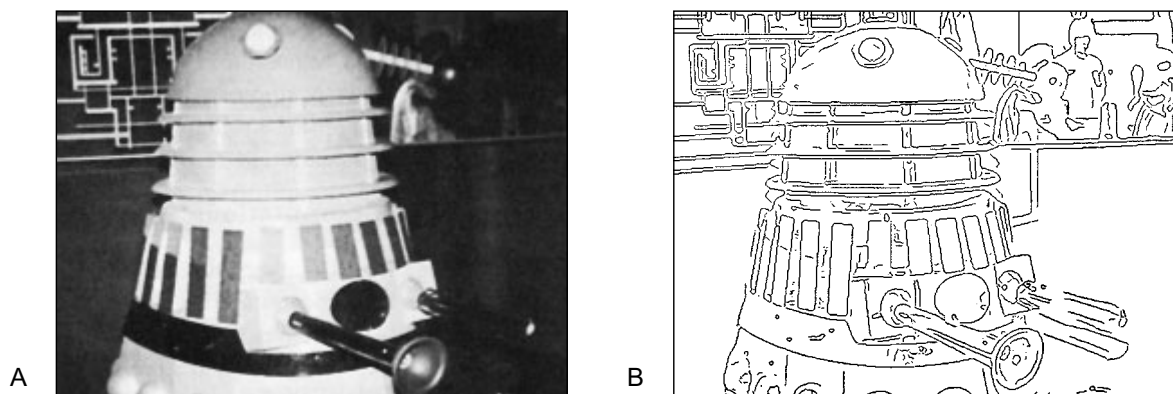
**Рис. 4.5.10** Край-операторы, функционирующие по алгоритму Канни. Край-оператор Канни определен профилем, параллельным краю (А) и перпендикулярным ему (В).

По материалам Sanny, 1986.

говорят, шкал. Сие весьма проблематичный момент, но Марр и Хилдрет никогда не упоминали о нем. Алгоритм Канни также требует интеграции по разным размерам край-операторов и чтобы оценить какова мера интеграции, требуемая для точной краевой детекции, необходим ряд процессов.

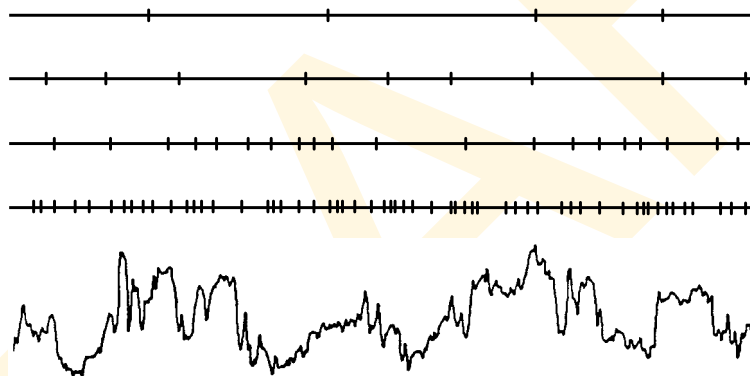
Вопрос интеграции шкал в своем исходном виде звучит так: зрительной системе необходимо определить, как согласуются разные края в разных шкалах — какие из них — это один и тот же край, а какие нет? Проблема проиллюстрирована рис. 4.5.12: пилообразная линия внизу рисунка представляет собой яркостный профиль некоего ретинального изображения, а деления на четырех линиях выше отмечают позиции краев, обнаруженных на шкалах четырех разных пространственных размеров. Вопрос интеграции шкал идет из попытки сложить деления вместе: какой край с каким делением идет? Наиболее очевидная тактика — это положить в основу согласования местоположение и ориентацию краев. Проблема в том, однако, что местоположение и даже ориентация данного края могут слегка разниться у разных шкал, несколько снижая возможность ясного решения.

Вопрос интеграции краевой информации по шкалам разного размера в какой-то мере решили с помощью т.н. *алгоритма шкального пространства*, придуманного американским исследователем искусственного интеллекта Эндрю Виткин (1983). Виткин заменил гипотезу Марра — Хилдрет о трех дискретных шкалах идеей непрерывного шкального пространства, простирающегося от очень мелких край-операторов, до



**Рис. 4.5.11** Яркостное изображение (А) в сравнении с картиной его краев (В), детектированных по канни-алгоритму.

По материалам Sanny, 1986.



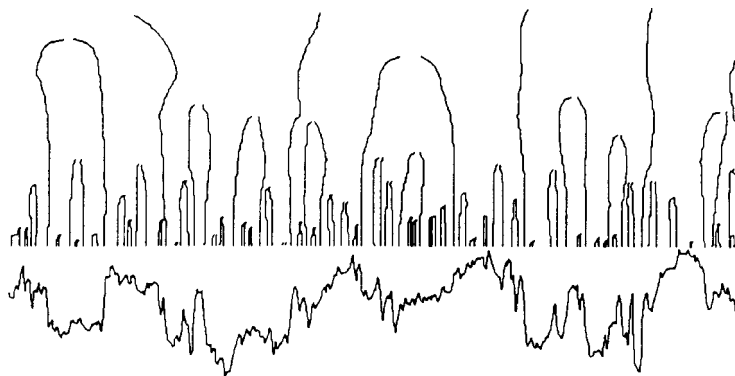
**Рис. 4.5.12** Проблема интеграции краевой информации на разных уровнях пространственного разрешения. Одномерный волновой сигнал проанализирован по алгоритму нуля-детекции на четырех пространственных частотах, в результате чего зафиксировано множество краев на разных уровнях (штрихи на горизонтальных линиях вверх). Суть проблемы интеграции в ответе на вопрос: как края на разных пространственных частотах соотносятся друг с другом?

По материалам Witkin, 1983.

очень крупных. Ежели существует виртуальный континуум шкальных размеров, то краевую информацию можно пропустить через этот континуум, дабы определить, какой край, какой толщины (рис. 4.5.13). Отметим, что ежели идти от верха (низкие пространственные частоты) к низу (высокие частоты), то на разных уровнях появляются новые края и что некоторые из них оказываются одиночными, а другие — парными, то есть, начинающимися вместе и заканчивающимися далеко друг от друга.

Одна из наиболее характерных особенностей виткин-алгоритма в том, что он предполагает непрерывность в представлении размеров. Сие значит, что если зрительная система человека использует именно такой алгоритм, то ей потребуется больший массив данных, чем предполагалось ранее. Однако идея непрерывного шкального пространства вполне совместима с экспериментальными данными, демонстрирующими, что рецептивные поля корковых клеток несут в себе весьма плотное представление пространственных частот (De Valois & De Valois, 1988). Возможно, одна из функций





**Рис. 4.5.13** Виткин-алгоритм шкального пространства, позволяющий решить проблему интеграции шкал. Виткин предложил идею непрерывного континуума шкал, автоматически связывающего края с разными пространственными шкалами.

По материалам Witkin, 1983.

рецептивных полей корковых клеток — это исполнение в каком-то виде виткин-алгоритма детекции и интеграции краев.

## 4.5.2 Опорные скетчи

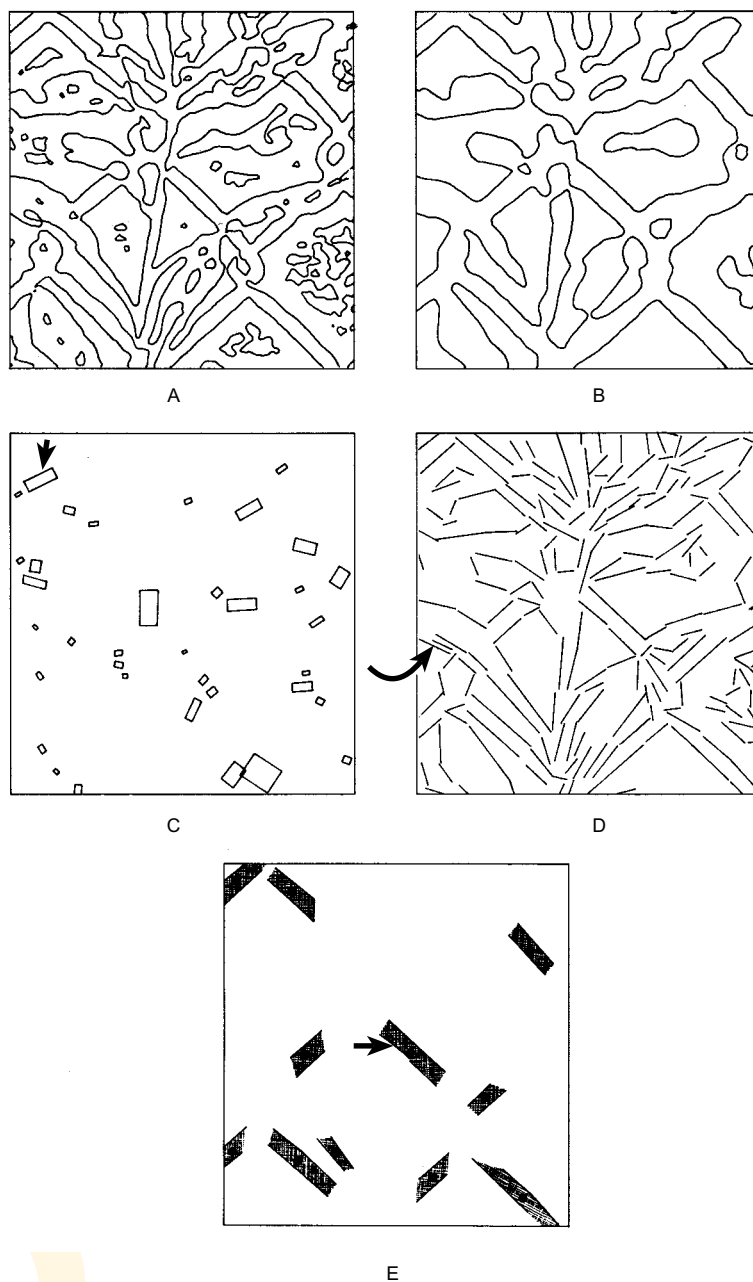
### 4.5.2.1 Простой опорный скетч

Напомним, что в концепции простого опорного скетча роль примитивов Марра отвел не только краям, но также линиям, пятнам и оконцовкам. Рис. 4.5.14 проиллюстрирует нам эти компоненты у решетки, взятой с рис. 4.5.6 А: на рис. 4.5.14 А и 4.5.14 В изображена нуль-детекция на шкалах двух разных размеров, а на рис. 4.5.14 С, 4.5.14 Д и 4.5.14 Е показаны пятна, края и штрихи, автоматически обнаруженные программами краевой детекции.

Штрихи Марром определены как короткие отрезки линий, чьи оконцовки лежат за пределами рецептивного поля; оконцовки определены как окончания краев и штрихов; пятна — как короткие штрихи с оконцовками по обеим сторонам. У перечисленных компонент, разумеется, в зрительной коре существуют детекторы, аналогичные детекторам краев: штрихи обнаруживаются простыми стрийными клетками, а оконцовки — сложными стопорными (см. раздел 4.1.3.2). Действительно, великие открытия Хьюбела и Визеля оказали серьезное влияние на решение Марра включить штрихи и оконцовки в опорный скетч. Последний можно легко толковать как количественный продукт нейронной обработки в V1.

Идея включения штриховых и краевых оконцовок в опорный скетч по началу выглядит неожиданной, поскольку элементы эти в отличие от краев, штрихов и пятен «невещественны», но лишь «завершают» их. Однако же существует ряд необычайно эффектных перцепционных феноменов, целиком построенных именно на «работе» оконцовок.

На рис. 4.5.15 даны два случая, в которых расположение оконцовок линий имеет явный перцепционный смысл. Отчасти смысл этот состоит в том, что оконцовки завершают нейронный сигнал детекции края или линии, что, вероятно, создает иллюзию контура (рис. 4.5.15 А и 4.3.15 В).



**Рис. 4.5.14** Марров простой опорный скетч. Яркостное изображение садовой решетки (рис. 4.5.6 А) проанализировано по алгоритму нуль-детекции на высоких (А) и низких (В) пространственных частотах, а результаты интегрированы в символьную карту направленных краев (D). Дальнейший анализ выявил наличие «пятен» (С) и «штрихов» (Е) в разных позициях.

По материалам Marr & Hildreth, 1980.

Иллюзорный контур — это зрительное ощущение края, возникающее в том участке поля зрения, где физического яркостного края нет.

Любопытно, что ежели некая фигура определена таким строго ориентированными краями как на рис. 4.5.15 В, то зачастую она выглядит светлее окружающего ее фона.

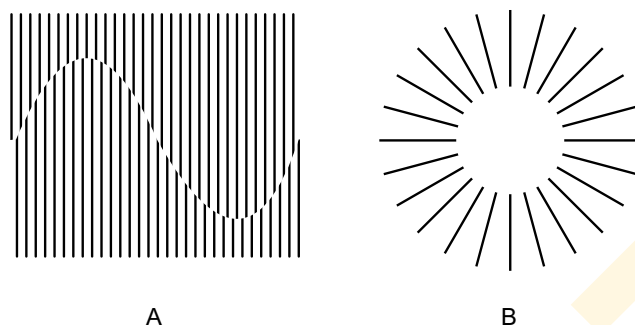


Рис. 4.5.15 Примеры иллюзорных контуров, определяемых оконцовками линий. Марр полагал, что такие контуры образуются за счет объединения оконцовок в опорном скетче.

К разговору об иллюзорных контурах мы вернемся в гл. 6, когда поведем разговор о перцепционной организации.

#### 4.5.2.2 Знаки и токены

Итак, простой опорный скетч — это первое символьное представление в марровой теории обработки визуальной информации, и выявление простого скетча зрительной системой — это момент, в котором определенные составляющие непрерывного ретинального изображения преобразуются в дискретное множество краев, штрихов, пятен и оконцовок. Сии четыре категории базовых компонентов ретинального изображения в символьном представлении именуются *знаками*, а каждый отдельный край, штрих, пятно или оконцовка в поле зрения наблюдателя — *токеном* соответствующего знака.

Каждый токен характеризуем рядом параметров, отличающих один возможный токен от другого: специфической позицией штрихов ( $x$ - и  $y$ -координаты), ориентацией, направлением контраста (светло-темный или темно-светлый), длиной и толщиной.

Отметим, что  $x$ - и  $y$ -позиции представлены на рис. 4.5.14 ретинотопически. Многие дальнейшие вычисления, которые нужно проделать в отношении опорного скетча в рамках локального участка, легко выполнимы поиском соседних участков на ретинотопической карте, в которой участки эти по определению стыкуются друг с другом. Поскольку бо́льшая часть обработки, необходимой для перехода на следующий уровень преставления — т.е. на полный опорный скетч — требует именно этого вида локального поиска, Марр оставил 2D-структуру ретинального изображения в опорном скетче неизменной. Также, когда на более поздних стадиях возникает необходимость в интеграции результатов, ретинотопический принцип позволяет легко находить необходимые элементы в соответствующих участках. Сие важно в контексте разговора об интеграции, который пойдет в 11-й главе.

#### 4.5.2.3 Полный опорный скетч

Марров полный опорный скетч образуется из простого опорного скетча в результате работы организационных процессов. Гипотетически эти процессы связывают короткие отрезки линий в длинные контуры, объединяют элементы изображения в группы и дробят его на участки, отличные друг от друга по текстуре. Мы отложим разговор об этих процессах до 6-й главы, в которой поведем детальный разговор о вопросах перцепционной организации.

## 4.6 АЛЬТЕРНАТИВНЫЕ КОЛИЧЕСТВЕННЫЕ ТЕОРИИ

Детекция таких оптических токенов сцены как края, штрихи и пятна — это не единственный вариант толкования функции клеток Хьюбела — Визеля. Мы знаем, что локальный анализ пространственных частот — это очень серьезный конкурент марровой концепции опорных скетчей, но опять же — не единственный. Одной из альтернатив является идея о том, что роль этих клеток критична в анализе текстурной информации, а также существует и такая точка зрения, что клетки Хьюбела — Визеля участвуют в восстановлении из яркостной информации фактической кривизны поверхностей. Мы рассмотрим эти гипотезы очень коротко.

### 4.6.1 Текстурный анализ

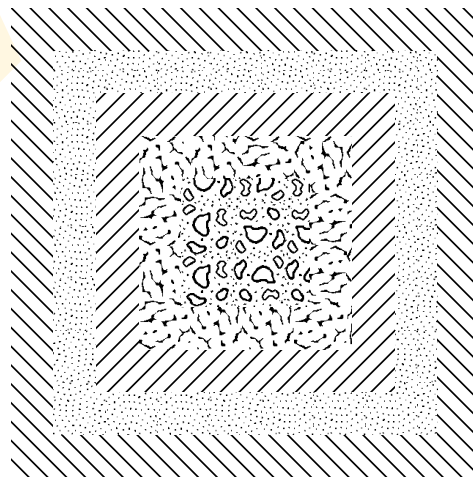
Итак, одной из альтернативных количественных теорий, толкующих возможную функцию клеток Хьюбела — Визеля, является концепция *текстурного анализа*, предложенная Джитендрой Малик и Петро Перона (Беркли).

⇒ Текстурный анализ (анализ текстур) — это процесс, посредством которого зрительная система выделяет участки поля зрения, отличные друг от друга по статистическим параметрам пространственной структуры.

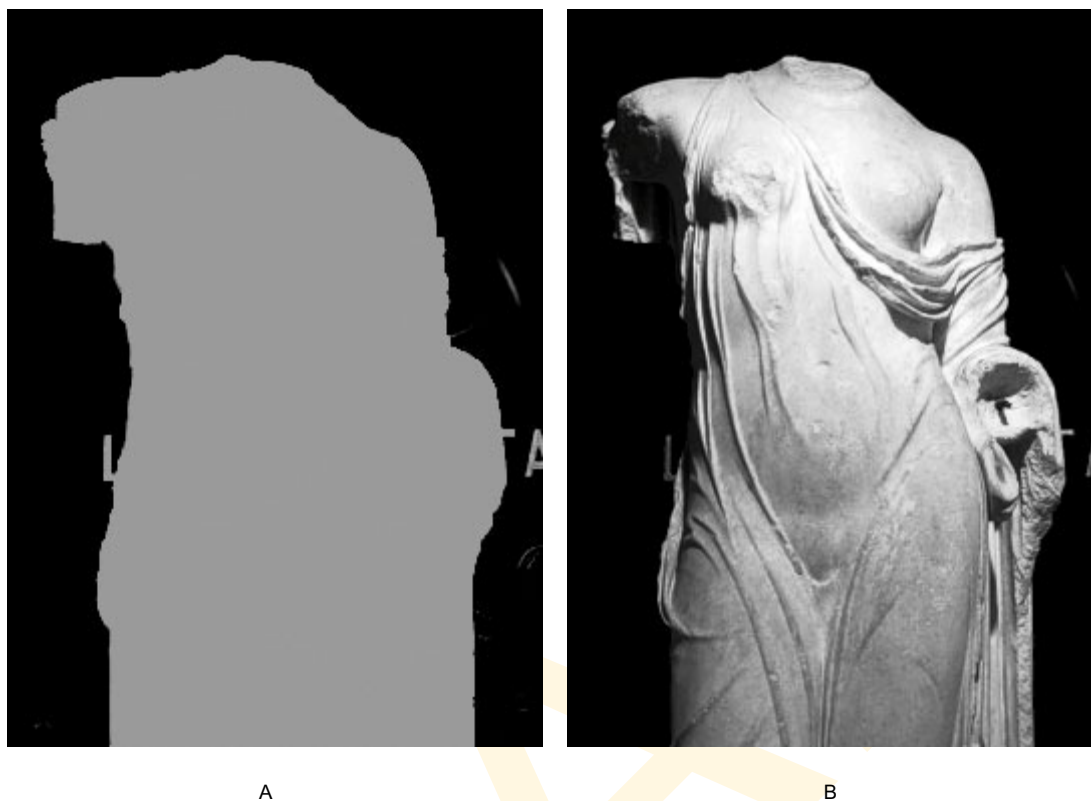
Исследователи привели весьма убедительный аргумент в пользу того, что клетки V1 запускают начальный этап сегрегации поверхностей по текстурному критерию.

На рис. 4.6.1 дан пример того, что визуальное отличие между квадратами зиждется на отличии в их текстуре, но не в окраске или общей фотометрической яркости. Детально мы опишем теорию Малик — Пероны в гл. 6, посвященной детекции текстур, здесь же отметим, что суть их догадки в следующем: многие из результатов текстурной перцепции в какой-то мере объяснимы идеей свертки ретинального изображения с рецептивными полями V1. При этом отличия в ориентации и пространственной частоте этих полей критичны в анализе Малик — Пероны.

На первый взгляд идея исследователей не сильно разнится с тем тезисом, что соответствующие клетки отвечают за обнаружение яркостных краев, и мы будем полагать, что края текстурных элементов играют весьма и весьма важную роль в текстурном анализе. Критично важным отличием от всего, что мы рассматривали ранее является то, что алгоритм Малик — Пероны не обнаруживает яркостные края в явном виде. Скорее, он просто использует выходной сигнал мультиориентационных мультишкальных клеток как входной сигнал процесса, обнаруживающего крупногабаритные неравномерности в текстуре, но не в



**Рис. 4.6.1** Текстурный анализ. У разных участков этого изображения примерно один и тот же средний коэффициент отражения и, следовательно, примерно одна и та же средняя фотометрическая яркость. Концепция текстурного различения, предложенная Маликом и Перона, основана на данных исследования выходных сигналов рецептивных полей V1.



**Рис. 4.6.2** Анализ яркостных модуляций (шейдинга). Жесткое отличие между перцептом яркостно-демодулированной версии изображения (А) и модулированной (В) наглядно демонстрирует нам то, что информацию об объеме предмета зрение извлекает из картины светов и теней. На снимке справа — статуя Афродиты.

По материалам Koenderink, Kappers, Pollik & Kawato, 1997)

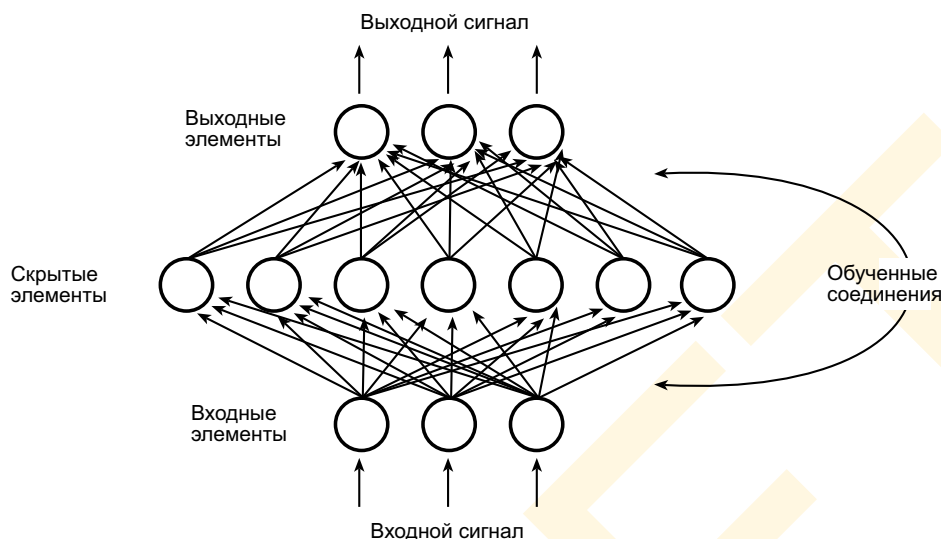
фотометрической яркости. Ежели оно так, и ежели это те же самые клетки, что задействованы в краевой детекции, то можно предположить, что они исполняют несколько различных функций. К этой мысли мы еще вернемся.

#### 4.6.2 Реконструкция рельефа из яркостных модуляций

Весьма оригинальную количественную гипотезу функции клеток V1 предложили на базе исследований нейронных сетей американские нейрофизиологи Сидни Леки и Терренс Сейновский (1988b, 1990). Они изучили вопрос т.н. *шейдинга* в ретинальном изображении, то есть вопрос о том, как информация о кривизне поверхностей объектов сцены извлекается зрением из модуляций фотометрической яркости в ретинальном изображении (рис. 4.6.2). Действительно, яркостно-демодулированная 2D-версия рис. 4.6.2 А выглядит плоско и нераспознаваемо (за исключением того, что это некое пятно неправильной формы). Когда к изображению добавлены яркостные модуляции (рис. 4.6.2 В) объект сразу обретает трехмерность и моментально распознается.

Принцип, с которым Леки и Сейновский подошли к решению вопроса о восприятии кривизны поверхностей, состоял в исследовании того, как трехслойная нейронная сеть (см. рис. 4.6.3) обучается «дешифровке» кривизны поверхностей из яркостных





**Рис. 4.6.3** Архитектура трехслойной прямопоточной нейронной сети, обученной шейдинг-анализу (анализу яркостных модуляций). Сеть построена из 122 входных элементов, имитирующих клетки ЛКТ с рецептивными полями класса центр/окружение (61 on-центр, 61 off-центр), 27 скрытых элементов и 24 выходных. Обучение выполнено с помощью автоматической процедуры, управлявшей весами связей между элементами.

модуляций в ретинальном изображении. Исследователи использовали обучающую процедуру, известную как алгоритм обратного распространения, автоматически меняющую веса связей так, чтобы средний, «скрытый» слой был способен к мэппингу входного представления в представление класса «форма из яркостных модуляций». Леки и Сейновский положили на входной слой сигнал от «сомбреро»-полей клеток колленчатого тела (см. рис. 4.1.2), а на выходной — характеристики, необходимые для кодирования кривизны отдельных участков поверхностей. Затем применили алгоритм обратного распространения, управляющий весами связей со срединным слоем и назад от него, что нужно для мэппинга входного представления (сигнал клеток колленчатого тела) в выходное представление кривизны поверхностей.

Однако мы не будем описывать алгоритм обратного распространения, обучающий нейронную сеть (см. Приложение В), равно как не станем говорить о тонкостях мэппинга рецептивных полей в скрытый срединный слой. Интересно другое — то, что после обучения сети, рецептивные поля множества «нейронов» скрытого слоя оказались поразительно схожими с таковыми у клеток V1, то есть — вытянутыми и с двумя (или более) долями возбуждения и торможения. И хотя оценка подобия рецептивных полей искусственной сети полям реальных клеток весьма и весьма субъективна, результат все же заставляет задуматься.

Сходу неясно, почему имитаты клеток обладают именно такой количественной структурой, но понятно по крайней мере то, что они имеют прямое отношение к модуляциям яркости в ретинальном изображении и, следовательно, модуляциям яркости у разного рода искривленных поверхностей. Возможно, наиболее важным является то, что сей анализ подтолкнул исследователей к мысли о возможной *полифункциональности* рецептивных полей V1.

Формальный математический анализ структуры, выведенной из яркостных модуляций, был выполнен американским исследователем Алексом Пентлэндом (1989).

Ученый показал, что в строго заданных условиях для восстановления примерной формы поверхности из ее проекционного изображения достаточен набор габоровых фильтров различных пространственных частот и ориентаций в обеих фазах и во множестве различных позиций. Сие оказался в точности тот самый набор рецептивных полей, что был принят многими исследователями при их толковании функции нейронов V1 (к примеру, Dangman, 1980; De Valois & De Valois, 1980). Следовательно, можно предположить, что выходной сигнал соответствующих клеток используется зрительной системой в качестве базиса дальнейших процессов, восстанавливающих форму поверхностей из яркостных модуляций как минимум по мелким участкам ретинального изображения (на что указывают результаты Леки и Сейновского). Так ли оно на самом деле — остается неясным по сей день.

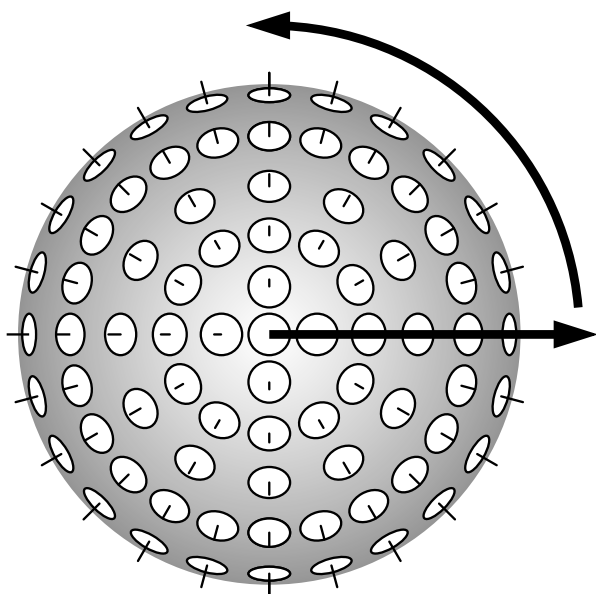


Такие естественные компоненты двумерных ретинальных изображений, как контуры и текстурные элементы (см. гл. 4), играют важнейшую роль в зрительном акте, потому что несут особую основную визуальную информацию об окружающем ее мире. Но при этом, чтобы зрение смогло сделать смысловой скачок от ретинальной проекции к реальности, компоненты эти следует интерпретировать в трехмерных показателях. Скачок сей необходим, поскольку особь нимало не озабочена контурами соприкосновения между ретинальными участками разной яркости, хроматики и текстуры — ее интересует все то же самое, но в реальной сцене. Следовательно, интерпретация ретинального изображения в логике зримых поверхностей трехмерного пространства — один из наиглавнейших шагов в попытке решить проблему инверсии (раздел 1.2.3). Процесс этот сложен и до сих пор толком не понят, но исследователи зрения прилагают фантастические усилия, пытаясь разобраться в его механизмах.

Мы помним, что две из трех пространственных размерностей внешнего мира в явном виде представлены в двумерном ретинальном изображении. Размерности эти можно толковать как совокупно описывающие *направление* от наблюдателя к поверхности. Но третья размерность — расстояние от наблюдателя до поверхности (*глубина*) — исчезает при образовании проекции трехмерной сцены на двумерную поверхность сетчатки. Коль скоро размерность эта утрачена, ее никогда не удастся восстановить с абсолютной точностью по причинам, которые уже обсуждались ранее (раздел 1.2). Однако же тот факт, что мы с вами вполне адекватно воспринимаем трехмерный окружающий мир и свободно ориентируемся в нем, свидетельствует о том, что информация об удаленных от нас поверхностях все же каким-то образом восстанавливается из двумерного ретинального изображения, и восстанавливается в большинстве случаев естественным путем. Более того — с очень высокой точностью. Как же сие возможно?

В контексте восприятия пространственной организации поверхностей сцены существуют две близких друг другу задачи: первая — это определение *глубины пространства*, то есть удаления поверхности от наблюдателя в трехмерном окружении; вторая — это восприятие ориентации поверхностей — их *поворота* и *наклона* относительно линии прямой видимости (зрительная ось). Хотя слова «наклон» и «поворот» часто звучат как синонимы, фактически они относятся к двум разным параметрам ориентации поверхности в глубину.

*Поворот (slant)* означает величину угла между зрительной осью наблюдателя и нормалью к интересующей поверхности (условный перпендикуляр к этой поверхности в данной точке). Чем больше этот угол — тем больше поворот. В *мерных кнопках* рис. 5.0.1 сплющивание диска соответствует повороту — чем больше это сплющивание, тем больше поворот плоскости от фронтальной (ее поворот равен нулю). Поворот также соответствует длине проекции единицы нормали к поверхности (нормаль к по-



**Рис. 5.0.1** Поворот и наклон локальных участков шара. Трехмерную ориентацию любого локального участка поверхности можно охарактеризовать двумя параметрами: поворотом и наклоном. Поворот — это угол между зрительной осью на данный участок поверхности и нормалью к этой поверхности (направление, перпендикулярное поверхности), на что указывает степень вытянутости овалов и длина отрезков. Наклон — это направление градиента глубины поверхности, на что указывает ориентация овалов и отрезков.

верхности единичной длины) на фронтальную плоскость. Чем длиннее проекция, тем больше поворот.

*Наклон (tilt)* относится к направлению градиента глубины относительно фронтальной плоскости. На рис. 5.0.1 наклон соответствует ориентации эллипса во фронтальной плоскости и направлению нормали к поверхности, спроектированной на фронтальную плоскость.

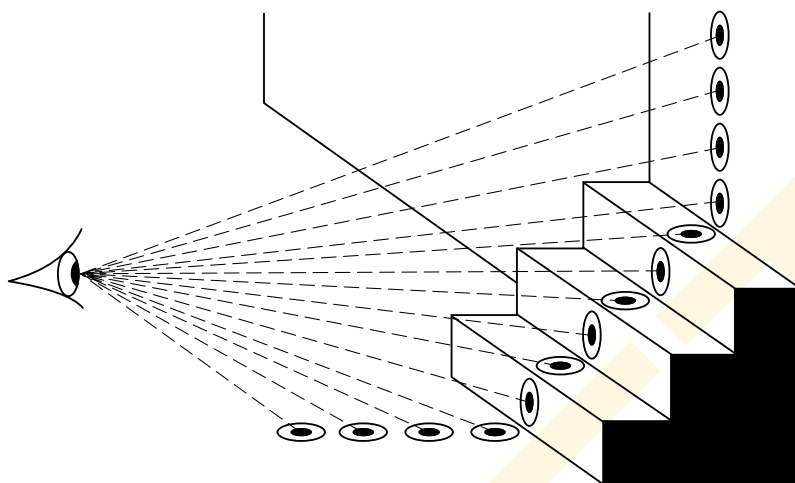
Понятно, что удаление участков поверхности и ориентация этой поверхности реконструируются зрением синхронно, бо сугубо взаимозависимы: трехмерная ориентация плоскости определяет то насколько разные ее участки удалены от наблюдателя, а расстояние между этими участками аналогично определяет трехмерную ориентацию. Мы можем сказать, что зрительной системе приходится «вычислять» ориентацию поверхности, отразившей свет в данном конкретном направлении (см. рис. 5.02), и расстояние до нее (глубину). Концепция визуального восприятия поверхностей по удаленности фрагментов и ориентации (или, как еще говорят, ориентации

на расстоянии) впервые была сформулирована Д. Д. Гибсоном в 1950 г. Он окрестил этот процесс — *восприятием планировки поверхностей*. В дальнейшем идею реконструкции ориентации и удаления в логике количественных показателей уточнили и углубили Дэвид Марр и Кейт Нишихара (1978). Как мы уже говорили в конце второй главы, Марр назвал сие фасетное представление *2.5-скетчем*. Однако наш сегодняшний взгляд на проблему несколько шире: безотносительно к тому, каково название и каково определение, представление пространственно-ориентированных поверхностей абсолютно необходимо зрению и требует понимания многих высокоуровневых перцепционных феноменов (Nakayama, He & Shimojo, 1995).

Почему первый шаг к зрительному восприятию пространственной структуры окружающего мира — это извлечение информации о характере видимых поверхностей? Вероятно, наиболее точный ответ таков: **видимые поверхности — это материальные сущности, фактически взаимодействующие со светом и определяющие тем самым то, каков характер оптического потока, воздействующего на сетчатку** (см. раздел 1.2).

Поскольку большинство поверхностей непрозрачны, мы не располагаем оптической информацией о подлежащих им предметах или расположенных позади их (если поверхности эти, разумеется, принадлежат трехмерным объектам). Сие значит, что





**Рис. 5.0.2** Ориентация на расстоянии. Расположение видимых поверхностей в сцене можно представить в виде локальных мелких фрагментов, каждый из которых имеет какое-то направление от глаза наблюдателя. Каждый фрагмент можно охарактеризовать его расстоянием от глаза наблюдателя и ориентацией относительно зрительной оси на него.

именно поверхности, в первую очередь, служат источником всей непосредственной зрительной информации о трехмерной структуре окружающего нас мира.

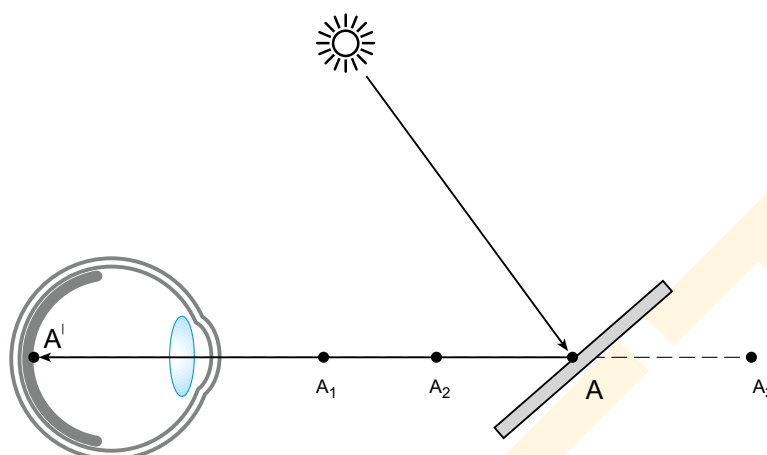
После того, как трехмерная структура видимых поверхностей воспринята зрением, особь может делать следующий вывод, например о том, каков характер частично или полностью закрытых поверхностей. И т.д.

## 5.1 ПРОБЛЕМАТИКА

В перцепционной реконструкции расстояния до поверхности первейшая сложность в том, что восприятие глубины пространства из двухмерного ретинального изображения — это т.н. *недоопределенная задача* (см. раздел 1.2.4). Причина же недоопределенности в том, что оптические процессы — т.е. процесс отражения света от поверхностей и формирования изображения на сетчатке — проектируют свет из трехмерного мира на двухмерную заднюю поверхность глаза. И, как мы показали в разделе 1.2, такие проекции инвертируемы в бесконечное число вариантов. То есть, мы говорим, что восприятие глубины пространства — это пример *парадигмы логической недоопределенности в восприятии*. Рис. 5.1.1 проиллюстрирует основную дилемму: предположим, свет отразился от точки  $A$ , попал в глаз и простимулирован рецептор в точке  $A'$ . Вопрос в том, откуда наблюдателю известно, что свет прибыл именно из точки  $A$ , а не из  $A_1$ ,  $A_2$ ,  $A_3$  и т.д.? В самом деле, он мог попасть в глаз с какой угодно позиции на линии  $AA'$ . Иными словами, поскольку все точки на этой линии дают проекцию в одну и ту же точку  $A'$ , обратная проекция может попасть в любую точку на отрезке  $AA'$ .

### 5.1.1 Эвристические допущения

Итак, формальная логика подталкивает нас к выводу о том, что веридикальное зрительное восприятие глубины пространства невозможно в принципе. Но как сей, казалось бы, неизбежный вывод согласуется с тем фактом, что мы ежедневно и ежеми-



**Рис. 5.1.1** Проблема недоопределенности глубины. Фотоны, отраженные от поверхности в точке  $A$ , ударяются о сетчатку глаза в точке  $A'$ , но они могут попасть туда и из точки  $A_1$ , и из точки  $A_2$ , или из любой другой точки на линии  $AA'$ .

нужно с высочайшей точностью оцениваем пространственную глубину пространства, в котором обитаем? Ответ таков: невозможным является лишь *абсолютно точное* восприятие при всех возможных обстоятельствах. Поскольку восприятие глубины пространства зрительной системой человека порой дает сбой -- логического противоречия нет. Как мы увидим далее, существует множество условий, в которых наше зрение буквально одурачивается неадекватным восприятием глубины. Собственно, примеров масса в повседневной жизни (хотя мы чаще всего не думаем о них, как о перцепционных ошибках): плоская фотография весьма убедительно передает объемные соотношения в изображенной сцене; кинофильмы -- еще эффектнее; а уж трехмерное кино -- фантастически реалистично. Однако во всех трех случаях наша зрительная система обманута, поскольку перцепты глубины сугубо иллюзорны и возникли от изображений, которые на самом деле совершенно плоски.

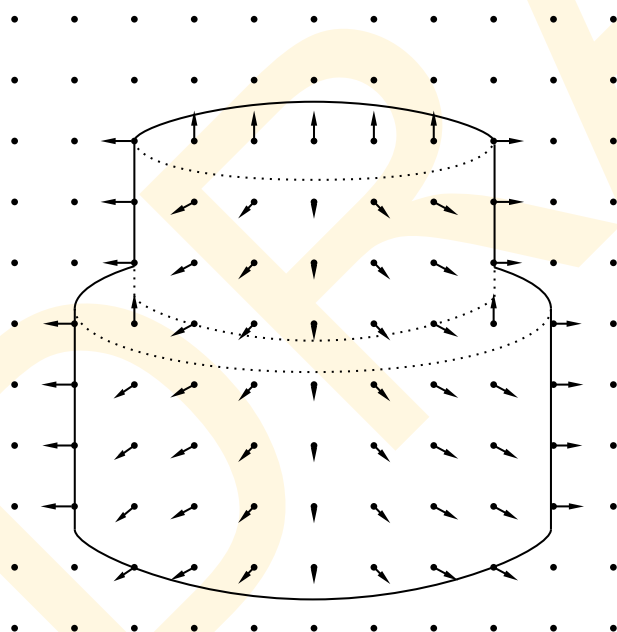
Нормальное, повседневное восприятие глубины пространства возможно благодаря тому, что зрительная система в неявном виде выполняет высокоточные эвристические предположения как о характере внешнего мира, так и об условиях его наблюдения. В совокупности со специфической информацией, извлекаемой нами из обоих ретиальных изображений, предположения эти оказываются достаточны для реконструкции информации о глубине пространства. Причем с точностью несоизмеримо большей, чем та, которая следовала из логического анализа данного выше. Однако же и от логической недоопределенности информации о глубине пространства есть некоторый толк -- в том смысле, что недоопределенность эта помогает найти условия, в которых систему восприятия глубины можно ввести в заблуждение. Условия эти дают критично важную информацию о характере допущений, сделанных этой системой. Например, наше зрение обманывается стереофильмами, поскольку зрительная система «по умолчанию» полагает, что оба глаза направлены в одну и ту же сцену. Объемный эффект возникает благодаря специальным очкам, создающим немного разные изображения на сетчатках обоих глаз: зрительная система интерпретирует эти отличия, как следствие глубины пространства, которая фактически не представлена в стимуляции. Однако же в нормальных условиях, т.е. в 99.99% случаев, такое предположение оказывается верным, позволяя нам, на основе легкого отличия между ретиальными изображе-

ниями в левом и правом глазу, воспринимать глубину пространства как фактически имеющую место.

Коль скоро зрительную систему можно обмануть средствами современной визуальной технологии, легко подумать, что ретинальная стимуляция не несет в себе существенной информации о глубине пространства. На самом же деле этой информации невероятное количество: в нормальных условиях наблюдения, то есть, когда глаза наблюдателя привычно «сканируют» равномерно освещенную сцену, заполненную непрозрачными поверхностями — информации более чем достаточно (о чем, кстати, настойчиво говорил Гибсон). Однако же равновелико важно и то, что различные факторы глубины одинаково интерпретируются зрением. И сие есть конвергенция множества источников информации, обеспечивающая ту удивительную точность восприятия глубины пространства, коей мы обладаем. Только в лабораторных условиях или в условиях, специально разработанных для иллюзиона, наше зрение обманывается в определении реального расстояния до поверхностей сцены.

### 5.1.2 Марров 2.5D-скетч

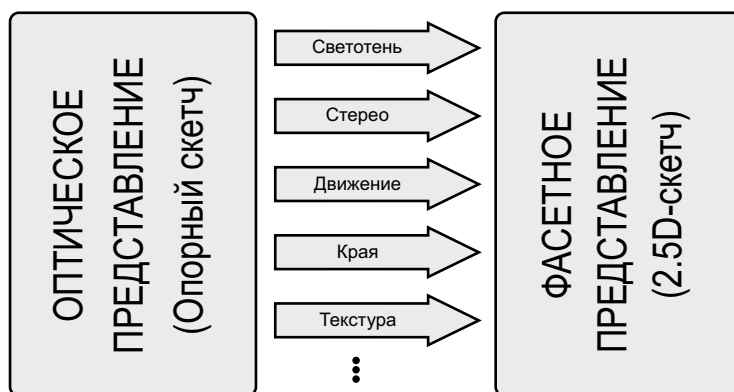
Каким образом в зрительной системе представлена визуальная информация о планировке поверхностей реальной сцены? По сей день наиболее убедительной остается концепция 2.5D-скетча, предложенная Дэвидом Марром в начале 80-х годов прошлого века (см. рис. 5.1.2). Как ясно из самого названия, 2.5D-скетч — это нечто среднее между двумерными параметрами оптического представления и трехмерными параметрами представления подлинно предметного. То есть, имеет место конвергенция



**Рис. 5.1.2** Марров 2.5D-скетч. В каждой позиции поля зрения показан вектор (нормаль к поверхности), перпендикулярный локальному участку поверхности в этой позиции и выглядящий как иголка, торчащая из нее. 2.5D-скетч также несет в себе информацию о расстоянии по зрительной оси (по материалам Marr & Nishihara, 1978).

выходных сигналов множества процессов, восстанавливающая информацию о глубине и ориентации отдельных участков поверхностей сцены в подходящее представление ориентации и расстояния. Об этом говорил Гибсон еще задолго до Марра, но Марр пошел дальше, предложив концепцию того, как такое представление можно фактически вывести из оптических ретинальных изображений.

Точка зрения на процессы реконструкции ориентации поверхностей и их удаления от наблюдателя, в основе которых лежит 2.5D-скетч, состоит в том, что существует множество независимых (или почти независимых) модулей обработки, «обсчитывающих» информацию о глубине, идущую из разных источников (рис. 5.1.3). Каждый модуль обрабатывает данные «своего» вида, внося индивидуальный вклад в общую интерпретацию глубины на



**Рис. 5.1.3.** Общая схема работы модуля восприятия глубины пространства. 2.5D-представление можно рассчитать из двухмерного оптического представления посредством множества параллельных и квазинезависимых процессов, извлекающих информацию об ориентации поверхности и глубине пространства из различных источников, таких как светотень, стерео, дви-



**Рис. 5.1.4.** Факторы визуальной глубины пространства.

уровне 2.5D-скетча. Специалисты по машинному зрению часто относят эти процессы к категории «форма-из-Х-модуля», где «Х» — это источник информации о глубине пространства, как, скажем «форма-из-светотени» или «форма-из-движения». Вместе с тем, хотя сия терминология и удобна, она все же несколько некорректна, поскольку на выходе этих модулей форма не представлена в явном виде. На выходе у них проходит информация либо о глубине, либо об ориентации поверхности, поэтому скорее подойдет формулировка «глубина-из-Х», либо «ориентация-из-Х». Представление же формы намного сложнее, что мы покажем в гл. 8.

Источников информации о глубине пространства множество. И множество вариантов их классификации (рис. 5.1.4).

Первым критерием в систематике является состояние глаз как таковых (*окулярная информация*), либо структура света, упавшего на сетчатку (*оптическая информация*).

Вторым критерием является то, исходит ли эта информация от обоих глаз (*бинокулярная информация*) или от одного (*монокулярная*).

Третий параметр — можно ли извлечь эту информацию из неподвижного изображения (*статическая информация*) или же требуется движение наблюдателя или объекта (*динамическая информация*).

Четвертый критерий — источник, определяющий фактическое расстояние до объекта (*абсолютная информация*), либо описывающий то, каково относительное пространственное взаиморасположение объектов (*относительная информация*).

Пятый параметр касается численных соотношений по расстояниям (*количественная информация*), либо порядковых отношений — ближе/дальше (*качественная информация*).

Итак, мы организовали разговор об источниках информации о глубине в логике четырех категорий: окулярной информации (аккомодация и конвергенция), стереоскопической информации (бинокулярная диспаратность и стереопсис Леонардо), динамическая информация (параллакс движения и динамика текстур) и пикториальная информация (всё остальное). При этом мы не считаем эту классификацию единственно верной.

## 5.2 ОКУЛЯРНАЯ ИНФОРМАЦИЯ

Окулярная информация о расстоянии до фиксированной поверхности исходит от факторов, зависящих как от состояния самих глаз, так и их отдельных компонент. В частности, критично важной в восприятии глубины пространства является фокусировка хрусталиков (*аккомодация*) и угол между зрительными осями (*конвергенция*). Аналогичная информация используется в системах машинного зрения, при электронном и/или механическом управлении аналогичными функциями пары видеокамер.

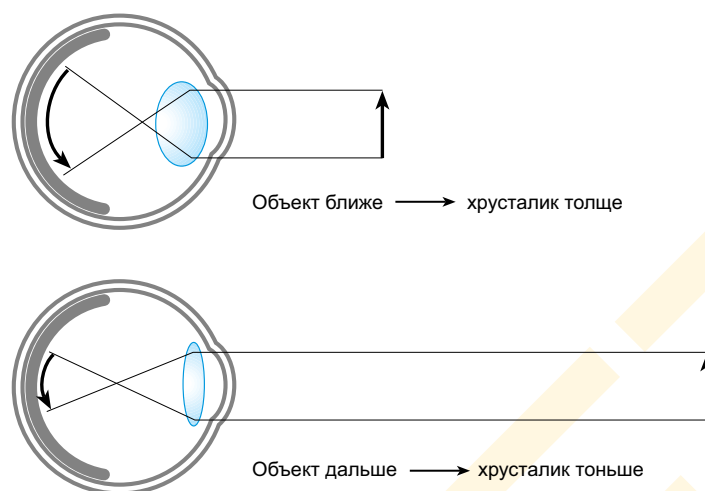
### 5.2.1 Аккомодация

⇒ Аккомодация — это процесс при котором цилиарные мышцы глаза, сокращаясь и расслабляясь, изменяют кривизну хрусталика посредством чего управляют его фокусным расстоянием.

Сие сугубо монокулярный фактор зрительного восприятия глубины пространства, поскольку может исходить из одного глаза, даже несмотря на то, что работают оба. Как было показано в разделе 1.3, хрусталик человека способен менять свое фокусное расстояние, уплотняясь при фокусировке по удаленным объектам и становясь толще — по близким (рис. 5.2.1). Когда зрительная система получает информацию о напряжении мышцы, меняющей кривизну хрусталика — она фактически получает информацию о расстоянии до объекта, по которому идет фокусировка.

Несмотря на то, что аккомодацию полагают весьма слабым источником визуальной информации о глубине пространства (Hochberg, 1971), эксперименты показали, что мы активно пользуемся ею на малых расстояниях: оценка собственно расстояния силами аккомодации малоэффективна, но весьма полезна в оценке размера предметов (Wallach & Floor, 1971). То есть, на близких расстояниях мы оцениваем размеры объектов с учетом информации от цилиарных мышц. Однако же за пределами 2–2.5 м от





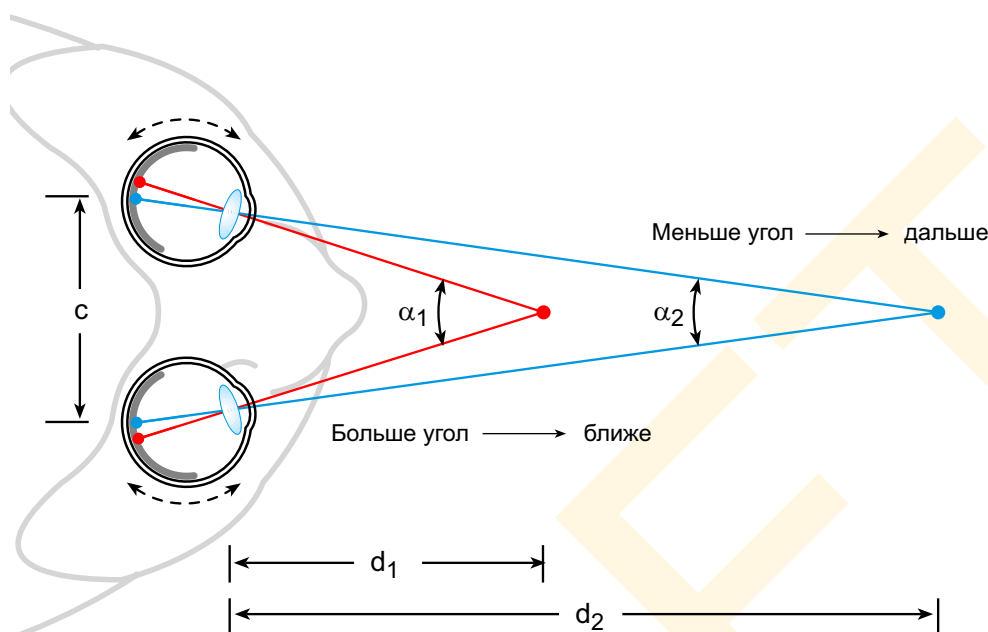
**Рис. 5.2.1** Аккомодация как фактор зрительного восприятия глубины пространства. Меняя толщину, хрусталик фокусирует свет на поверхности сетчатки. Когда объект близко — хрусталик утолщается за счет напряжения цилиарной мышцы; когда объект далеко — утончается за счет ее расслабления.

наблюдателя аккомодация уже не дает (либо почти не дает) какой-либо значимой информации о глубине пространства: на этих расстояниях цилиарные мышцы находятся в полностью расслабленном состоянии и хрусталик уже не может стать плосче.

Мы говорим, что аккомодация дает информацию об абсолютной глубине, то есть описывает фактическое расстояние до статического объекта, позволяя зрительной системе, если угодно, самокалиброваться. Как видно по рис. 5.1.4, большинство оптических факторов восприятия глубины пространства несут либо информацию об *относительном* расстоянии (т.е. о том, который из двух предметов ближе к наблюдателю, который дальше), либо об отношении расстояний до двух объектов. Относительная информация о глубине бесспорно необходима, но информация абсолютная требуется особи для установки фактического расстояния до предмета сцены. Если бы мы не могли определять фактическое расстояние, особенно на малых удалениях, мы бы постоянно спотыкались о предметы, которые оказались ближе, чем мы думали, и тщетно тянулись бы к тем, что на самом деле далеко.

Чтобы аккомодация дала точную информацию об удалении объекта, объект этот должен быть сфокусирован на сетчатке. Сие предполагает, что зрительная система должна каким-то образом «узнать», когда объект оказался в фокусе. Наилучшим индикатором фокусировки является появление острых контуров. То есть, мы говорим, что изображение размытое является весьма эффективным пусковым механизмом аккомодации. Как уже обсуждалось в разделе 4.3, информация о размытости/резкости краев проходит по высоким пространственным частотам: у острых краев больше энергии на этих частотах, нежели у размытых. Поэтому, вероятнее всего, зрительная система управляет аккомодацией, меняя напряжение цилиарных мышц и добываясь при этом максимального сигнала на выходе высокочастотных каналов.

Однако из-за ограниченности диапазона применения, аккомодация у приматов не является критично важным источником информации о глубине пространства. Тогда как, скажем, у африканского хамелеона она играет наиважнейшую роль в поиске пищи (Harkness, 1977). Хамелеон ловит добычу, молниеносно выкидывая липкий язык, цеп-



**Рис. 5.2.2** Конвергенция как фактор зрительного восприятия глубины пространства. Угол между глазами ( $\alpha$ ) меняется при изменении расстояния до объекта, на котором фиксирован взгляд: чем дальше объект, тем меньше угол конвергенции ( $\alpha_2$ ) и наоборот ( $\alpha_1$ ).

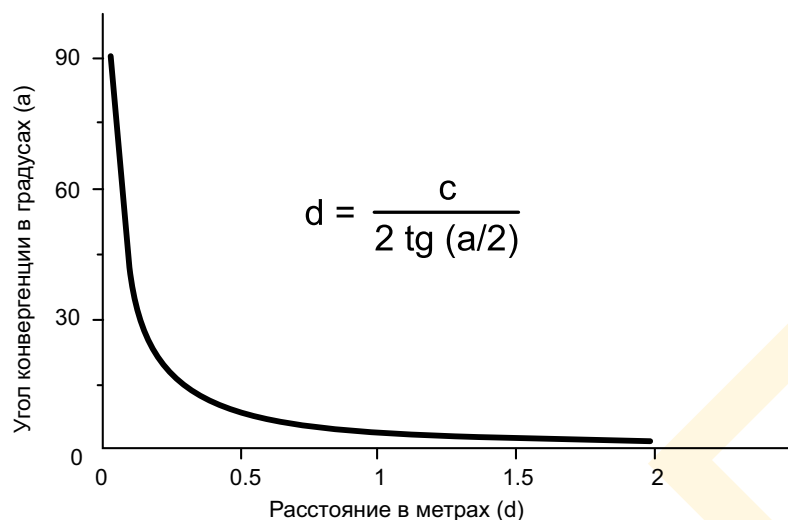
ляющий насекомое, и быстро втягивает его назад в рот. Отметим, что сие возможно потому, что малый диапазон расстояний релевантен открытой пасти животного.

### 5.2.2 Конвергенция

Еще одним источником информации о глубине пространства является т.н. *конвергенция*, при которой глазные яблоки подворачиваются внутрь (друг к другу), фиксируя взгляд на объекте. *Фиксация* глаз на точке во внешнем мире возникает тогда, когда оба они ориентированы так, что свет от данной точки сцены одновременно падает в центральные ямки обеих сетчаток. Действительно, коль у fovea один центр, на нем в данный момент времени может зафиксироваться лишь одна точка (см. рис. 5.2.2).

Критично важным фактом является то, что угол между двумя зрительными осями систематически меняется по мере изменения расстояния между наблюдателем и точкой фиксации. Фиксация взгляда на близко расположенной точке ведет к большому углу конвергенции, а на дальней точке, соответственно, к малому (рис. 5.2.2). Поскольку конвергенция — это сочетанная работа двух глаз, мы говорим, что сие, в отличие от аккомодации, — *бинокулярный* источник визуальной информации о глубине пространства. Однако подобно аккомодации, конвергенция дает информацию об абсолютном расстоянии до объекта.

Геометрия бинокулярной конвергенции дана на рис. 5.2.2: тригонометрическое уравнение, соотносящее расстояние до объекта с углом конвергенции — на рис. 5.2.3. Очевидно, что ежели расстояние между центрами глаз известно, то угол конвергенции можно использовать при расчете расстояния до объекта. Но так ли это? Эксперименты подтвердили — да, так (например, Hofsten, 1976), но только в диапазоне нескольких метров. Объяснение сего ограничения становится понятным, если угол конвергенции



**Рис. 5.2.3** Конвергенция как функция от расстояния. Угол конвергенции весьма активно меняется в диапазоне до 1-2 метров и очень слабо на больших расстояниях.

Действительно, если объект фиксации взгляда сдвинулся ближе или дальше, даже когда один из глаз закрыт, то степень конвергенции изменится должным образом, управляемая при этом аккомодацией хрусталика открытого глаза. Сие создает сложности в раздельной оценке вкладов аккомодации и конвергенции в восприятие глубины пространства. Однако еще раз подчеркнем, что аккомодация и конвергенция — весьма и весьма важные источники о глубине на малых удалениях, особенно среди тех немногих источников, что дают информацию об абсолютном расстоянии до объекта.

дать как функцию от расстояния до объекта фиксации (рис. 5.2.3): на малых расстояниях угол этот сильно меняется, даже когда точка фиксации смещается незначительно; однако за пределами 2.5–3 метров он меняется крайне слабо и уходит практически в ноль, когда взгляд устремляется на бесконечно удаленную точку.

Мы говорим, что конвергенция и аккомодация обычно ковариантны: если расстояние до объекта фиксации меняется, синхронно меняются степень и аккомодации, и конвергенции.

### 5.3 СТЕРЕОСКОПИЧЕСКАЯ ИНФОРМАЦИЯ

Возможно, наиболее острое ощущение глубины дает нам т.н. *стереопсис* — физиологический механизм зрительного восприятия относительного расстояния до объектов, основанный на латеральном смещении их ретинальных изображений. Стереопсис возможен благодаря тому, что у нас два латерально разделенных глаза, поля зрения которых накладываются друг на друга по центру взгляда.

Поскольку глазные яблоки на несколько сантиметров разнесены в пространстве, ретинальные изображения большинства объектов в участке наложения слегка разнятся. То есть, одна и та же точка внешнего оптического массива дает проекции на левую и правую сетчатки немного по-разному, и характер этого отличия зависит от того, насколько ближе или дальше эта точка от точки фиксации взгляда. Сие относительное латеральное смещение именуется *бинокулярной диспаратностью*.

В 1838 г. британский физик и изобретатель, сэр Чарльз Витстон первым обратил пристальное внимание на тот факт, что наши глаза смотрят на мир со слегка разных пространственных позиций:

«Ныне вполне очевидно, почему художники не в силах дать нам добротное представление близкорасположенных объектов, то есть рисунок, который был бы умом

нашим неотличим от самого объекта. Когда рисунок и объект рассматриваются обоими глазами, то в случае рисунка на сетчатке образуются две *почти одинаковых* картинки, но в случае реального объекта — эти картинки *разные*. Следовательно, в этих двух ситуациях имеется ощутимая разница между воздействиями на органы наших чувств и, соответственно между формами перцептов в сознании. Таким образом, рисунок объекта никак не может отождествляться с реальным объектом».

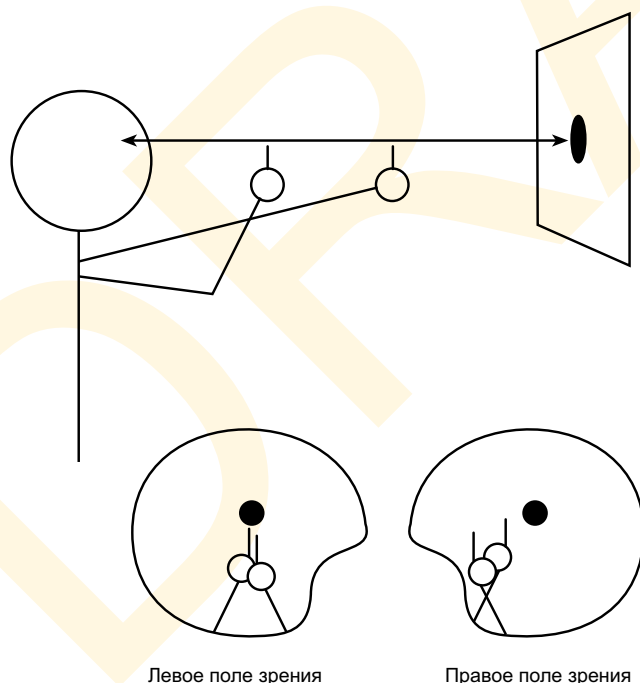
### 5.3.1 Бинокулярная диспаратность

Ретинальные изображения достаточно похожи, и вряд ли мы увидим существенную разницу в видимой сцене, если вначале закроем один глаз, а потом второй, открыв первый. Однако продемонстрировать латеральную диспаратность одной и той же точки внешнего оптического массива можно с помощью специального опыта: закроем правый глаз и полностью вытянем вперед левую руку с поднятым вверх указательным пальцем; затем на половину этого расстояния вытянем правую руку также с поднятым указательным пальцем; зафиксируем взгляд на какой-то точке в пространстве и совместим с нею оба указательных пальца (рис. 5.3.1). Теперь, не меняя положения рук, быстро откроем правый глаз и закроем левый. Что произошло? Оба пальца резко «сдвинулись» влево относительно точки фиксации. Так произошло потому, что пространственный разнос глазных яблук дает несколько отличные друг от друга ретинальные изображения внешнего мира, и латеральное «смещение» пальцев — яркая демонстрация тому, что бинокулярная диспаратность имеет место. Когда такое смещение регистрируется нашей зрительной системой в двух одновременно существующих

ретинальных изображениях, оно интерпретируется так, что оба пальца и предмет фиксации расположены на разных расстояниях от наблюдателя. То есть, результатом бинокулярной диспаратности является ощущение того, что объекты находятся от нас на разном удалении (но при этом не того, что они смещены друг относительно друга).

Визуальная информация, которую нам удастся извлекать благодаря бинокулярной диспаратности, намного точнее, чем может показаться на первый взгляд: *направление* смещения говорит нам о том, какая точка в пространстве ближе точки фиксации, а какая дальше от нее; *величина* смещения — о том на сколько ближе или дальше.

Чтобы продемонстрировать количественные отношения



**Рис. 5.3.1** Демонстрация бинокулярной диспаратности. Тот факт, что наши глаза смотрят на мир с немного разных позиций, можно продемонстрировать пальцевым опытом, описанном в тексте.

между степенью диспаратности и глубиной, повторим эксперимент с пальцами, но при этом обратим внимание на то, насколько они "смещаются" от точки фиксации: ближний палец «сместится» больше, дальний — меньше. Чем ближе палец к лицу — тем сильнее смещение. Количественный характер бинокулярной диспаратности исключительно важен для восприятия, поскольку, давая информацию лишь об относительной глубине, он описывает *соотношение* расстояний, а не только номинальное ближе/дальше.

Чтобы продемонстрировать взаимосвязь между направлением смещения и глубиной, еще раз повторим опыт, но на сей раз сосредоточим взгляд на ближнем пальце, а не на дальнем объекте: вначале левым глазом, затем правым (это существенно сложнее, но вполне осуществимо, если поднапрячься). Мы увидим при этом, что дальний палец и дальний объект «сдвинулись» на сей раз вправо (а не влево) относительно объекта фиксации (которым теперь является ближний палец). Следовательно, когда мы переключаемся с левого глаза на правый, «смещение» объектов влево интерпретируется как то, что они расположены ближе точки фиксации, а вправо — как то, что дальше. Сие демонстрирует взаимосвязь между направлением смещения и глубиной относительно точки фиксации. Поразмыслив над незатейливой геометрией бинокулярной диспаратности, мы поймем, каковы эти взаимоотношения и какой информацией пользуется зрительная система, когда интерпретирует смещение ретинальных изображений в глубину.

#### 5.3.1.1 *Согласованные ретинальные позиции*

Первым шагом в нашем понимании геометрии бинокулярной диспаратности является определение т.н. *согласованных позиций* в обеих сетчатках:

- ⇒ Согласованными мы назовем такие позиции, которые совпадут, если простым латеральным смещением суперпонировать центральные ямки левого и правого глаз.

Для иллюстрации представим каждую сетчатку в виде половинки яичной скорлупы с *fovea* в центре. Затем вообразим, что мы совмещаем обе сетчатки, вставив одну скорлупу в другую, причем так, что центры совпали. Теперь представим, что мы проткнули иголкой обе скорлупки, а затем вынули одну из другой и вернули на свои места — на заднюю поверхность глазных яблок. Дырочки от иголки расположатся в согласованных одна с другой пространственных позициях обеих сетчаток.

Диспаратность же возникает тогда, когда проекция данной точки внешнего оптического массива попадает в *несогласованные* позиции левой и правой сетчаток.

**Кросс- и анкроссдиспаратность.** Рассмотрим схему, на которой глазные яблоки фиксированы на точке «Р» (рис. 5.3.2). По определению проекция точки «Р» падает в центральные ямки обеих сетчаток и стимулирует фоторецепторы в согласованных позициях. Теперь посмотрим на проекции ближней точки «С» в момент, когда взгляд фиксирован на «Р»: как видно по рис. 5.3.2 проекции падают в несогласованные позиции — на правой сетчатке справа от центральной ямки, на левой — слева. Разбегание проекций кнаружи именуется *кроссдиспаратностью* (или *перекрестной диспаратностью*). Кроссдиспаратность двух изображений точки «С» указывает на то, что точка эта ближе точки фиксации («Р»). Насколько ближе — зависит от степени разбегания. Глядя на рис. 5.3.2, мы можем представить себе, что чем ближе точка С к наблю-



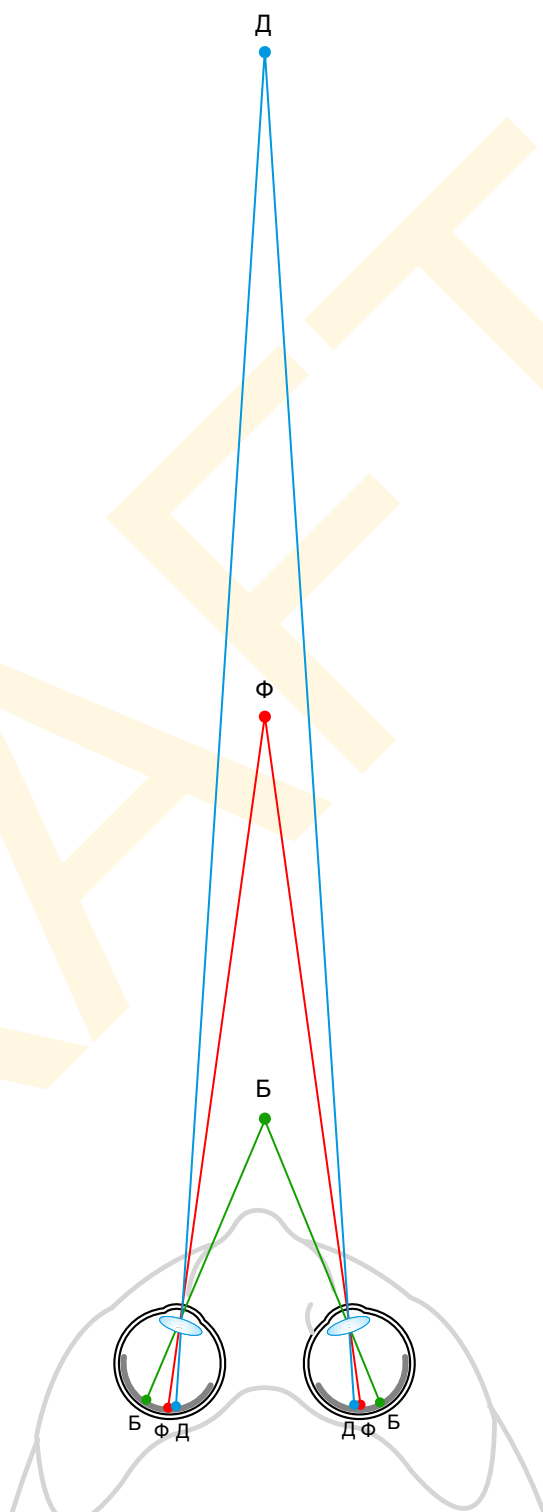
дателю, тем выше перекрестная диспаратность ее позиции в ретинальных изображениях (т.е. тем больше разнос между правым и левым ретинальными изображениями).

Теперь рассмотрим ретинальные проекции точки *Г*, которая расположена дальше точки фиксации взгляда (*Р*). Как видно из рис. 5.3.2, проекции этой точки также падают на диспаратные участки, но в этом случае в правой сетчатке проекция оказывается слева от центральной ямки, а в левой — справа. Сие внутреннее (назальное) направление смещения именуется *анкроссдиспаратностью*, указывая на то, что точка ее вызывавшая (в нашем примере «*Г*») находится дальше точки фиксации. И опять же: чем дальше точка, тем выше диспаратность.

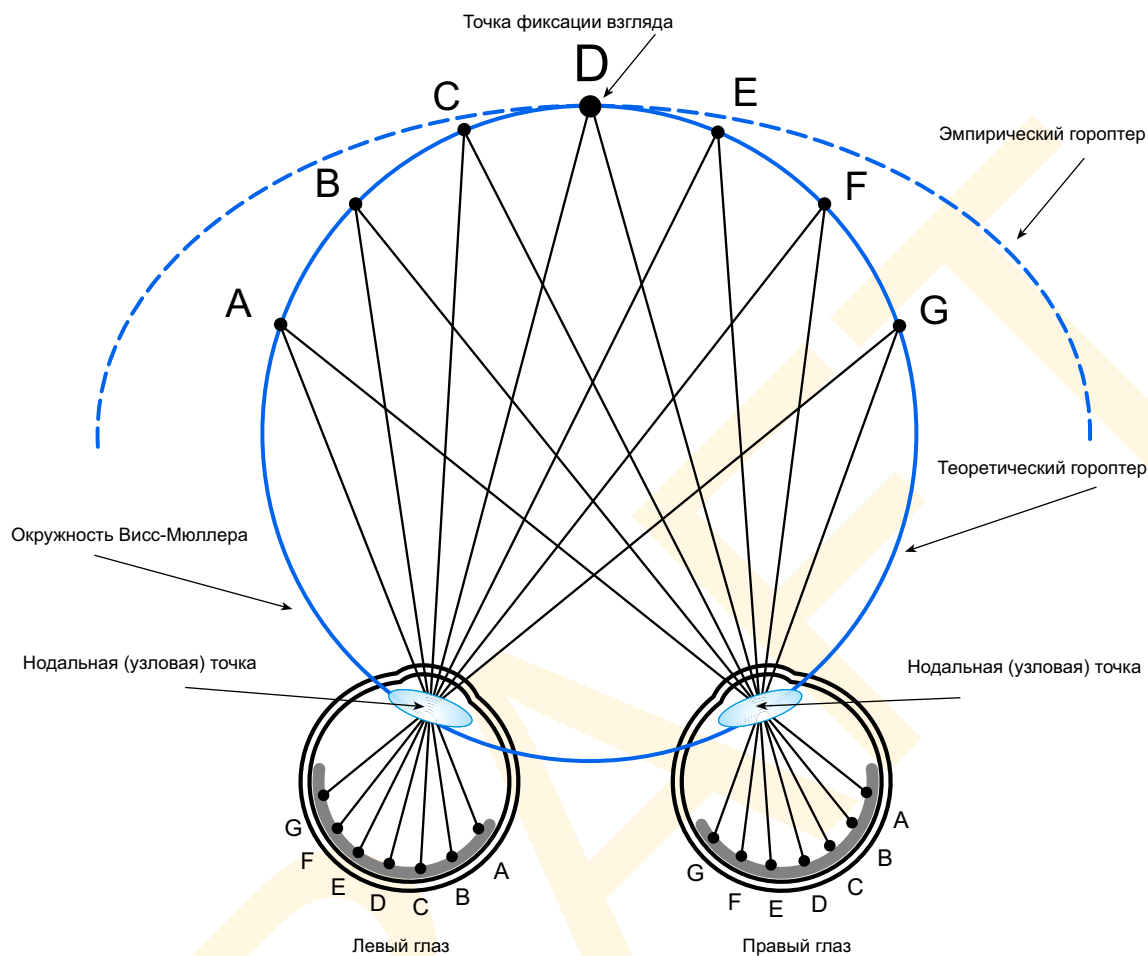
### 5.3.1.2 Гороптер

Не все точки внешнего оптического массива дают диспаратные проекции: наиболее типичный пример — это точка фиксации, дающая проекции строго в центр обеих foveae. Множество точек внешнего оптического массива, посылающих свет в согласованные позиции сетчаток, именуется *гороптером*. Существуют два варианта определения гороптера: теоретический (геометрический) и эмпирический (экспериментальный).

**Теоретический гороптер** можно определить проекцией пар согласованных ретинальных точек во вне через узловые точки обоих глаз. Точки пере-



**Рис. 5.3.2** Кросс- и анкроссдиспаратность. Если «*Ф*» — точка в фиксации взгляда, то те точки, что ближе нее к наблюдателю (например, точка «*Б*») ложатся на сетчатку кроссдиспаратно — т.е. кнаружи; те, что дальше (например, «*Д*») — анкроссдиспаратно, т.е. кнутри.

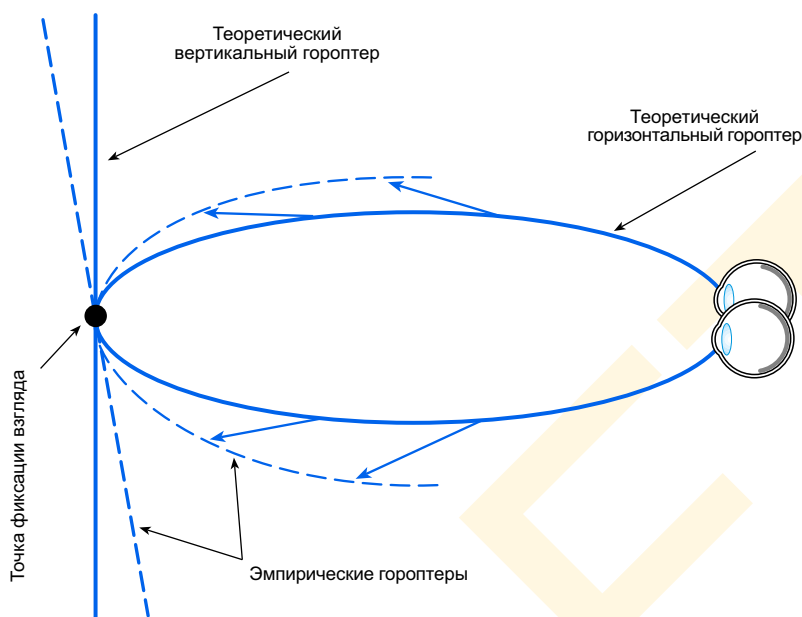


**Рис. 5.3.3** Гороптер в горизонтальной плоскости. Множество точек в сцене, дающих проекции в согласованные точки сетчатки обоих глаз, именуется *гороптером*. Теоретический гороптер в горизонтальной плоскости является окружностью, именуемой *окружностью Висс-Мюллера*. Окружность эта проходит через нодальные точки обоих глаз и точку фиксации взгляда (D). Эмпирический гороптер (пунктирная кривая) находится немного кзади от гороптера теоретического.

сечения всех возможных пар линий образуют во внешней среде множество точек, которое и составит теоретический гороптер. В горизонтальном сечении он явится собой окружность, именуемую *окружностью Висс-Мюллера*. Окружность эта проходит через точки фиксации и узловые точки обоих глаз (рис. 5.3.3). Диспаратность каждой из этих точек, теоретически, равна нулю. В вертикальной плоскости, локали, дающие проекции в согласованные точки сетчаток, лежат на прямой линии перпендикулярной зрительной оси (рис. 5.3.4).<sup>1</sup>

**Эмпирический гороптер** определен результатами психофизических экспериментов, выполненных разными способами, но в целом, позиционирующих гороптер как

<sup>1</sup> Отметим, что вертикальные и горизонтальные плоскости позиционируются относительно ориентации глазных яблок, но не объектов сцены. Горизонтальный гороптер окажется в вертикальной плоскости, когда наблюдатель смотрит строго вниз.



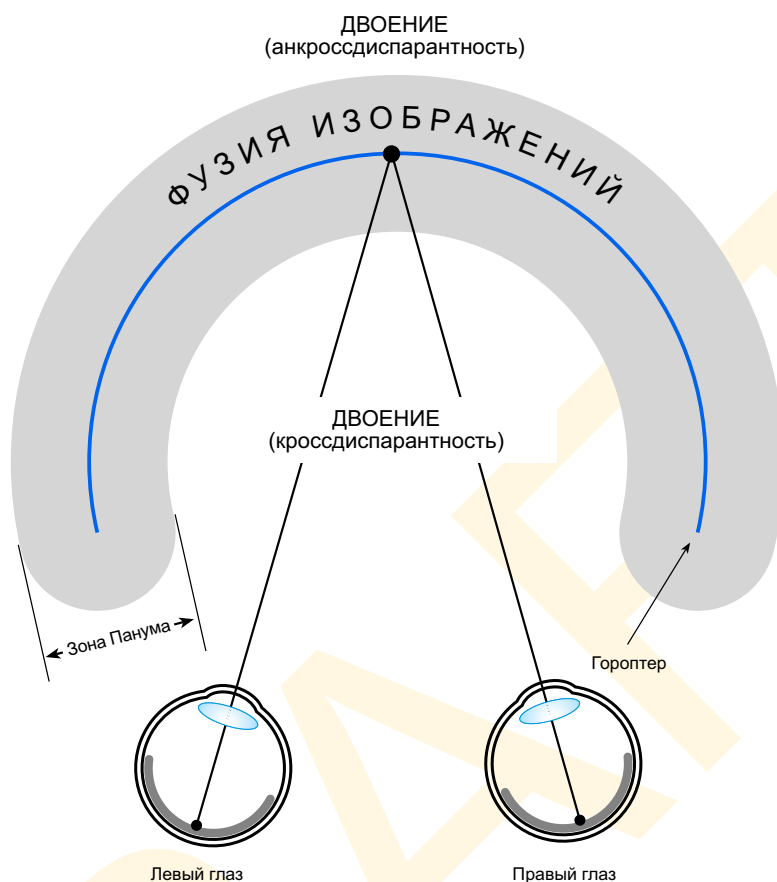
**Рис. 5.3.4** Теоретический и эмпирический гороптеры в вертикальной плоскости. В вертикальной плоскости теоретический гороптер представляет собой вертикальную прямую, перпендикулярную зрительной оси в точке фиксации взгляда. Эмпирический гороптер в вертикальной плоскости (пунктир) — это прямая, слегка отклоненная кзади.

горизонтальную плоскость, лежащую слегка кзади от гороптера теоретического — рис. 5.3.3 (Ogle, 1950). В вертикальной плоскости эмпирический гороптер отклоняется от наблюдателя выше точки фиксации и к наблюдателю — ниже ее (рис. 5.3.4). По сей день не совсем ясно, отчего же эмпирический гороптер не идентичен теоретическому. Правда, отличие это невелико и в большинстве практических ситуаций им можно смело пренебречь.

Таким образом, стереоскопический компонент в восприятии глубины возникает из различных направлений и углов ретиальной диспаратности точек внешнего оптического массива, лежащих перед и позади гороптера. Поразмыслив над этим, приходим к мысли, что бинокулярная диспаратность — это источник скорее относительной, нежели абсолютной информации о глубине пространства: направление и величина диспаратности определяет то, что данная точка находится ближе или дальше от гороптера.

Абсолютное расстояние до гороптера, меняется при изменении расстояния до точки фиксации взгляда. Следовательно, чтобы зрительная система могла добыть абсолютные данные об удалении предмета — информацию, извлекаемую из бинокулярной диспаратности, следует объединить с информацией, поступающей от аккомодации и конвергенции. Если известно абсолютное расстояние до одного объекта, то абсолютные расстояния до других можно вывести, шкалировав относительные величины расстояний.

Прямым следствием из сего разговора является вполне закономерный вопрос: почему в наших глазах не двоится (*диплопия*)? В самом деле, точки сцены, попадающие бинокулярную зону поля зрения и образующие два ретиальных изображения, редко когда делятся в нашем сознании. Причин сего как минимум две: во-первых, точки, находящиеся на гороптере, или возле него, перцепционно сливаются в единый образ. Об-



**Рис. 5.3.5** Фузионная зона Панума. Точки, лежащие в панумовой области сливаются в сознании наблюдателя в единый образ. Точки, лежащие перед и после зоны Панума, дwoятся (соответственно кросс- и анкроссдиспарантно).

ласть возле горюптера, в которой диспарантные точки сцены сливаются в общий перцепт, именуется *фузионной зоной Панума* (рис. 5.3.5). Во-вторых, диспарантность точек, лежащих вне зоны Панума, дает перцепт глубины пространства. Однако мы все же *сможем* ощутить двоение, если диспарантность чрезмерна, — например, когда пытаемся посмотреть на кончик собственного носа.

Чтобы через бинокулярную диспарантность продемонстрировать двойственный характер зрения, повторим «пальцевой» опыт (напомним — один палец на вытянутой руке, второй — на середине расстояния от кончика носа до первого пальца), но на сей раз оба глаза будут открыты. Если мы сосредоточим взгляд на удаленном объекте, то оба пальца «раздwoятся», поскольку окажутся вне зоны Панума. Ежели мы повторим опыт с объектом, лежащим не далее чем на 2-2.5 см от дальнего пальца, то мы увидим раздвоенное изображение ближнего пальца, но не дальнего. Так происходит потому, что дальний палец оказался вплотную к горюптеру, заданному дальним объектом, лежащим в панумовой зоне. Ближний палец окажется раздвоенным потому что по-прежнему находится вне зоны Панума.

Стереоскопическое зрение исключительно чувствительно к бинокулярной диспарантности. Большинство людей способны увидеть отличие в нескольких секундах зрительного угла ( $1/3600$  градуса, поскольку одна угловая секунда — это  $1/60$  угловой ми-

нуты, которая, в свою очередь — это  $1/60$  градуса). Единственное ограничение в том, что работает стереопсис лишь в диапазоне 30 метров. За пределами этого диапазона диспаратность уже слишком мала для обслуживания стереоскопии. От стереоскопии мало толку и по центральной части поля зрения, где ретинальные проекции сцены накладываются друг на друга.

Однако же не всем двуглазым даровано стереоскопическое восприятие глубины: *стереослепотой* поражены примерно от 5 до 10% особей в популяции (Richards, 1970) и, чаще всего, она вызывается болезнями глаз, перенесенными в детстве. Первая и наиболее общая причина стереослепоты — т.н. *страбизм* (косоглазие). Если сия патология купирована хирургически в раннем возрасте, то стереоскопическое зрение будет нормально развиваться (Banks, Aslin & Letson, 1975). Стереопсис отсутствует также у детей, нормально видящих лишь одним глазом (по причине либо монокулярной близорукости, либо катаракты). Однако же стереослепые, разумеется, воспринимают глубину пространства (мы же воспринимаем ее, закрыв один глаз) но эти люди не могут извлечь полезную информацию из бинокулярной диспаратности.

### 5.3.1.3 Стереограммы

Возможно, наиболее эффектной демонстрацией того, что бинокулярная диспаратность способна вызывать зрительное ощущение глубины пространства, являются т.н. *стереограммы*: парные изображения, отличные друг от друга относительным латеральным смещением элементов и которые при специальном стереоскопическом рассматривании дают мощную иллюзию глубины. Стереограммы предложил в 1838, упомянутый выше, Чарльз Витстон, занимавшийся в свое время вопросами геометрии бинокулярной диспаратности. Ученый обнаружил, что если правому и левому глазу предъявить изображения, отличные друг от друга лишь некоторым латеральным смещением изображенных объектов, объекты эти будут восприниматься расположенными на разном удалении от наблюдателя. Простейшие стереоскопы Витстона были очень модны в конце XIX века, а спустя время их несколько усовершенствованные версии стали весьма популярной детской игрушкой. Стереограммы можно рассматривать без всяких специальных приспособлений, что мы сейчас и продемонстрируем.

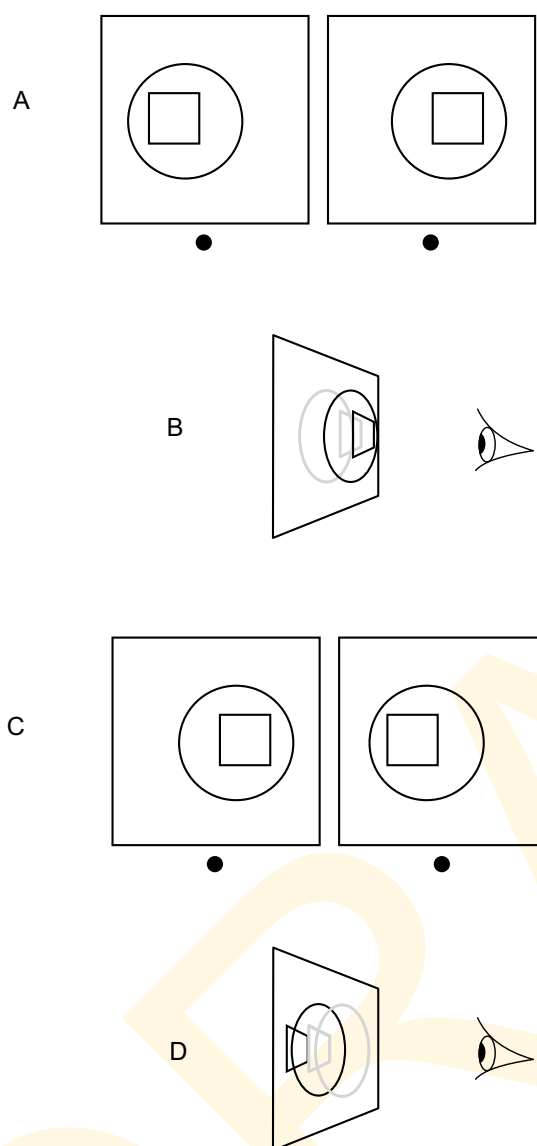
Ключевой особенностью стереограмм является то, что проекции объектов в левом и правом изображениях латерально смещены друг относительно друга, создавая бинокулярную диспаратность. Направление диспаратности (кросс- или анкросс-) и ее мера задают степень визуальной глубины. На рис. 5.3.6 А показана одна из таких стереопар. Отметим, что относительно наружных квадратов круги лишь слегка смещены кнаружи, а внутренние квадраты — смещены существенно. Если изображения стереопары стереоскопически совместить так, как предложено ниже, латеральная диспаратность даст перцепт, аналогичный тому, что показан на рис. 5.3.6 В.

На рис. 5.3.6 С показана стереопара с анкроссдиспаратностью, дающая аналогичный перцепт, но с обратным взаиморасположением элементов по глубине (рис. 5.3.6 D).

Отметим, что результат обоих опытов в точности соответствует нашим прогнозам результатов кросс- и анкроссдиспаратности.

**Кросс- и анкроссконвергенция.** Итак, чтобы добиться ощущения глубины при разглядывании стереограмм, мы должны так поставить наши глаза, чтобы мозг объединил нейронные изображения, индуцированные изображениями ретинальными. Существует два способа: *кросс-* и *анкроссконвергенция*. Изображения на рис. 5.3.6 А и





**Рис. 5.3.6** Бинокулярная диспаратность и стереограммы. Если изображения в части А стереоскопически слились кроссконвергентно (см. в тексте), круг и квадрат будут восприниматься над страницей, то есть так, как показано в части В. Если изображения слиты по методу анкросконвергенции, то круг и квадрат будут восприниматься позади страницы (D).

ложенной строго по середине между двумя точками.<sup>1</sup> Теперь сосредоточим внимание на среднем изображении и присмотримся к кругу и квадрату, чудесным образом по-

5.3.6 С построены так, чтобы давать перцепт глубины при кроссконвергенции глаз, который мы опишем первым. Рассматривая наши стереопары по методу анкросконвергенции, получим перцепты обратной глубины, то есть стереопара рис. 5.3.6 А даст перцепт, похожий на тот, что вызывается рис. 5.3.6 D, а стереопара на рис. 5.3.6 С — на тот, что вызывается рис. 5.3.6 В.

Чтобы добиться стереоскопического слияния по методу кроссконвергенции, нам придется приложить такое волевое усилие к глазодвигательным мышцам, чтобы левая половина стереопары в правом глазу сошлась с правой ее половиной в левом. Если мы промахнемся с усилием, то увидим четыре фигуры, поскольку каждым глазом схвачены по паре изображений. Цель кроссконвергенции в том, чтобы получить *три* изображения, среднее из которых — плод слияния двух, вместе дающих иллюзию глубины.

Свести глаза волевым усилием напрямую без тренировки довольно сложно, но есть обходной маневр. Сосредоточим взгляд на кончике карандаша, расположенном между точками под изображением, а затем начнем постепенно приближать карандаш к глазам, удерживая фокус на его острие. По мере выполнения сего мы обнаружим, как каждая из половин стереограммы начнет двоиться (диплопия), в результате чего появятся четыре изображения и четыре точки. На определенном расстоянии от страницы границы двух (из четырех) изображений четко сядут друг на друга — это момент, когда нужно остановить карандаш, острие которого оказалось теперь на точке, распо-

<sup>1</sup> Если изображения на странице и наши зрительные оси оказались в разных горизонтальных плоскостях, нужно аккуратно подвигать лист вверх-вниз.

висшими над квадратным фоном. Иллюзия может продлиться секунды, но если она возникла — это будет мощный перцепт визуальной глубины.

Чтобы свести стереограмму по методу анкроссконвергенции, начнем с того, что поднесем книжку вплотную к лицу, чтобы нос коснулся страницы. Расслабим глазодвигательные мышцы — то есть, не будем пытаться сфокусироваться на странице — а затем потихонечку начнем отдалять страницу от лица. Мы вновь увидим четыре изображения и четыре точки, но в определенный момент изображений станет три и точек тоже три. Остановимся в этом положении и обратим внимание на среднее изображение, которое стало восприниматься объемным. Отметим, что глубина, которую мы воспринимаем таким образом, обратна той, что мы видели, когда прибегали к методу кроссконвергенции.

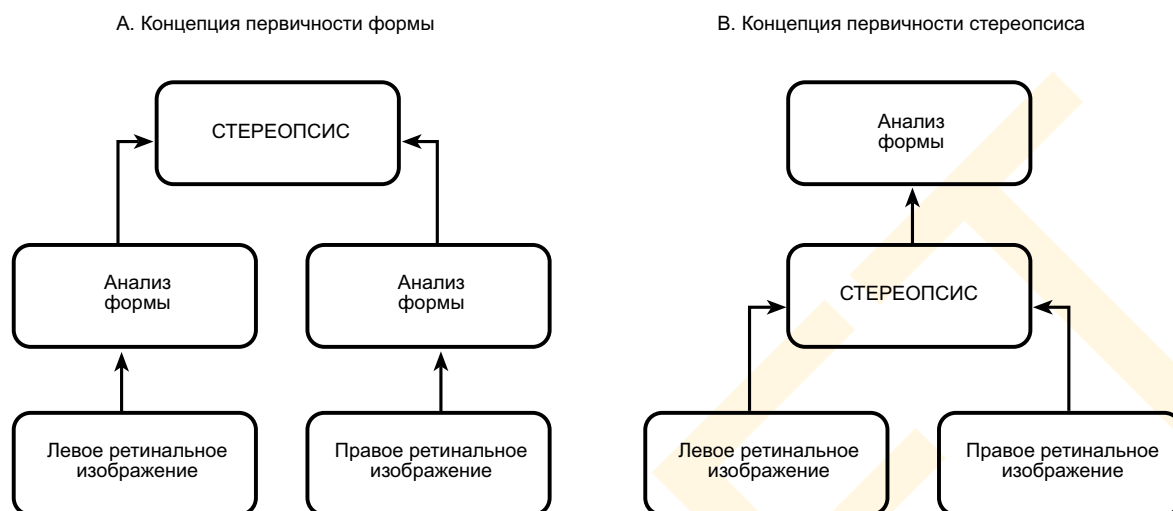
### 5.3.2 Задача согласования

До сего момента мы говорили о стереозрении лишь как о необходимости оценивать меру и направление диспаратности согласованных элементов в ретинальных изображениях. Но мы покамест не касались много более сложных проблем, в частности определения того, какие элементы в одном ретинальном изображении с какими элементами согласуются в другом. Сие именуется *задачей согласования*. Мы избегали этого разговора ранее, поскольку мысленно отслеживали путь световых потоков из внешнего оптического массива (то есть из определенной точки в пространстве и далее) до проекций этой точки на сетчатку левого и правого глаз. Выполняя сие, мы априори исходили из соответствия ретинальных изображений в левом и правом глазу. Но зрительная система сталкивается с куда более сложной задачей — задачей согласования: начинается с двух изображений и ищет, какой элемент в одном изображении какому элементу соответствует в другом. Как именно зрительная система находит такие соответствия, требует объяснения.

Много лет исследователи зрения полагали, что вопрос сей решается особого рода анализом формы, проходящим *до* стереопсиса. На рис. 5.3.7 А дана схема, иллюстрирующая то, какой представлялась им последовательность анализа глубины и формы. Полагали, что вначале, раздельно по правому и левому ретинальным изображениям, анализируется форма, а результаты используются затем в решении вопроса. Объяснение было таким, что хотя и трудно определить, какая точка из сотен тысяч в одном ретинальном изображении согласуется с такой же точкой в другом, но зато легко вывести то, что, скажем, кончик носа немецкой овчарки в одном ретинальном изображении сошелся с кончиком ее же носа в другом. Таким образом, вопрос согласования радикально упрощается, если анализ формы выполнен первым. Но так ли происходит на самом деле? Вполне вероятно (рис. 5.3.7 В), что стереопсис фактически может осуществляться первым и отрабатывать свою функцию, не прибегая к помощи монокулярной информации о форме объекта.

#### 5.3.2.1 Стохастические стереограммы

Известный венгро-американец, основатель Bell Telephone Laboratories — Бела Юлеш, предложил в свое время проверить идею анализа формы как предшественника стереопсиса с помощью т.н. *стохастической стереограммы*. Стохастическая стереограмма — это пара изображений, состоящих из тысяч хаотично разбросанных точек, чье латеральное смещение провоцирует правдоподобную иллюзию глубины, когда рассматривается стереоскопически, то есть, когда первое изображение стимулирует



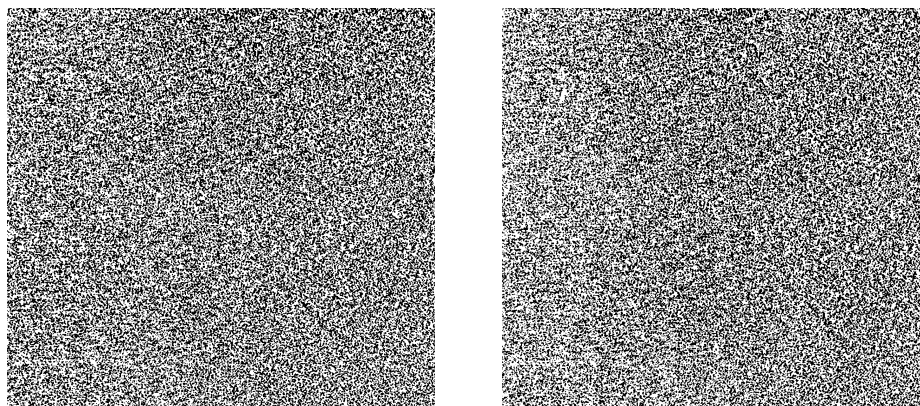
**Рис. 5.3.7** Две концепции стереопсиса. А — концепция первичности формы предполагает, что ретинальные изображения согласуются путем сравнения результатов детекции формы стимулов. В — концепция первичности стереопсиса исходит из того, что стереопсис предшествует анализу формы. Восприятие стохастических стереограмм убедительно свидетельствует в пользу второго варианта.



**Рис. 5.3.8** Стохастическая стереограмма. Оба изображения получены из общего источника — множества случайно (стохастически) разбросанных точек путем латерального смещения определенного участка (см. описание в тексте). Когда эти изображения рассматриваются по методу кросконевергенции, то есть так, что правый глаз смотрит строго на левое изображение, а левый — строго на правое, то в сознании появляется квадрат, как бы висящий над фоном.

сетчатку одного глаза, а второе — другого. На рис. 5.3.8 дан пример такой стереопары, в которой закодирован квадрат, «висящий» над страницей.

Когда каждое из изображений стохастической стереограммы рассматривается по отдельности, точки выглядят случайно (стохастически) расположенными (в том смысле, что в них нет какой-либо информации о форме предмета). Концепция первичности формы в стереоскопическом согласовании прогнозирует то, что восприятие глубины



**Рис. 5.3.9** Стохастическая стереограмма спирали. Если оба изображения слить по методу кроссконвергенции глаз, то «появится» спираль, выходящая из страницы. Ощущение появляется благодаря легкому латеральному смещению тысяч мельчайших точек.

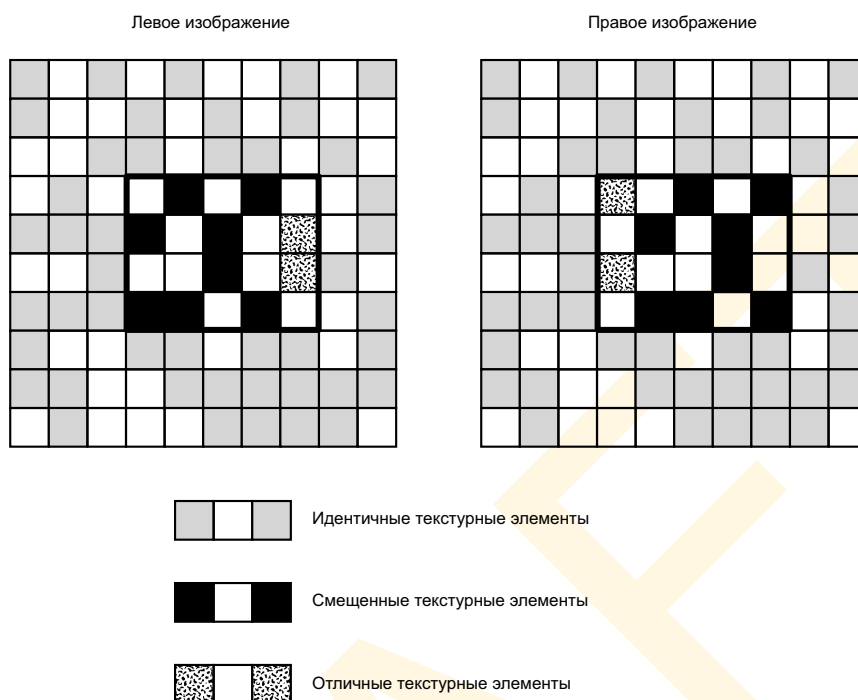
пространства невозможно путем фузии стохастической стереопары, поскольку предполагает, что согласование должно основываться на распознавании монокулярной информации о форме объекта.

Чтобы самостоятельно проверить данный прогноз, прибегнем к кроссконвергенции. Вполне вероятно, что квадрат появится не сразу, поскольку стохастические стереограммы воспринимать тяжелее, нежели обычные. Но когда наконец удастся — зернистый квадрат «повиснет» над зернистым фоном.<sup>1</sup> Итак, поскольку никаких монокулярных форм к сравнению нет в обоих ретинальных изображениях, наш предварительный вывод таков, что гипотеза «первоформы» неверна: очевидно, что зрительная система решает вопрос согласования без монокулярных данных о форме, поскольку юлешевы стохастические стереограммы таких данных в себе не имеют.

На рис. 5.3.9 к нашему любопытству дана еще одна стереограмма. Следует заранее предупредить, что она много труднее к восприятию, чем простой квадрат, но результат стоит усилий: мы, в итоге, увидим спираль, «выходящую» из страницы.

Однако же важно не переусердствовать с выводами. Из того факта, что с помощью стохастической стереограммы мы можем заполучить перцепт глубины, вовсе не следует, что перед стереопсисом зрительная система не анализирует форму. Мы показали лишь то, что перцепт стереоскопической глубины *может* возникнуть без использования зрительной системой монокулярной информации о форме предметов сцены: есть вероятность, что перед стереопсисом имеет место некий примитивный анализ формы или контура, гарантирующий решение вопроса согласования, ежели монокулярные данные о форме представлены в стимуле. Ориентация отрезков и краев — пример информации, которая может стать полезной в решении этого вопроса, то есть: отыскиваются отрезки линий и края одинаковой ориентации и по ним оценивается соответствие. Такая информация становится доступна зрительной системе из выходного сигнала простых клеток V1 на ранних этапах зрительного акта (Hubel & Wiesel, 1962). Фактически, затруднения, которые мы испытываем в получении перцепта глубины при

<sup>1</sup> Если при фузии мы прибегнем к анкроссконвергенции, то увидим квадрат позади фона в квадратном отверстии.



**Рис. 5.3.10** Создание стохастической стереограммы. Левое и правое изображения идентичны (светло-серые и белые квадратики) за исключением латерально смещенного участка и самостоятельных участков (текстурированные и белые квадратики). (В фактических стереограммах текстурные элементы черные и белые — рис. 5.3.8).

рассматривании стохастической стереограммы в сравнении с обычной фотографической стереопарой, указывает на то, что монокулярная информация о форме действительно играет какую-то роль. Но при этом перцепт глубины, инициируемый стохастическими юлешевыми стереограммами — яркое свидетельство того, что даже в отсутствие такой информации — восприятие глубины вполне возможно.

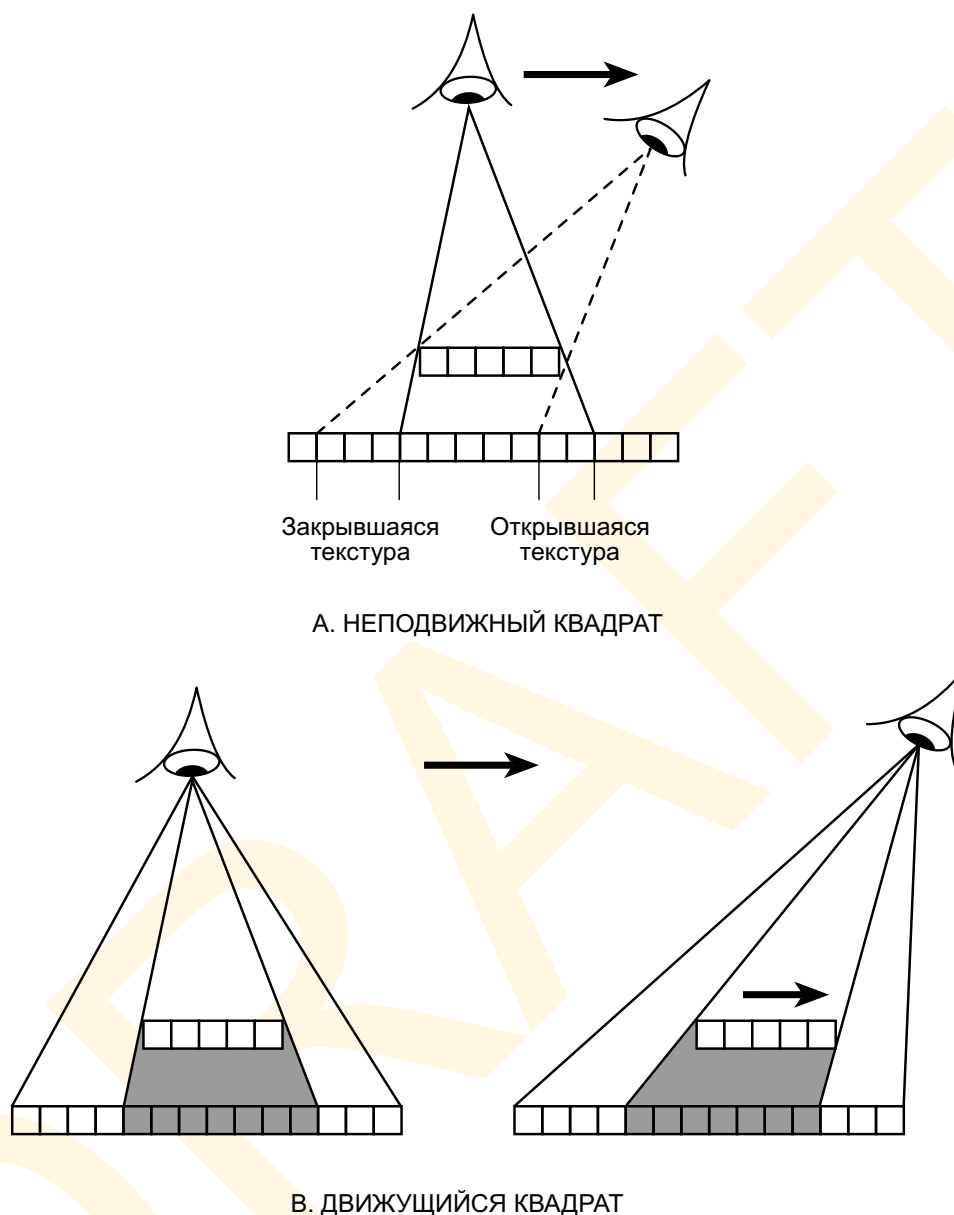
Чтобы понять, за счет чего при восприятии стохастических стереограмм зрительная система оказывается способной к решению задачи согласования, мы в первую очередь должны понять, как они построены. Создание стереограммы — это весьма простая процедура, которую мы можем выполнить самостоятельно. Первое изображение строится простым заполнением случайными точками (можно взять любое из двух изображений рис. 5.3.8). Второе — получается из первого следующим образом:

- первое изображение клонируется;
- из дубля вырезается фигура нужной формы и сдвигается вправо примерно на 0.5 см (см рис. 5.3.10). Величина сдвига определит силу ощущения глубины;
- образовавшееся пустое место заливается случайными точками;
- изображения располагаются сторона к стороне.

Стереограмма готова. Теперь, если рассматривать ее стереоскопически, то возникнет диспаратность по точкам сдвинутой области. От направления сдвига будет зависеть то, как расположится вырезанная фигура — перед фоном или позади него.

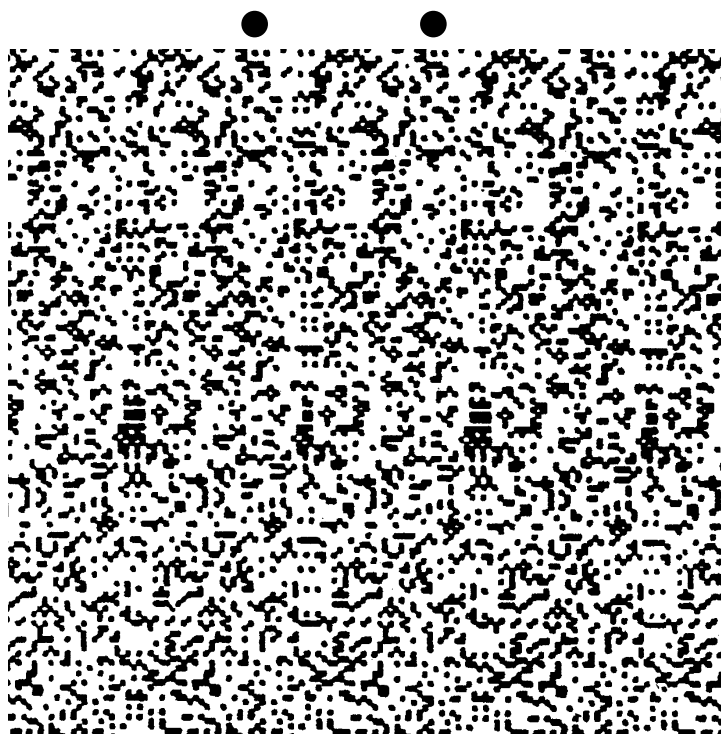
Следующий опыт со стереограммой рис. 5.3.8 позволит нам продемонстрировать еще одно интересное свойство стереоскопии: попробуем немного поводить головой вверх-вниз по плоскости бумаги, когда смотрим на стереограмму, точнее, собственно в





**Рис. 5.3.11** Феномен движения в стохастической стереограмме. Когда наблюдатель, рассматривающий стохастическую стереограмму начинает двигаться, то «ближняя» поверхность тоже начинает движение в направлении фактического движения наблюдателя (В), поскольку нет никаких закрывающихся и открывающихся текстурных элементов (как в реальности, когда ближняя поверхность неподвижна, а движется сам наблюдатель) (А).

период существования иллюзии (это трудно и может потребовать нескольких попыток): квадрат чудным образом начнет смещаться вперед-назад согласно нашим движениям. Сей любопытный феномен есть результат отличия фактической структуры стереограммы от того какой на самом деле должна быть физическая структура реальных поверхностей. К примеру, рассмотрим, что произойдет, если квадрат фактически расположен перед фоном, а наблюдатель при этом смещает голову вправо (рис. 5.3.11 А). В реальности по правому краю квадрата появились бы дополнительные точки, потому



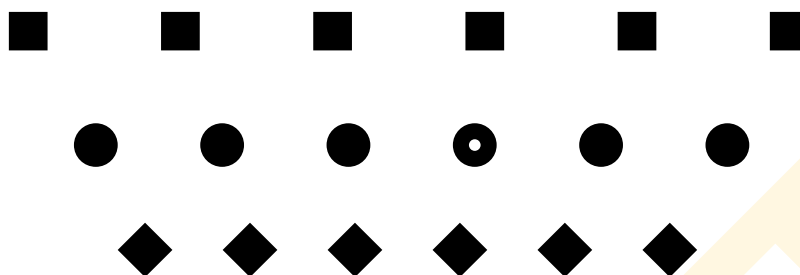
**Рис. 5.3.12** Автостереограмма. Если данное изображение рассматривать согласно указаниям в тексте, то без всяких специальных приспособлений (очков) возникнет иллюзия глубины: над фоном повиснут «шашечки».

что теперь они уже не прикрыты им, а по левому краю часть точек пропала бы, поскольку он закрыл их. Однако же в стереограмме квадрат фактически копланарен окружающему полю, поэтому никаких новых точек не появится и не исчезнет, как ни крути головой. Аналогично, когда квадрат на самом деле находится перед фоном, возникает т.н. *параллакс движения* (см. раздел 5.4.1). Если параллакса не возникает, то стимуляция выглядит несовместимой с восприятием глубины пространства. Почему же тогда квадрат не сливается с фоном, как только мы двинем головой?

Дело в том, что существует альтернативное объяснение тому, что при отсутствии динамической информации, у наблюдателя может возникнуть перцепт глубины. Вообразим, что по каким-то причинам квадрат сместился вправо синхронно с нашей головой (рис. 5.3.11 В): сие полностью ликвидирует эффект движения головы, то есть — никакие новые элементы из под квадрата не появляются, а старые не исчезают — параллакса движения нет. Поскольку такая ситуация вполне возможна и не конфликтует с обычной проксимальной стимуляцией, сие есть альтернативное восприятие, позволяющее нам держать стереоскопическую иллюзию.

### 5.3.2.2 Автостереограммы

Еще одной разновидностью стереограмм (ставших весьма популярными в последние несколько лет потому, что не требуют специальных очков и аппаратов) являются т.н. *автостереограммы*, названные так Тилером и Чангом (1977), но нынче именуемые чаще *стереограммами «магический глаз»*. На рис. 5.3.12 дан пример: при нор-



**Рис. 5.3.13** Упрощенная автостереограмма. Если следовать инструкциям в тесте, возникнет иллюзия глубины: будет казаться, что ряды расположены на разной глубине — квадраты «окажутся» ближе кругов, а ромбы — дальше.

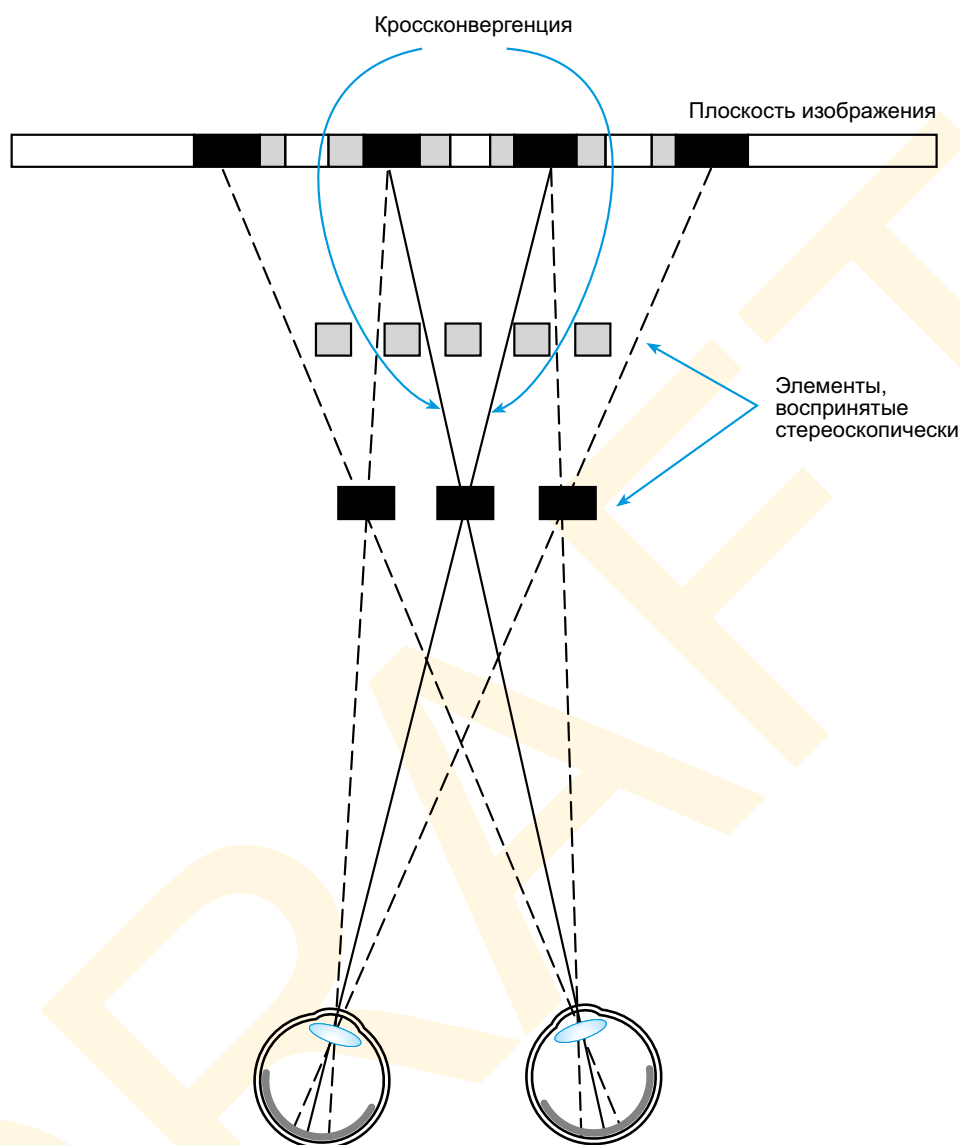
мальном просмотре стереограмма воспринимается как есть, то есть — как плоская картинка с регулярным горизонтальными элементами; при специальном просмотре, однако, появляется иллюзия глубины. Чтобы воспроизвести эту иллюзию, мы должны усилием глазодвигательных мышц свести глаза так, чтобы две точки вверху превратились в три, расположенных по горизонтальной линии. Как только нам это удастся, так сразу возникнет иллюзия того, что группа шахматно расположенных четырехугольников, возвышается над фоном.

Чтобы понять каким образом подобные паттерны вызывают перцепт глубины, разберем намного более простую автостереограмму, показанную на рис. 5.3.13. Давайте посмотрим на нее по методу кроссконвергенции и когда мы увидим два соприкасающихся круга с белыми точками внутри, мы обнаружим, что ряды фигур как бы находятся на разном удалении: квадраты ближе к нам, ромбы — дальше, а круги — по середине. Легко догадаться, что ежели посмотреть на стереограмму по методу анкроссконвергенции, то перцепт глубины окажется обратным.

Иллюзия глубины возникает тогда, когда каждый из глаз фиксируется на своем объекте и перцепты этих объектов сливаются в один общий перцепт.<sup>1</sup> Волевым усилием мы можем заставить наши глаза зафиксироваться на разных объектах. В каждом из рядов фигуры одинаковы по форме, что способствует возникновению т.н. *фузионной ошибки*, исход которой в том, что перцепты объектов фиксации сливаются в один, «расположенный» на плоскости глубины, по которой сведены глаза. Плоскость эта «лежит» перед фактической плоскостью рисунка (рис. 5.3.14). То же происходит с фигурами в других рядах.

В сравнении со сложным компьютерным изображением на рис. 5.3.12 автостереограмма рис. 5.3.13 очень проста, но ее принцип ровно тот же. Разумеется, можно построить куда более замысловатые рисунки на базе более плотных текстур из случайных элементов, более сложного расстава этих элементов и их расцветок (Tyler, 1990). В каждом случае эффект стереоскопической глубины будет зависеть от того, насколько нам удастся обмануть зрение, совместив элементы двух ретинальных изображений, и устроив их горизонтальную диспаратность как в фактическом стимуле с целевыми отношениями по оси глубины.

<sup>1</sup> Мы можем в этом убедиться, потому что фактически в картинке присутствует всего лишь один круг с белой точкой по его центру, но когда иллюзия в ходу, мы видим два таких круга.

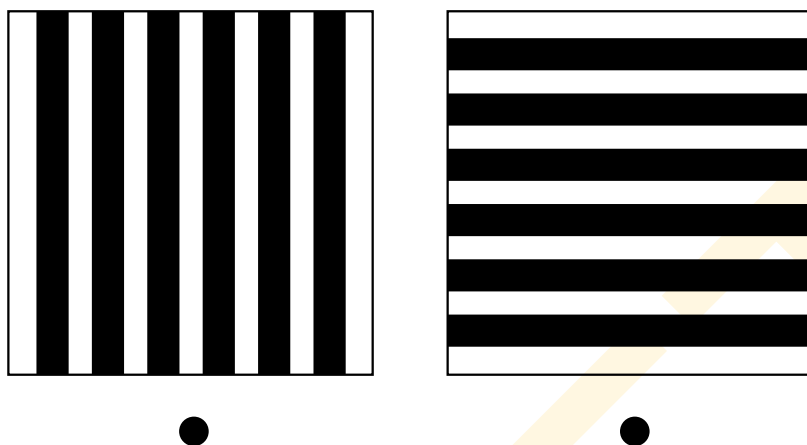


**Рис. 5.3.14** Объяснение механизма автостереограмм. Иллюзия глубины пространства возникает, когда глаза неконвергированы, то есть, когда разные элементы в регулярном паттерне слились (см. пояснения в тексте).

### 5.3.2.3 Бинокулярная конкуренция

До сих пор в нашем разговоре о бинокулярном зрении мы исходили из концепции согласованных точек. Но что происходит, когда ретинальные изображения в правом и левом глазу сильно разнятся, например, когда на одну сетчатку проектируется вертикальная, а на другую горизонтальная решетка?

Сильное несоответствие оптических изображений на левой и правой сетчатках ведет к тому, что в данный момент времени наблюдатель видит лишь что-то одно. Сие феномен, названный *бинокулярной конкуренцией* или *бинокулярной супрессией* (Levelt, 1968). По стимулам малых угловых размеров — от  $1^\circ$  и менее — конкуренция стремится к абсолютной — перцепты перемежаются с периодом в 1 – 4 с. По изображениям превышаю-



**Рис. 5.3.15** Бинокулярная конкуренция. Если слить оба изображения (по одному из методов), станет ощущаться конкуренция между вертикальными и горизонтальными штрихами: в одном и том же участке поля зрения штрихи будут выглядеть то вертикальными, то горизонтальными.

щим  $1^\circ$  конкуренция становится, если угодно фрагментарной, кусочковой — мы одновременно видим разные участки обоих изображений. Испытать на себе проявление бинокулярной конкуренции можно с помощью рис. 5.3.15 по кросс- и анкроссметоду: фузия этих двух решеток в единый стабильный перцепт совершенно невозможна, как ни старайся, но при этом если держать конвергенцию какое-то время, то можно наблюдать, как ощущения меняются в разных участках квадрата и в разное время. Временная изменчивость перцепта, судя по всему, возникает из-за того, что нейроны отвечающие на одну стимуляцию — утомляются после продолжительной пульсации, в результате чего на авансцену выходит сигнал от неустоленных нейронов.

Бинокулярная супрессия — это скорее неестественный феномен, поскольку глаза наши обычно нацелены на одну и ту же сцену, в которой все соответствия веридикальны. Однако довольно часто случается так, что нормальные условия дают различные изображения в согласованных точках обеих сетчаток, в частности, когда между согласованными точками возникает выраженная диспаратность. Поскольку зрение наше крайне редко стимулируется такими изображениями, причиной может стать бинокулярная супрессия, ведущая к появлению единичного перцепта, когда, казалось бы, он должен быть двойным.

### 5.3.3 Количественные подходы

Давайте вернемся к разговору о том, как зрительная система решает вопрос согласования в стохастических стереограммах, когда в ее распоряжении нет глобальной информации о форме. Чтобы прочувствовать сложность проблемы рассмотрим детально, что нужно сделать. Большинство точек (но не все) благодаря общей локализации в сцене в левом изображении имеют соответствующую согласованную точку в правом изображении. Зрительной системе нужно каким-то образом расшифровать, какие пары идут вместе. Чтобы хоть как-то упростить задачу, рассмотрим только те пары точек, что лежат вдоль линии бинокулярного сдвига, когда голова наблюдателя в вертикальном положении. Если в одном изображении по данной горизонтали 100 точек, то количество



логически возможных вариантов образования пар равно  $100!$ <sup>1</sup>. Это воистину большое число и, заметим, что оно относится только к одной из сотен горизонтальных рядов.<sup>2</sup>

Существует множество количественных подходов к решению задачи согласования: одни исследователи рассматривают отдельные точки (пиксели), другие — края и линии, третьи пытаются согласовывать целые фрагменты. Мы рассмотрим отдельные небольшие примеры каждого из подходов.

### 5.3.3.1 Первый алгоритм Марра — Поджио

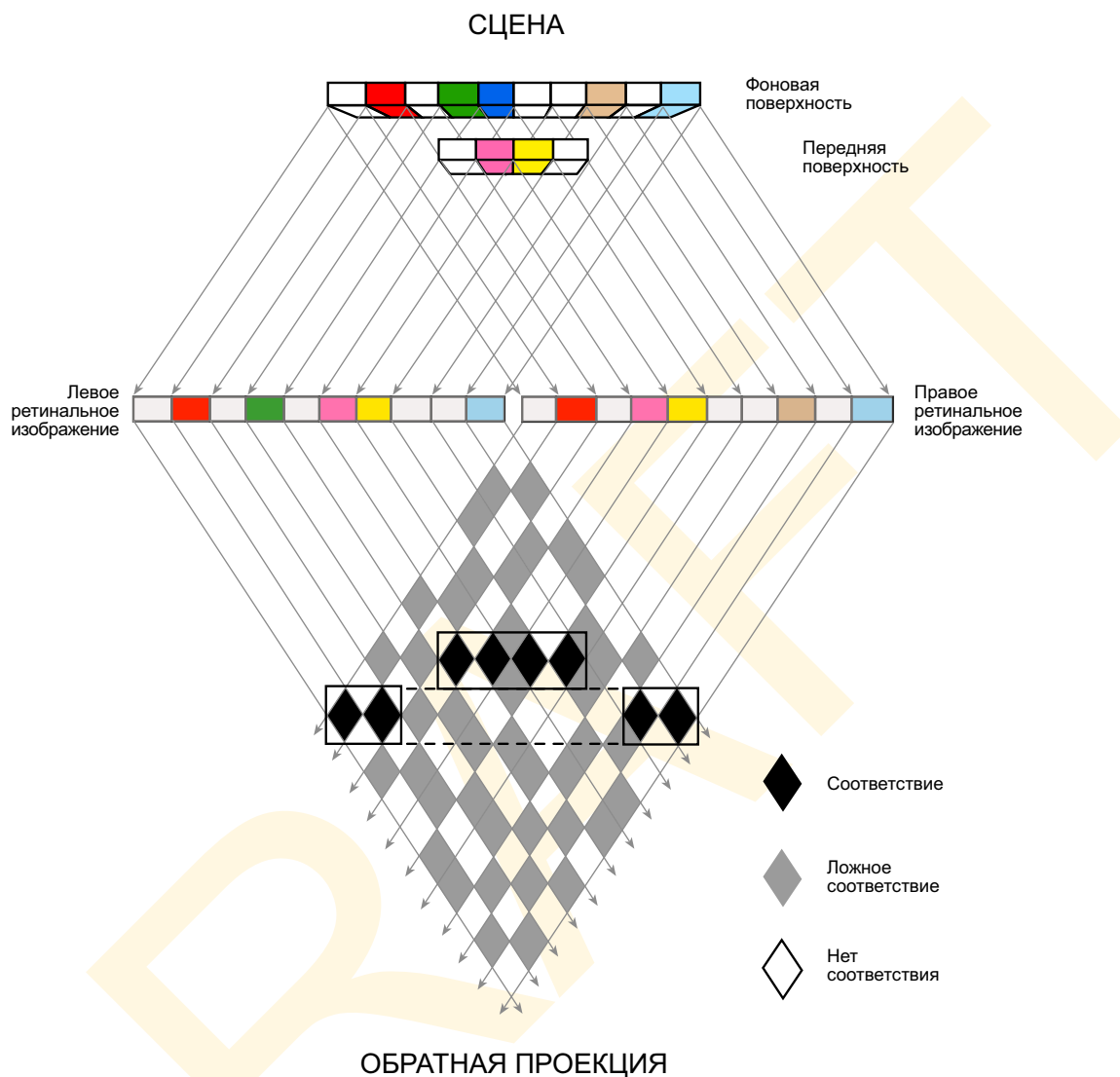
Одним из хорошо известных и весьма интересных алгоритмов является схема, предложенная в 1977 г. Дэвидом Марром и Томазо Поджио (Массачусетский технологический институт). Сие далеко не новая и не самая лучшая концепция стереопсиса, но это весьма любопытный и исторически важный пример того, как динамическая нейронная сеть может количественно решить сложную зрительную задачу. Это также отличная демонстрация того, как эвристические допущения или ограничения можно внедрить в такую сеть. Поэтому рассмотрим вопрос детально.

Идея, на которую опирается алгоритм Марра — Поджио (1977) — это согласование позиций путем уравнивания отдельных пикселей в левом и правом ретинальных изображениях. Отправной точкой в нашем понимании этого подхода является концепция обратной проекции из обоих ретинальных изображений назад — в сцену: рис. 5.3.16. Схема рис. 5.3.16 демонстрирует взгляд сверху на две полосатых поверхности в сцене, одна (меньшая) расположена перед другой (большой) — см. вверху иллюстрации. Свет, отраженный этими поверхностями, образует оптические изображения на левой и правой сетчатках, которые, для простоты, изображены плоскими (в центре схемы). Каждый пиксел этих изображений спроектирован обратно в «зеркальный мир», показанный внизу диаграммы. Закрашенные клетки в получившейся матрице представляют позиции, по которым имеется яркостное соответствие между пикселями изображений: точки соответствуют друг другу (закрашены), если оба пикселя черные (либо белые) и не соответствуют (не закрашены) если одна точка черная, а вторая белая (и наоборот). Если дать проекцию от закрашенных клеток назад к пикселям левого и правого ретинальных изображений, то мы увидим, что у обоих яркость одинакова. Среди множества соответствий найдутся такие, что относятся к видимым участкам фактических поверхностей сцены. Эти пары на рис. 5.3.16 закрашены темнее, дабы отличать их от лжепар, которые тоже имеют место.<sup>3</sup>

<sup>1</sup> Факториал сотни —  $100 \times 99 \times 98 \times \dots \times 2 \times 1$ .

<sup>2</sup> Число это вычислено следующим образом. Поставим точку в одном из рядов правого изображения и назовем ее «А». Если в согласованном ряду левого изображения 100 точек, то имеется  $100!$  возможных вариантов образования пары точки А с одной из точек в левом изображении. Теперь поставим еще одну точку в правом изображении и назовем ее «В». Существуют 99 вариантов образования точкой В пары с одной из оставшихся точек в левом изображении, поскольку одна из точек «ангажирована» точкой А. То есть, существует  $100 \times 99$  образования пар точками А и В в правом изображении с точками того же ряда в левом. Если мы поставим третью точку — С, то она может образовать пару в левом с любой из оставшихся 98 точек, что дает  $100 \times 99 \times 98$  вариантов образования пар точками А, В и С. По всем 100 точкам правого изображения получаем:  $100 \times 99 \times 98 \times \dots \times 2 \times 1$ .

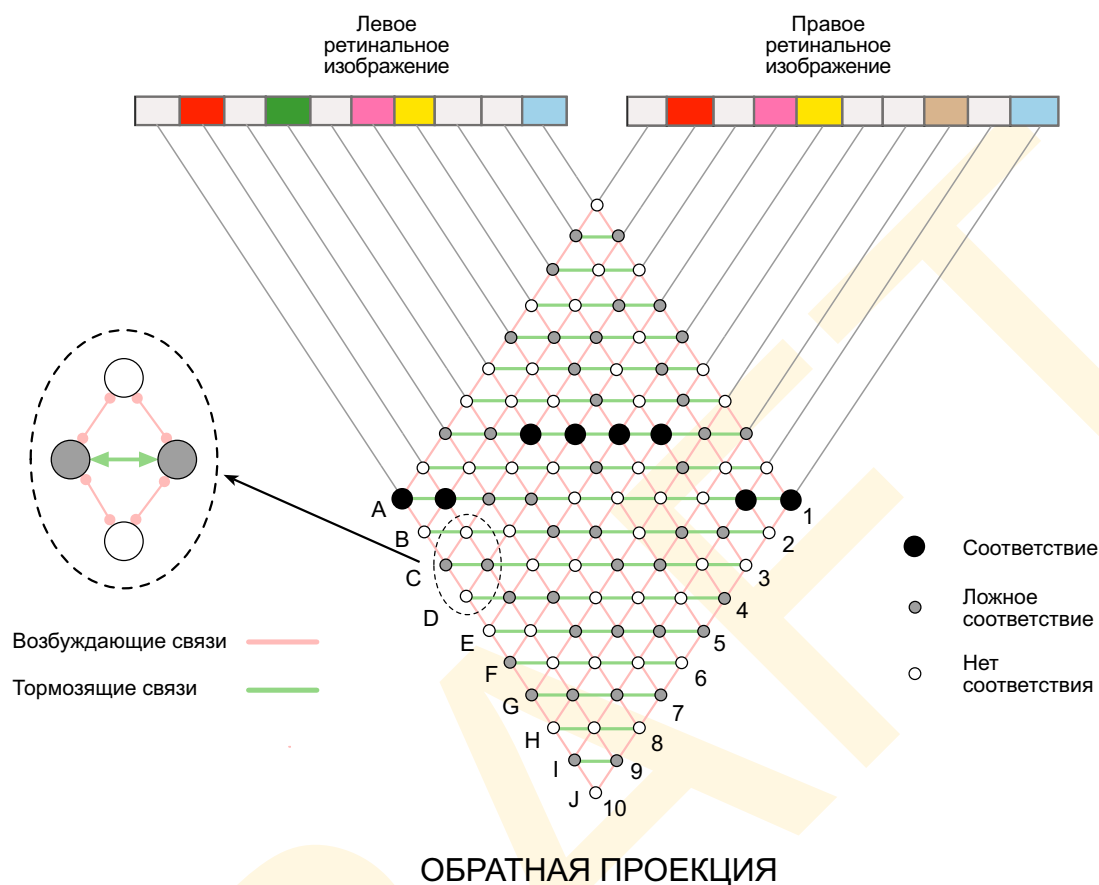
<sup>3</sup> Отметим, что лжепары возникают по дальним поверхностям, частично прикрытым ближними. Сие происходит из-за того, что ни один из глаз не получает оптической информации от закрытых точек. По этой причине Марр ограничил свой 2.5-скетч лишь *видимыми* участками поверхностей.



**Рис. 5.3.16** Обратная проекция двух поверхностей. В верхней части иллюстрации показан вид сверху двух поверхностей, дающих проекции на левую и правую сетчатки. Внизу показано, как ретинальные изображения можно использовать для образования обратной проекции, в которой множество яркостных соответствий, часть из которых согласуется с веным восприятием глубины (темные элементы) и множество тех, что дают ложные соответствия (серые элементы).

Тот факт, что наряду с истинными существуют лжепары, свидетельствует о том, что имеет место недорешенность вопроса инверсии, и решить его можно разными способами. Главная задача зрительной системы — определить, какие соответствия корректны, какие нет.

Марр и Поджио предположили, что сие выполнимо с помощью т.н. *динамической нейронной сети*, показанной на рис. 5.3.17 и на которой даны левое и правое изображения с рис. 5.3.16. Ретинальные изображения, активирующие внутренние узелки в нейронной сети, что представляет множество всех возможных соответствий. То есть, каждый узелок представляет потенциальное равенство пикселей: один в левом изобра-



**Рис. 5.3.17** Первый стереоалгоритм Марра — Поджио. Нейронная сеть выполняет обратную проекцию двух ретинальных изображений, активирующих узелки, согласованные с позициями, в которых левое и правое изображения закрашены одинаково. Ингибиторные (тормозящие) связи (см. вставку) вводят условие непрозрачности, а возбуждающие — непрерывности (подробнее см. в тексте).

жении и один в правом. Только те пересечения, что образовались от пикселей одной яркости — это возможные соответствия, поскольку пиксели, суть проекция от одной и той же точки в сцене и, разумеется, должны иметь одну яркость (оба черные или оба белые). Следовательно, эти соответствия, отмеченные серыми узелками, составляют наибольшее ограничение в решении вопроса соответствия, поскольку все белые узелки можно исключить из рассмотрения.

Однако уравнивание по яркости — это не единственно возможное решение проблемы согласования, поскольку таких равенств может быть множество по каждой точке левого и правого ретинальных изображений. И как раз по этой причине стохастическая стереограмма создает столько трудностей для ее восприятия зрительной системой. Вопрос в том, какие эвристические условия следует ввести в формулировку проблемы, дабы найти, наконец, однозначное, корректное решение на все случаи? Марр и Поджио (1977) предложили два таких условия:

1. *Непрозрачность поверхностей.* Условие непрозрачности позиционировано потому, что большинство поверхностей реального мира непрозрачны, и видимыми могут быть только ближайшие к нам. То есть, если соответствие A10 на рис. 5.3.17 истинно

(то есть, если пиксел А в правом изображении фактически согласуется с пикселом 10 в левом изображении), то соответствия В10, С10, D10 и т.д. не могут быть истинными, а также А9, А8, А7 и т.д.

*2. Непрерывность поверхностей.* Условие непрерывности означает, что поскольку поверхности в реальном мире стремятся к локальной непрерывности в глубину (за исключением кроющих краев), корректное решение стремится быть одним, при котором соответствия «уплотняются» в глубину, т.е. какими они и должны быть на локально непрерывных поверхностях.

Отметим, что оба условия — это эвристические допущения, которые обычно истинны, но все же не всегда. Если они истинные — алгоритм Марра — Поджио работает. Если они ложны, то есть, если одна или несколько поверхностей прозрачны и/или если поверхности локально прерывисты — решение будет в той или иной мере некорректным.

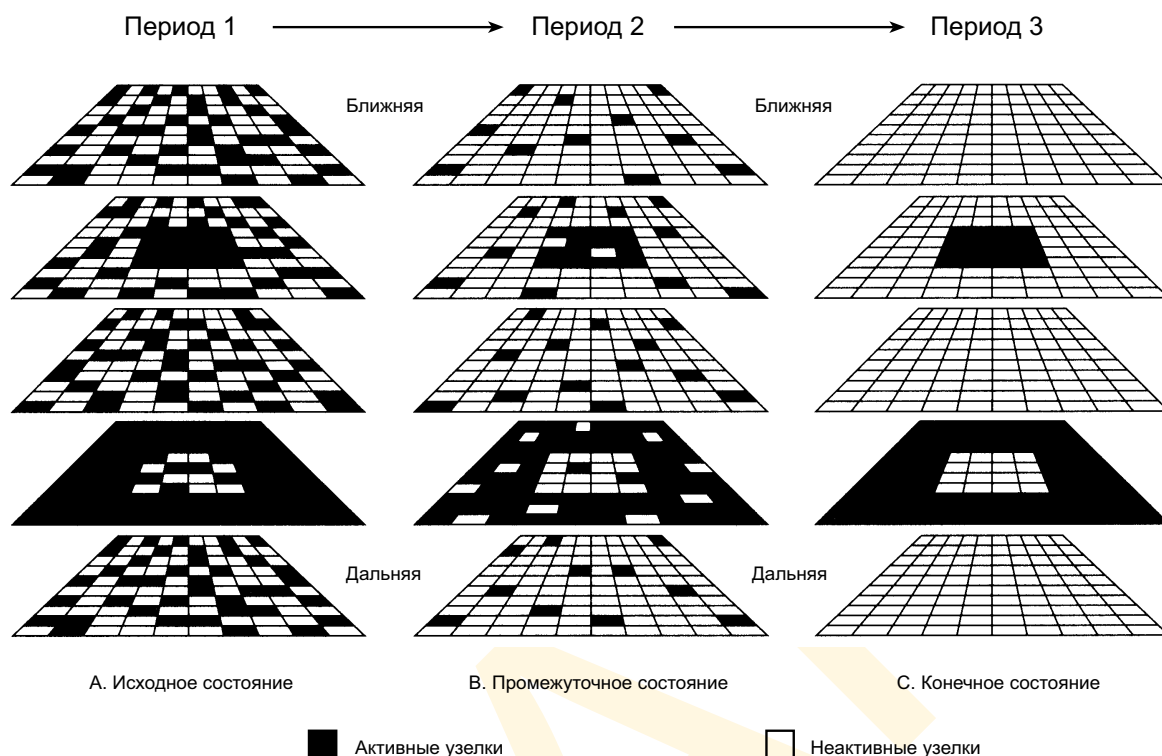
Марр и Поджио заложили свои условия в межузелковые соединения нейронной сети рис. 5.3.17. Модель вначале активировывает все узелки в «перекрестной сети», что представляет пиксели сходной яркости в правом и левом изображениях. Это будут только серые узелки на рис. 5.3.17, и это значит, что они были активированы в начальной фазе алгоритма. Сие множество возможных соответствий затем сокращается по критерию непрозрачности и неразрывности и происходит это за счет характера связей в сетке, показанной на увеличенной четырехузелковой вставке в левой части рисунка. Эта часть архитектуры, отмеченная диагональными соединениями на рис. 5.3.17 именуется сетью класса «Победителю достается всё», поскольку только один узелок в каждой диагонали остается работающим после активации, имевшей место в устойчивом состоянии системы.<sup>1</sup>

Условие по непрерывности заложено в сеть за счет взаимного возбуждения пикселей в плоскостях сходной удаленности от наблюдателя (или одной и той же). Эти взаимодействия между узелками отмечены на рис. 5.3.17 толстыми отрезками. Их эффект, состоит в том, что возможные соответствия на одном и том же удалении от наблюдателя — т.е. на одной и той же глубине — стремятся «поддерживать друг друга», активируя один другого посредством межвзаимодействий. Когда принимается условия по непрерывности, решение смещается в сторону копланарности.

Совокупный эффект обоих условий прогрессивно сокращает множество возможных равенств до единичного множества наиболее совместимых с (1) диспаратной информацией, извлекаемой из ретинальных изображений, (2) условием непрозрачности и (3) условием непрерывности. Сеть работает и в обратном направлении через возбуждающие и тормозящие связи, активируя и тормозя узелки до тех пор, пока наконец система не перейдет в устойчивое состояние.

Сие финальное состояние сети обычно (но не всегда) согласуется с фактическим положением дел в сцене. На рис. 5.3.18 показано, как активация в перекрестной сети меняется со временем в простейшем случае, когда квадратная фигура расположена перед фоном (как на стереограмме рис. 5.3.8). Исходный паттерн активации полностью оп-

<sup>1</sup> Нейронные сети класса «Победителю достается всё» дают такой эффект, поскольку любой узелок, имеющий наибольшую активацию, будет подавлять другие узелки больше, чем они его. Если взаимодействия между узелками в сети непрерывны во времени, взаимное подавление вынуждает этот перекося в активации усиливаться — большинство активных узелков становятся еще активнее, а прочие — менее активными — до тех пор пока не будет достигнуто устойчивое состояние, при котором только один узелок активен.



**Рис. 5.3.18** Динамическое поведение стереосети Марра — Поджио. Активные узелки соответствуют различным удаленностям плоскостей, представленных черными квадратиками. А — исходно все позиции, совпадающие друг с другом по яркости  $t$  (т.е. и верные, и ложные соответствия) — активированы. В — появление возбуждающих и тормозящих связей, среди которых есть верные и ложные. Последние минимизированы. С — дальнейшие взаимодействия существенно среди верных связей.

ределен множеством пар пикселей по яркостному соответствию. Как только узелки сети начинают межвзаимодействовать через связи активации и торможения, условия непрозрачности и непрерывности начинают давать эффект. Наконец, исходные соответствия редуцируются до множества, представляющего квадрат перед фоном.

Первый алгоритм Марра — Поджио — это весьма любопытный пример того, как процесс бессознательной инференции внедряется в нейронную сеть. Логически неразрешимая задача инверсии оказывается решаемой путем эвристических допущений, дающих верные решения, когда допущения эти истинны. «Задача» в данном случае — это поиск корректных соответствий в стохастической стереограмме. Это есть «решение» через точные «предположения» о структуре поверхностей в сцене, такие, например, как непрозрачность и непрерывность. Сии предположения мы назовем «эвристическими», потому что они лишь вероятно (нежели наверняка) адекватны предметной ситуации в реальной сцене. Таким образом, алгоритм Марра — Поджио, это пример того, как бессознательная инференция реализует восприятие без привлечения дедукций, основанных либо на количественных преобразованиях, либо последовательном применении правил символической логики. Вместо того, допущения влияют на силу связей внутри нейронной сети и работают они, оказывая благоприятный эффект на поведение этой сети (в смысле достижения веридикальности восприятия).



Стереоалгоритм Марра — Поджио (1977) — это также интересный пример работы динамической нейронной сети, как физического гештальта. Вспомним, что Кёхлер говорил о физических гештальтах как о динамических физических системах, конвергирующих с минимальным энергетическим состоянием (см. раздел 2.1.2). В случае мозга воспринимающего, полагали гештальтисты, минимум энергии отражен в принципе Прягнанца: перцепт должен оказаться настолько «точен» и «прост», насколько позволяют внешние обстоятельства. Нейронная сеть Марра — Поджио аналогичным образом конвергирует к состоянию энергетического минимума (Hopfield, 1982). Более того, решение берется из неопределенных условий стимуляции («превалирующие условия») в связке с дальнейшими ограничениями, заложенными в нейронную сеть (например, что поверхности непрозрачны и непрерывны). Тот аргумент, что непрерывные поверхности проще чем прерывистые — один из самых очевидных. То, что непрозрачные поверхности проще прозрачных доказывается тем, что проще их оптические свойства.

### 5.3.3.2 Краевой алгоритм Марра — Поджио

Спустя несколько лет после публикации своего первого алгоритма Марр и Поджио (1979) придумали второй вариант алгоритма, отличный от первого по ряду важных параметров:

1. *Краевое соответствие.* Второй алгоритм Марра — Поджио ловит стереоскопические соответствия не по отдельным пикселям, но путем поиска краевых равенств в правом и левом ретинальном изображениях. Что намного результативнее, поскольку позволяет пользоваться информацией недоступной при попиксельном уравнивании, в частности такой как ориентация краев и их полярность (светло-темные или темно-светлые). Края, которые не согласуются по ориентации и полярности мы можем исключить из рассмотрения, добавив тем самым в решение вопроса еще одно упрощение.

2. *Мультишкалы.* Второй алгоритм задействует пространственно-частотные каналы зрительной системы (см. гл. 4), проверяя соответствие краев на низких пространственных частотах (крупная шкала) и только после этого — на высоких (тонкая шкала). В логике стохастической стереограммы сие значит, что на низких частотах работают не с отдельными точками (как это происходило в первом алгоритме), но с крупными областями изображения. Только после того, как на данном общем уровне все возможные краевые соответствия будут найдены, ищутся тонкие соответствия на уровне высоких пространственных частот, т.е. отдельных точечных контуров.

3. *Однопроходная обработка.* Мы говорим, что краевой алгоритм — это неитеративный процесс. Сие значит, что конвергенция достигается с первого раза и в нескольких попытках нет необходимости (как в первом алгоритме). Вместо этого по нескольким этапам (но при этом в один проход) ищется наилучшее краевое соответствие. Поскольку итерации (попытки) требуют времени, машинное исполнение алгоритма оказывается намного «шустрее» своего предшественника.

Уравнивание по краям (на базе выходного сигнала край-детекторов) в новом алгоритме дало очевидные преимущества перед уравниванием по отдельным пикселям. Сие очевидно правдоподобнее биологически, поскольку бинокулярная обработка начинается в V1, т.е. после того, как выходной сигнал отдельных рецепторов объединится в сложные протяженные рецептивные поля коры. Если, как полагал Марр, клетки эти действительно отвечают за краевую детекцию, то решение проблемы согласования

сверткой краев вполне обоснованный подход. Он также дает ряд количественных преимуществ, поскольку детекторы краев несут информацию о более сложных компонентах ретинального изображения (то есть, ориентированных краях данной полярности), нежели отдельные рецепторы, и сложность эта позволяет исключить многие потенциально лишние соответствия.

Результаты работы второго алгоритма Марра — Поджио оказались ближе к результатам психофизических экспериментов с людьми (Marr, 1982). То есть, мы говорим, что второй алгоритм Марра — Поджио — это очень хорошая модель стереозрения человека. Однако же, к сожалению, есть ряд моментов, которые ни один алгоритм, в том числе и краевой, не в состоянии истолковать.

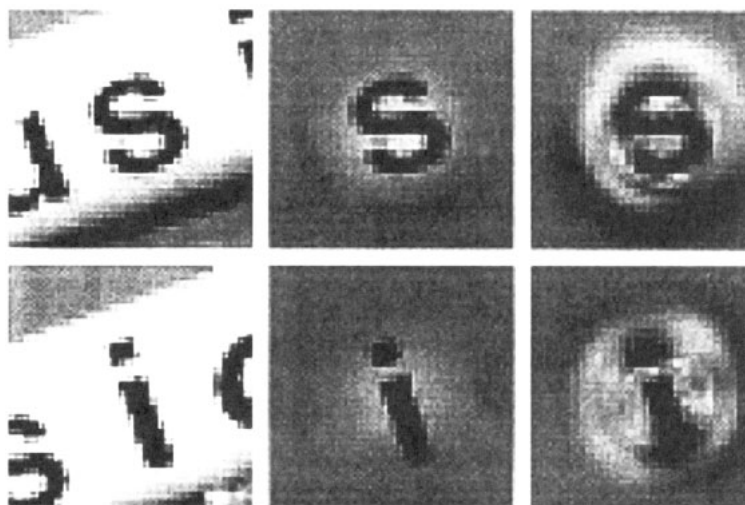
Одна из наиболее серьезных проблем состоит в том, что результирующий вывод о глубине пространства весьма разрежен, поскольку краевой алгоритм описывает соответствия только на краях. Пиксельный алгоритм лишен этого недочета, потому как описывает плотное попиксельное соответствие в изображениях. Во втором алгоритме Марра — Поджио информация о глубине залегания поверхностей между краями заполняется путем интерполяции, что априори неточно.

### 5.3.3.3 Фильтрующий алгоритм

Много позже Дэйвид Джонс и Джинтендра Малик (Беркли, 1992) предложили алгоритм, в котором, фактически, выполнялись уравнивание по областям, лежащим вокруг интересующей нас точки. Это не есть непосредственное уравнивание по локальным областям (что было бы с количественных позиций неоправданно затратно), но скорее согласование выходного сигнала множества биологических пространственных фильтров, разнящихся по пространственно-ориентационной настройке (см. раздел 4.2). Вспомним, что результаты психофизических и физиологических экспериментов жестко указывают на то, что зрительная система выполняет нечто сходное с локальным пространственно-частотным анализом, при котором клетки селективно откликаются на пространственные решетки определенных размеров и ориентаций (De Valois & De Valois, 1988). Джонс и Малик показали, что выходной сигнал таких *мультиориентационных мультишкальных фильтров (МОМШ)* фактически дает намного более прочный базис стереосогласования, чем отдельные точки или ориентированные края. Последние не столь четко согласуются с характером рецептивных полей, работающих по типу габоровой функции, вейвлет-функции и пр. (определяется вариациями в пространственной позиции, ориентации и пространственном размере стимуляции).

Отправной точкой концепции Джонса и Малик явился тезис о том, что в одной, отдельно взятой, гиперколонке существует множество клеток, чьи рецептивные поля центрованы по одним и тем же ретинальным позициям и что каждая клетка настроена на различные ориентации и пространственные размеры стимуляции. Паттерн активности в такой гиперколонке можно смоделировать математически в виде вектора: упорядоченного множества значений, соответствующих частотам пульсации каждого из нейронов. Этот вектор несет на удивление полную информацию о пространственной структуре внутри области, центрованной по данной позиции. На рис. 5.3.19 даны два примера, центры которых — это середины левых изображений. Правые изображения — плоды реконструкции пространственной информации из вектора, представляющего выходной сигнал 60 клеток гипотетической гиперколонки. Мы видим, что клетки эти захватывают значительную часть информации о пространственной структуре изображения в околоцентральных областях.

Ядром МОМШ-модели Джонса и Малик является процесс согласования вектора,



**Рис. 5.3.19** Реконструкция изображения с помощью МОМШ-фильтра Джонса — Малик. Изображения в крайней правой колонке — плод реконструкции из откликов мультиориентационных мультишкальных пространственных фильтров, позиционированных по центру изображений в крайней левой колонке. Результаты справа демонстрируют то, что выходной сигнал таких фильтров несет в себе весьма важную локальную пространственную информацию. Исходные изображения, маскированные по Гауссу, даны в центре.

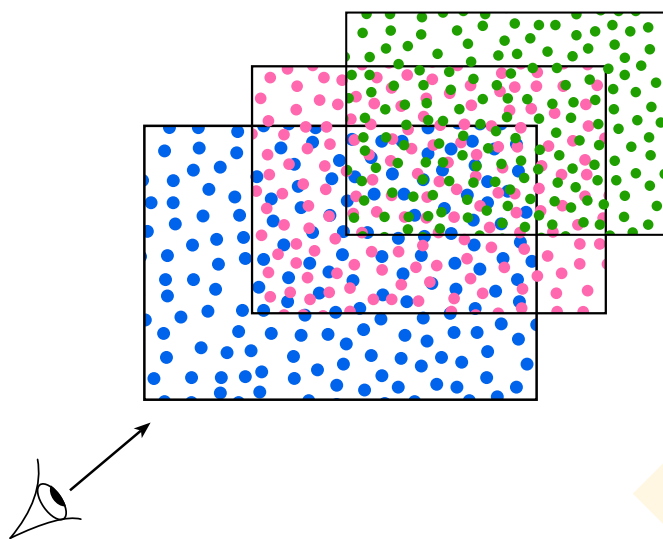
По материалам Jones & Malik, 1992.

представляющего данную позицию в одной сетчатке с каждым из векторов, представляющих латерально смещенные позиции в другой сетчатке. Математические детали этого сравнения нам не интересны, а значимым является то, что латерально смещенная позиция во второй сетчатке, вектор которой более всего схож с таковым в первой сетчатке, описывает наибольшее соответствие по этой позиции. Когда аналогичное сравнение выполняется по всем позициям на всех возможных смещениях, результат дает четкое решение задачи согласования.

Отметим, что соответствия, которые отыскивает алгоритм Джонса — Малик намного точнее и добротнее, чем найденные по точечному или краевому алгоритмам, поскольку МОМШ-векторы несут намного более важную пространственную информацию о локальной структуре изображения, чем выходной сигнал единичных рецепторов или краевых детекторов. То есть, он дает гораздо более селективные и точные согласованные пары.

Подобно обоим алгоритмам Марра — Поджио концепция Джонса — Малик использует нечто подобное условию непрерывности. В результате, как и все рассмотренные нами концепции, идея Джонса — Малик спотыкается о стохастические стереограммы.

Однако же зрение человека *может* решить проблему согласования при разглядывании стереограмм. Американский исследователь К. Праздный (1984) построил стереограммы, изображающие полупрозрачные поверхности, как бы расположенные одна над другой и заполненные стохастическими точками (рис. 5.3.20). Обратим особое внимание на то, что в данном примере не соблюдаются условия непрозрачности и непрерывности — «поверхности» дискретны, большей частью прозрачны и соседние проекции на сетчатку могут дать точки, расположенные на разных удалениях по глубине. Однако зрительная система человека в состоянии восстанавливать зазор по глу-



**Рис. 5.3.20** Стереограмма Праздного. Зрение человека способно «разбираться» в стохастических стереограммах аналогичных показной на данном рисунке, то есть с поверхностями, покрытыми рыхло рассеянными точками и, следовательно, не отвечающими условию непрозрачности и непрерывности.

каких-то еще элементов ретинальных изображений не способен истолковать то, как можно получить стереопсис от смесевых пар. Такие пары требуют поэлементного согласования на неких более абстрактных уровнях, на которых различные типы краев представлены независимо от специфики элемента изображения, в котором они находятся. Сие не значит, что задача согласования решается только на этом абстрактном уровне, но что описанные выше схемы решения нуждаются в дополнении некими последующими механизмами.

### 5.3.4 Физиологические механизмы

Тот факт, что зрение много эффективнее, чем работа любой искусственной схемы стереопсиса, неизбежно подводит нас к вопросу о том, каким же образом мозг решает эту задачу? К сожалению, на том уровне, на котором мы сегодня рассматриваем количественные модели стереопсиса — ответ неизвестен. Но есть кое-какие любопытные факты.

Во-первых, когда Хьюбел и Визель проводили экстраклеточные записи в V1 кошек, то обнаружили бинокулярно сенситивные клетки (1962). Т.е. ученые нашли нейроны, которые сильнее откликаются, когда правая и левая сетчатки одновременно стимулируются одним и тем же штрихом или краем, нежели когда это происходит поочередно — вначале одна, потом другая. Однако клетки, обнаруженные в V1, очевидно настроены на свойства стимула, воздействующего на согласованные области сетчатки, нежели на бинокулярно диспаратные. То есть, несмотря на то, что клетки бинокулярны, они, вероятнее всего, не реагируют на диспаратную стимуляцию.

Несколько лет спустя британцы Хорейс Барлоу, Колин Блэйкмо и Джон Петтигрю

бине между точками на этих поверхностях. Но до тех пор, пока расстояние в глубину между ними достаточно.

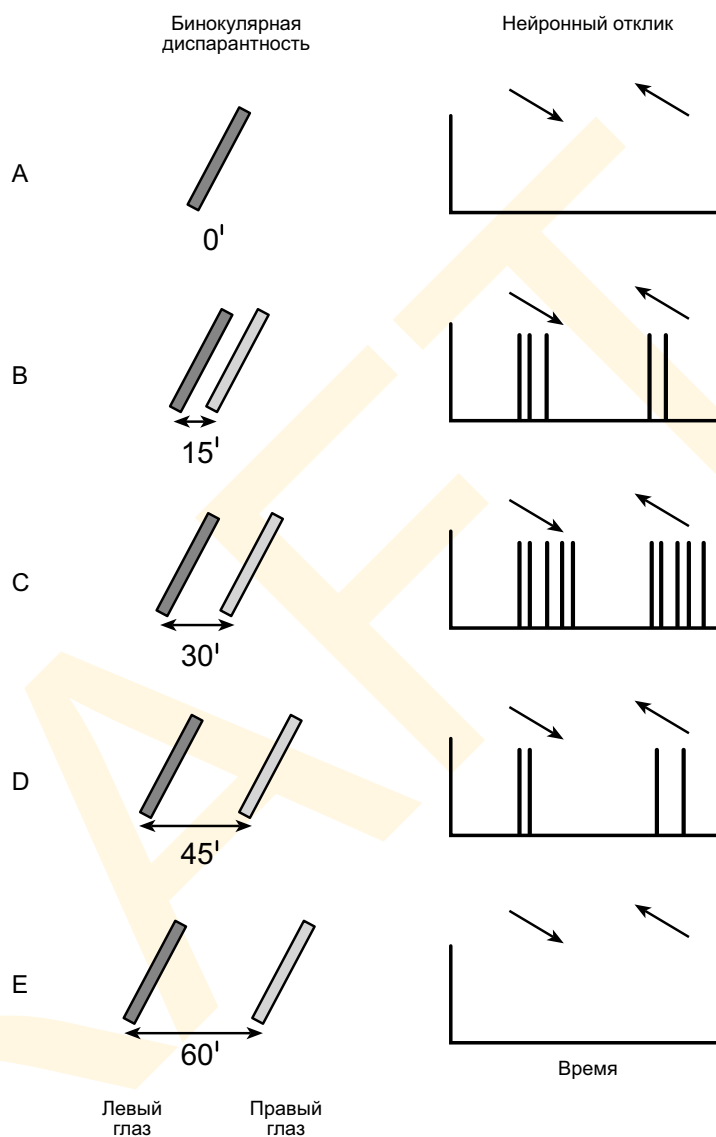
Еще одним интересным фактом, создающим трудности в обработке гипотез о решении задачи согласования является то, что зрительная система способна добиваться стереоскопического восприятия посредством фузии яркостных границ на одной сетчатке с текстурными или хроматическими границами на второй (Ramachandran, Rao & Vidyasagar, 1973). Сие наводит на размышления, что механизмы решения проблемы стереосогласования находятся на более высоких уровнях, чем предполагалось. Ни один алгоритм, основанный на согласовании яркостей пикселей, краев, локальных участков и или

(1967), тоже изучавшие этот вопрос, доложили, что некоторые бинокулярные клетки в V1 дают пик отклика на стимуляцию диспаратных локалей сетчатки. Результаты исследований Хьюбела и Визеля (1970), проводившихся практически в это же время на обезьянах, поддержали версию о том, что клетки V1 чувствительны к стимулам нулевой или околонулевой диспаратности, то есть к стимуляции в согласованных или почти согласованных локалях сетчатки. Они также представили доказательства того, что в V2 имеются клетки, дающие пик отклика на присутствие одних и тех же элементов ретинальных изображений в локалях с относительно высокой диспаратностью.

На рис. 5.3.2.1 дан типичный отклик нейрона V2 макаки, чувствительного к бинокулярно диспаратной стимуляции. Слева даны относительные позиции и ориентации двух стимулирующих штрихов — один на левой сетчатке, второй — на правой. Справа — отклик клетки на движение штрихов в противоположных направлениях (на что указывают стрелки). Рис. 5.3.21 С иллюстрирует то, что клетка эта дает максимальный отклик при диспаратной стимуляции порядка 30 минут зрительного угла.

Прочие измерения показали, что клетка откликается меньше на стимуляцию близкой диспаратности (рис. 5.3.21 В и 5.3.21 D) и полностью теряет отклик, когда диспаратность либо много меньше (рис. 5.3.21 А), либо много больше (рис. 5.3.21 Е) 30 угловых минут, либо же когда стимул предъявляется монокулярно.

Споры ученых о том, настроены бинокулярные клетки на определенную диспаратность стимуляции или нет, вскрыли ряд серьезных методологических сложно-



**Рис. 5.3.21** Отклик нейронов коры макаки, чувствительных к факторам глубины в стимуляции. Максимум чувствительности приходится на движущиеся штрихи (любого направления) с угловым зазором в 30 угловых минут (С), чуть менее — с зазором в 15 и 45 минут (В и D), и совсем не реагируют, когда стимулируются согласованные ретинальные позиции (А) или когда зазор составляет 60 угловых минут (Е) и более (Е).



стей. Главная из них — это проблема диагностики того, какие ретинальные позиции согласуются в данный момент времени. Сие можно установить более или менее точно, если взгляд животного на чем-то фиксирован. Однако большинство экстраклеточных опытов проводилось на анестезированных животных, которые даже не были в сознании, не говоря уже о фиксации взгляда. Вопрос удалось решить американцам Джиану Поджио и Буркхарту Фишеру (1977), производившим экстраклеточную запись у обезьян, специально обученных фиксировать взгляд на определенной точке.<sup>1</sup> Используя данную методику для точного управления расположением глазных яблок в течение сессии измерений, было выявлено, что клетки V1 все же обладают чувствительностью к диспаратным стимулам (как и предполагали Барлоу, Блэйкмо и Петтигрю), но уровень диспаратности оказался весьма мал (на чем, кстати, и настаивали Хьюбел и Визель). Также в подтверждение результатов Хьюбела и Визеля было обнаружено, что клетки V2 настроены на ощутимо большую пространственную диспаратность в стимуляции. При этом некоторые из них оказались чувствительны к кросс-, а другие к анкроссдиспаратным парам стимулов.

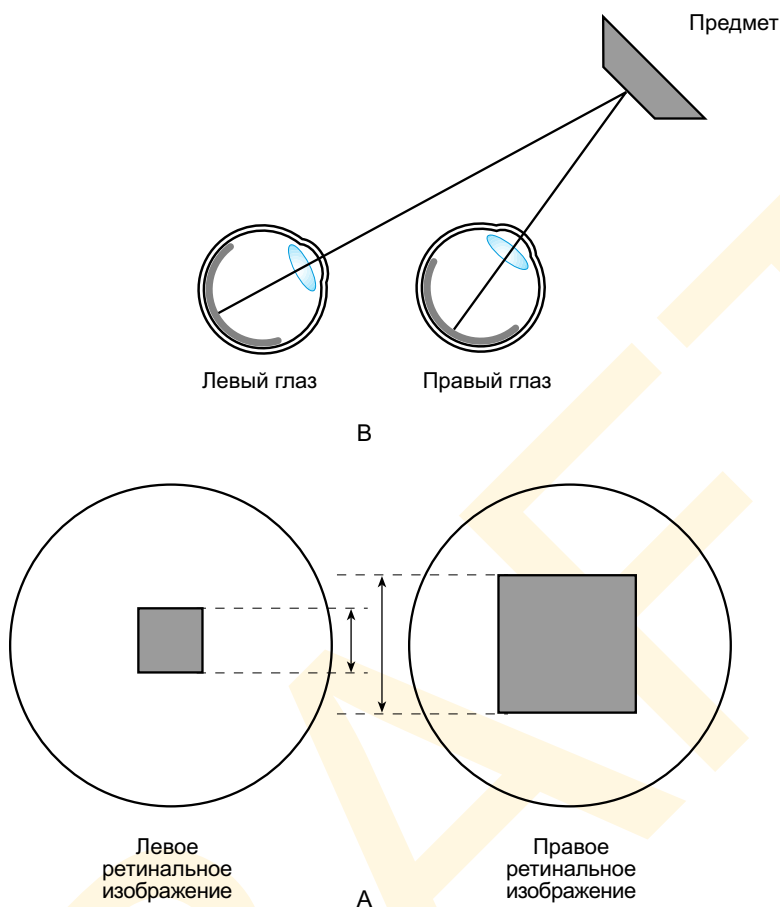
Возможно наиболее любопытным выводом из этих результатов является то, что в стереоскопию вовлечены отдельные, самостоятельные физиологические системы: одна, расположенная в V1, высокочувствительна к парам низкой или нулевой диспаратности; другая, локализованная в V2 — к парам высокой диспаратности обоих направлений (кросс и анкросс). Сие деление указывает на то, что клетки V1 могут откликаться при бинокулярной фузии внутри панумовой области вокруг горютера, и что клетки V2, перед и позади этой области, могут реагировать на двойные изображения.

Несмотря на то, что благодаря этим (и множеству других) экспериментов наше знание о стереоскопическом зрении всерьез углубилось, мы до сих пор не имеем понятия о том, за счет чего зрительная система решает задачу согласования. Но все же у нас есть надежда лучше разобраться в тонкостях обработки нейронной информации, идущей внутри бинокулярных клеток коры. Количественные теории бинокулярного согласования, использующие биологически дружественные представления, такие, в частности как МОМШ-модель Джонсона и Малик (1992) (см. раздел 5.3.3), призваны дать физиологам понимание того, механизмы каких видов они могут обнаружить, и как эти механизмы будут работать. Однако же модели эти еще столь слабы, что речь не идет даже об их апробации.

### 5.3.5 Вертикальная диспаратность

До сего момента мы говорили о бинокулярной диспаратности, полагая ее сугубо горизонтальной. Однако существует и вертикальная бинокулярная диспаратность, но которой до сегодняшнего дня почти не уделялось научного внимания. Возможно, простейший способ понять, почему вертикальная диспаратность присутствует в бинокулярных ретинальных изображениях — это рассмотреть отличия в размере проекций одного и того же объекта на левой и правой сетчатках. Рис. 5.3.22 А иллюстрирует нам зрительную геометрию ситуации при которой целевой объект находится справа от наблюдателя. Если оба глаза зафиксировались в центре целевого объекта (скажем, квадрата), то изображение в правом глазу окажется крупнее своего собрата в левом, поскольку правый глаз ближе к объекту внимания. Сие отличие в размерах ведет к

<sup>1</sup> Запись сигнала кортикальных клеток у бодрствующих макаков не причиняет им страданий, поскольку, как это не удивительно, в самом мозге нет болевых рецепторов.

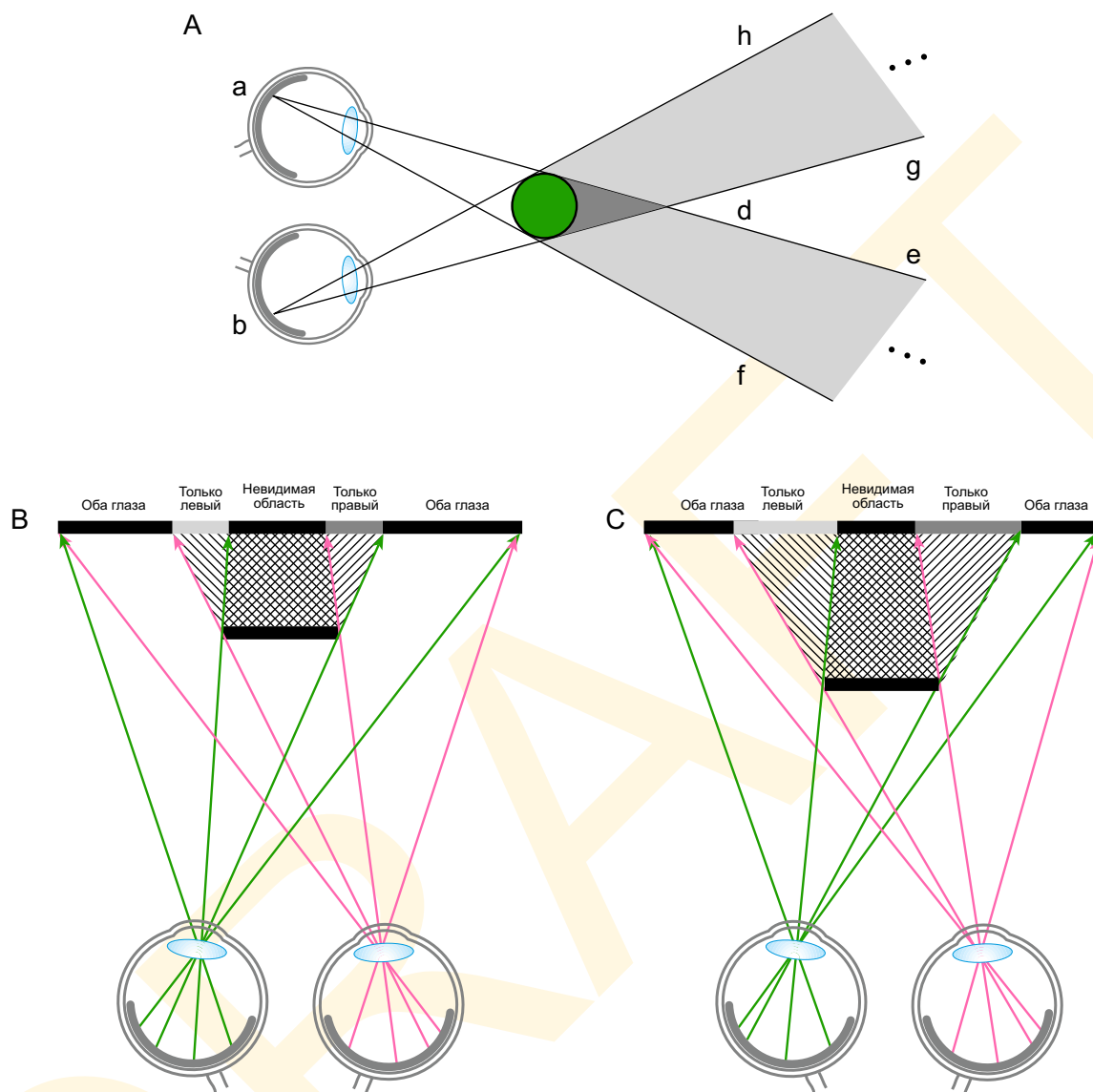


**Рис. 5.3.22** Вертикальная бинокулярная диспаратность. Когда целевой объект ближе к одному глазу (А), отличия в габаритах согласованных ретинальных изображений дают как горизонтальную, так и вертикальную диспаратность (В).

возникновению вертикальной диспаратности между согласованными точками на обеих сетчатках (рис. 5.3.22 В). Мы говорим при этом, что сия дистанционно-индуцированная диспаратность касается и горизонтального направления, поскольку отличие в размере ведет к масштабированию оптического ретинального изображения по всем направлениям.

### 5.3.6 Стереопсис Леонардо

Еще один источник стереоскопической информации, происходит из того факта, что при бинокулярном рассматривании поверхностей, расположенных на разном удалении от наблюдателя некоторые фрагменты дальней поверхности мы воспринимаем лишь одним глазом (к примеру, Nakayama & Shimojo, 1990). На рис. 5.3.23 дан пример. Отметим, что в дальней поверхности имеется участок — справа от ближней поверхности — который виден только правым глазом. Аналогично — слева от передней поверхности имеется участок, который виден только левым. Сие значит, что по этим участкам задача согласования не имеет решения, просто потому что согласованных точек на сетчатке другого глаза не существует.



**Рис. 5.3.23.** Полузакрытые участки. В нормальных условиях всегда найдутся фрагменты сцены, которые видны обоими глазами, одним глазом, другим, а также не видны. А — когда в поле зрения находится небольшой близкорасположенный предмет, то полностью невидимый участок находится сразу позади этого предмета (темная заливка). Существуют также две полузакрытых области (светло-серая заливка), из-за чего предмет становится как бы прозрачным (пояснения в цитате из Леонардо). В — полузакрытые области, видимые правым глазом, оказываются возле правого края ближнего объекта; видимые левым глазом — возле левого края. С — размер полузакрытого участка зависит от расстояний между наблюдателем и ближним объектом, также наблюдателем и фоном. То есть, информация о глубине пространства рождается из того, что определенные участки левого ретинального изображения не имеют согласованной пары в правом ретинальном изображении.

Тем не менее, участки монокулярной стимуляции дают важную бинокулярную информацию об относительных расстояниях между поверхностями. Накаяма и Шимойо (1990) назвали данный вид информации о глубине пространства *стереопсисом Леонардо* — в честь великого живописца Леонардо да Винчи, который первым описал его. Ин-

формацию о глубине зрение извлекает из зрительной геометрии, соседствующей с закрывающими краями: монокулярные участки — это всегда лишь часть *дальней* поверхности. Если монокулярный участок представлен в правом ретинальном изображении, он обязательно окажется лежащим справа от закрывающего края; если в левом ретинальном изображении — слева. Накаяма и Шимойо продемонстрировали эффективность леонардового стереопсиса экспериментально: они обнаружили, что он дает прогнозируемый перцепт относительной глубины, когда собственно присутствует (как только что описано), но не когда геометрические отношения искусственно реверсированы (то есть, когда монокулярный участок в правом ретинальном изображении находится слева от закрывающего края или когда монокулярный участок в левом изображении справа от него). С момента повторного открытия данной формы стереопсиса Накаямой и Шимойо, он был внедрен в некоторые количественные модели бинокулярной диспаратности (например, Jones & Malik, 1992).

То, что при восприятии стохастической стереограммы в сознании наблюдателя возникает объемный перцепт вовсе не означает того, что детекция краев и интерпретация поверхностей не играют роли в визуальной оценке глубины пространства.

Рассмотрим роль т.н. *окклюзии поверхностей*. В восприятии глубины окклюзия — это самый надежный монокулярный фактор глубины пространства, поскольку когда объекты загораживают друг друга — это верный признак их расположения в глубину. Накаяма и Шимойо (1990), а также Хи и Накаяма (1994) показали, что окклюзионные взаимоотношения играют важную роль в оценке бинокулярной информации.

Их демонстрация роли окклюзии в стереопсисе основана на простейшем физическом факте, продемонстрированном на рис. 5.3.23 А. В своей знаменитой *Trattato della Pittura* Леонардо приводит аналогичный рисунок, когда описывает взаимоотношения между окклюзией, полуокклюзией и прозрачностью (цитата из Витстона, 1838):

«Если объект С рассматривать одним глазом с точки А, то все объекты в пространстве позади него окажутся этим глазом невидимы; но когда другой глаз в точке В открыт — часть этих объектов становится видимой; и только те окажутся невидимыми для обоих глаз, что отмечены двойной тенью CD... Поскольку ангулярное пространство EDG позади D будет всегда видимым обоими глазами, объект С, рассматриваемый обоими глазами, окажется как будто прозрачен, согласно обычному определению прозрачных вещей, то есть который не прячет за собой ничего.»

Леонардо показал, что когда мы смотрим на объект, каждый глаз кодирует некий фрагмент сцены, который невидим для другого глаза. Такие участки он назвал полузакрытыми (Belhumeur & Mimford, 1992). Когда мы смотрим на маленький предмет, то позади него имеется небольшое пространство, полностью закрытое этим предметом, а за ним еще одно пространство, просматриваемое обоими глазами. Если мы сосредоточимся на дальнем пространстве, то сам объект станет как бы прозрачным.

Существует несколько простых правил, описывающих положение и свойства полузакрытых участков: во-первых, полузакрытые участки, видимые левым глазом, всегда расположены на левом крае кроющего объекта, а видимые правым глазом — на правом. Иначе просто невозможно физически (рис. 5.3.23 В); во-вторых, относительный размер полузакрытых участков систематически меняется по мере изменения расстояния до кроющего и кроемого объектов — если кроющий предмет оказался ближе к наблюдателю, полузакрытые участки становятся крупнее (рис. 10.16 С). Следовательно и положение, и размер полузакрытых участков несут информацию, которую можно использовать для реконструкции глубины.

Накаяма и Шимойо (1990) обнаружили, что окклюзионная информация влияет на наблюдательскую оценку бинокулярной глубины. В их экспериментах испытуемым предъявлялась стереограмма с физически невозможными полуокклюзиями. Исследователи установили, что когда полуокклюзия представляла собой, скажем, некий паттерн, видимый только правым глазом, но возле левого края кроющего объекта, видность этого паттерна подавлялась. Когда этот же паттерн предъявлялся возле правого края (то есть так, как возможно в действительности) — стимул становился легко видимым. Андерсон и Накаяма (1994) обобщили данные по ряду наблюдений и пришли к выводу, что конфигурация окклюзии (т.е. свойства поверхностей и объектов) влияют на ранние этапы стереопсиса.

Итак, мы понимаем теперь, что восприятие глубины пространства (подобно восприятию движения и цветовому восприятию) — это зрительная инференция. Результаты вышеупомянутых исследований показали, что инференция эта зависит от фундаментальных свойств поверхностей и объектов, а именно от того, что они способны оптически закрывать друг друга.

## 5.4 ДИНАМИЧЕСКАЯ ИНФОРМАЦИЯ

Динамическая зрительная информация относится к изменениям в зрительной структуре, возникающим с течением времени, т.е. благодаря *оптическому потоку*, как говорил Гибсон. Когда наблюдатель движется относительно предметов сцены, направление и скорость, с которой смещаются ретинальные проекции различных объектов, зависят не только от скорости движения наблюдателя, но и от того, насколько объекты эти удалены от наблюдателя и на чем он зафиксировал взгляд. Информацию о глубине зрение в том числе извлекает из т.н. *параллакса движения* — разнице в характере смещения пар точек, возникающей из-за их различной удаленности от точки фиксации взгляда. Несмотря на то, что основной разговор о восприятии движения мы отложим вплоть до 10-й главы, коротко мы коснемся его именно здесь, поскольку движение — мощнейший источник информации о глубине пространства.

### 5.4.1 Параллакс движения

Убедиться в существовании параллакса движения, а также оценить его характер мы можем с помощью «пальцевого опыта», весьма схожего с тем, что мы провели, исследуя бинокулярную диспаратность.

Схема такова:

1. Расположим оба указательных пальца перед собой — один на вытянутой руке, один на полпути до носа.
2. Закроем правый глаз и наведем оба пальца на удаленный объект, сфокусировавшись на этом объекте.
3. Удерживаем пальцы так твердо, как только можем, начнем слегка смещать голову вправо. Отметим при этом, что оба пальца визуальнo сдвинулись влево относительно удаленного объекта, но при этом ближний палец смещается больше и быстрее.
4. Теперь сдвинем голову влево и отметим, что наши пальцы визуальнo сместились вправо, но вновь ближний сдвинулся быстрее и больше.

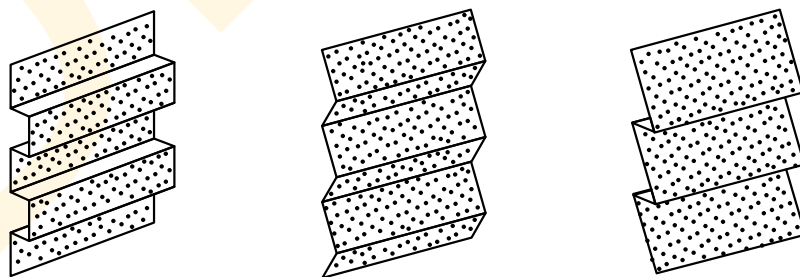
Разница в движении пальцев в данном примере демонстрирует характер параллакса движения: изображения точек, расположенных на разных удалениях от наблюдателя движутся с разными ретинальными скоростями, когда смещается точка стояния.



Тесное информационное родство между параллаксом движения и бинокулярной диспаратностью можно обнаружить, сравнив оба пальцевых эксперимента. В случае бинокулярной диспаратности мы неподвижно держали голову и сравнивали левое ретинальное изображение с правым, при этом каждое из них существовало на сетчатках одновременно. В случае параллакса движения мы перемещаем голову в течение некоторого времени и сравниваем первое изображение с последним. Таким образом, бинокулярная диспаратность касается отличия в паре смещенных *одновременных* ретинальных изображений, а параллакс движения — того же, но в паре изображений *последовательных*. Если во втором случае мы сместим голову латерально на расстояние строго равное расстоянию между центрами глаз, то отличие между первым и последним изображениями будет ровно таким же как между правым и левым изображениями в опыте по демонстрации бинокулярной диспаратности.

Эксперименты показали, что параллакс движения обретает особое значение, когда нет иных источников трехмерной информации. Если, к примеру, два неподвижных самосветящихся диска, находящихся на разных расстояниях от наблюдателя, рассматривать монокулярно в темной комнате и при этом двигать головой вправо-влево, перцепты этих объектов утратят статичность и начнут смещаться один относительно другого. Однако более сложные паттерны светящихся точек могут дать любопытные перцепты глубины пространства. К примеру, стохастические картинки, состоящие из тысяч точек, движущихся так, как если бы они лежали на рифленых поверхностях различной пространственной структуры (рис. 5.4.1), давали перцепт определенной глубины, когда двигался либо наблюдатель, либо когда двигалось экранное изображение, имитируя параллакс (Rogers & Graham, 1979).

Как и в случае бинокулярной диспаратности, характер ретинального параллакса движения зависит не только от расстояния до объектов сцены, но также до точки фиксации взгляда наблюдателя. Чтобы увидеть отличие, повторим опыт с пальцами, но на сей раз попытаемся вначале сфокусироваться на удаленном объекте, а затем на ближнем пальце. Если мы сдвинем голову вправо, сфокусировавшись на дальнем объекте, наши пальцы визуальнo сместятся влево — в направлении противоположном направлению движения головы. При этом ближний палец будет «двигаться» быстрее. Теперь сдвигаем голову вправо, но при этом удерживая фокус на *ближнем* пальце (что довольно трудно). Отметим, что на сей раз и дальний палец, и удаленный объект смещаются *вправо* — то есть, в направлении головы. При этом объект движется быстрее. Таким образом, паттерн ретинального параллакса движения меняется на обратный при разных



**Рис. 5.4.1.** Параллакс движения. Чтобы продемонстрировать эффективность параллакса движения при восприятии глубины пространства Б. Роджерс и М. Грехэм (1979) использовали мятые поверхности, покрытые стохастическими точками. Показаны некоторые из таких поверхностей. В фактическом опыте видимых краев не было, кроме тех, что определены относительным движением других случайных точек.

удаленностях от точки фиксации, то есть в зависимости от того, ближе или дальше от этой точки расположен объект. Отличия по направлению и скорости параллакса движения — суть точная аналогия степени кросс- и анкроссдиспарантности при стереоскопическом зрении.

Полезно отметить, что, аналогично бинокулярной диспарантности, в норме наши ощущения от параллакса движения — это перцепты объектов, расположенных скорее на разных расстояниях от нас, нежели их проекций, движущихся с разными скоростями по сетчатке. Мы ощущали разную скорость движения в опыте с пальцами, поскольку явно нацелились на данный аспект зрительной стимуляции (т.е. на движение, а не на глубину) и поскольку в данной ситуации движение особенно заметно. Однако когда мы сами перемещаемся в сцене, очень трудно ощутить зрительный параллакс в явном виде — неподвижные предметы сцены воспринимаются как неподвижные, даже несмотря на то, что их ретинальные проекции движутся. Сия способность воспринимать фактическое положение предметов в сцене, несмотря на непрерывное движение их проекций по сетчатке называют *константностью положения*, и отметим, что последняя играет важнейшую роль в возникновении перцепта глубины из параллакса движения. Эту и другие формы константности мы рассмотрим в гл. 7, когда будем детально разбирать зрительную диагностику физических свойств объектов внешнего мира.

Параллакс движения еще тем похож на бинокулярную диспарантность, что дает лишь относительную информацию о глубине. То есть, он не описывает фактического расстояния до предмета, но лишь то, насколько ближе или дальше этот предмет находится от точки фиксации. В отличие от бинокулярной диспарантности, однако, параллакс движения дает эффективную информацию о глубине на больших расстояниях: к примеру, когда мы едем в поезде мимо двух холмов, из которых один удален от нас на 2 км, а второй на 4, то именно благодаря параллаксу движения мы понимаем, что один ближе другого. Бинокулярная диспарантность в данном случае не даст ничего, потому что расстояние между центрами глазных яблок слишком мало в сравнении с расстоянием до объектов.

#### 5.4.2 Оптический поток от движения наблюдателя

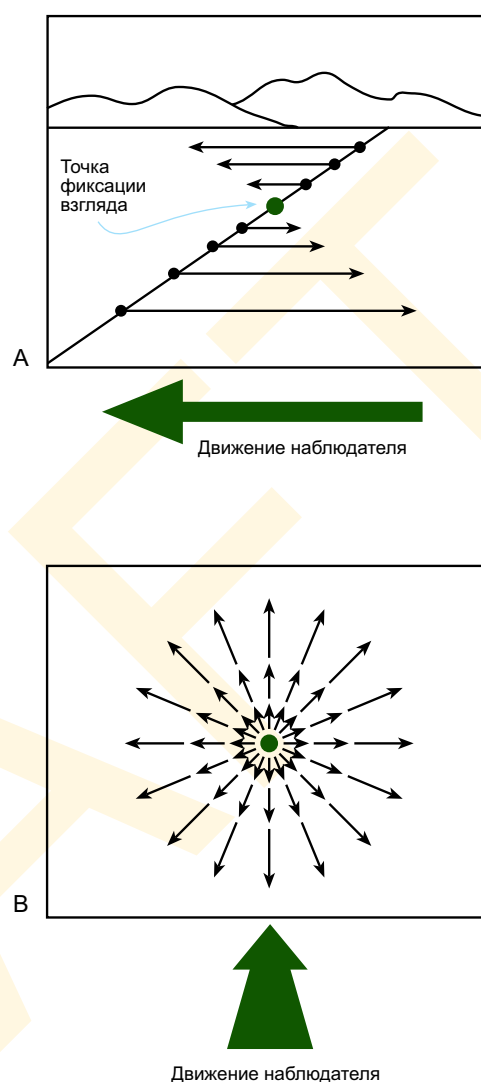
В реальном восприятии относительный параллакс движения двух изолированных точек встречается крайне редко, если вообще возможен. Как особо отмечал Гибсон (1966), двигается обычно сам наблюдатель и при этом активно, деятельно исследует нагромождение предметов в окружающем пространстве. Сие рождает сложные паттерны оптического потока.

Ученый весьма доказательно продемонстрировал, что движение ретинального изображения, вызванное движением наблюдателя, отнюдь не хаотическое, но высокоструктурированное и зависит как от расположения поверхностей в трехмерной сцене, так и от движения самого наблюдателя. Он также показал, что сию общую карту оптического потока можно использовать в восприятии пространственного расположения поверхностей. Систематические изменения в движении ретинального изображения по большим областям оптического массива Гибсон назвал *градиентами движения*, дабы подчеркнуть то, что они весьма (но при этом плавно) изменчивы по скорости и направлению. Ученый проанализировал характер этих градиентов в определенных экологически значимых ситуациях.

На рис. 5.4.2 даны несколько примеров гибсоновских градиентов движения: стрел-

ки указывают на то, как меняется картина точек ретинального изображения по мере движения наблюдателя (как если бы отслеживались их траектории с помощью специальной серийной фотосъемки). На рис. 5.4.2 А дана картина оптического потока при плавном движении наблюдателя в направлении, показанном большой стрелкой внизу и при точке фиксации взгляда в центре сцены. Таким же окажется паттерн оптического потока, если мы, двигаясь на машине по проселочной дороге среди полей, остановимся на корове (точка фиксации), стоящую возле длинного забора, перпендикулярного этой дороге (диагональная линия). Ретинальная проекция точки фиксации (коровы) неподвижна, поскольку наблюдатель следит за ней за счет движения головы и глаз, чем удерживает проекцию на центральной ямке сетчатки. Содержимое оставшейся части поля зрения меняется с какой-то определенной скоростью и направление смещения каждого участка зависит от положения объектов по отношению к точке фиксации взгляда: все точки, расположенные ближе к наблюдателю от этой точки будут смещаться в направлении, противоположном направлению движения наблюдателя, а все точки, лежащие за точкой фиксации — в направлении движения наблюдателя. В обоих случаях скорость оптического потока зависит от разницы в глубине: она максимальна по точкам максимально отличным по удалению от точки фиксации и минимальна по точкам наименее удаленным от нее. Таким образом, стрелки на рис. 5.4.2 А демонстрируют то, как точки, расположенные вдоль забора, будут смещаться, если мы зафиксируем взгляд на корове при нашем движении влево. Ежели мы вернемся к опыту с пальцами, мы поймем, что сии градиенты движения вполне согласуются с тем, что мы уже наблюдали.

Еще одним типичным паттерном оптического потока, описанном Гибсоном, является *оптическая экспансия* или *люминг* (см. рис. 5.4.2 В). Возникает, когда наблюдатель движется прямо в направлении поверхности, расположенной фронтально, зафикси-



**Рис. 5.4.2** Градиенты движения. Гибсон говорил о паттернах относительного движения, вызванного перемещением наблюдателя. А — оптический поток, образовавшийся от движения наблюдателя влево (крупная стрелка) при точке фиксации взгляда в центре линии. В — паттерн оптической экспансии (люминг), возникший в результате движения наблюдателя прямо на точку фиксации взгляда, к примеру, прямо на отвесную стену.

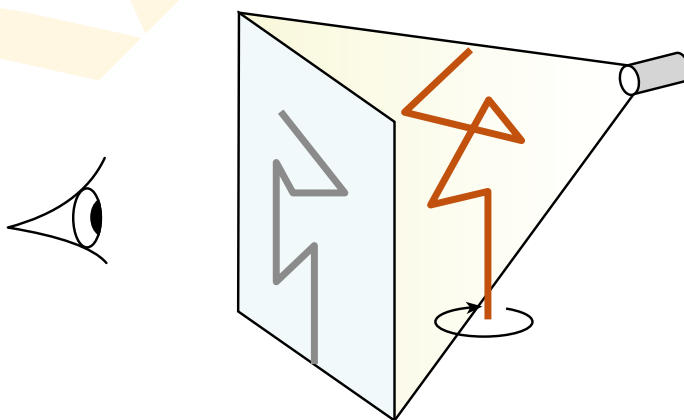
сировав при этом взгляд строго по курсу движения. Оптическая экспансия имеет место, например, когда мы идем прямо на стену, глядя строго перед собой. Проекция точки фиксации, как всегда, неподвижна на сетчатке, а проекции все остальных точек разбегаются кнаружи от нее со скоростью, растущей пропорционально удаленности от точки фиксации. Как нетрудно себе представить, если отдаляться от плоскости, то картина оптического потока будет обратной, в которой все точки поверхности начнут сбегаться к точке фиксации.

Описанные варианты градиентов движения — это лишь два случая очень простого оптического потока, возникающего при очень простом движении по отношению к единичной поверхности сцены. В повседневных условиях оптические потоки радикально сложнее. Они — плод изменений в направлении взгляда благодаря движениям головы и глаз, движениям нашего тела вверх-вниз при ходьбе, кривизны пути и сложной организации сцен, содержащих поверхности разной ориентации, на разных расстояниях, под разными углами и разной искривленности. Картины реального оптического потока, возникающие при нормальной жизни человека в нормальных условиях так сложны, что не поддаются делению на простые группы. Их структуру можно классифицировать лишь путем углубленного математического анализа.

### 5.4.3 Оптический поток от движения объектов

Информация о третьей пространственной размерности объекта — глубине, становится доступной нашему зрению в том числе и тогда, когда в движении относительно наблюдателя находится сам объект. Если различные фрагменты объекта располагаются на разном расстоянии от наблюдателя (то есть, если объект объемен) и если фрагменты эти в процессе движения меняют свое относительное расстояние до наблюдателя (как, например, когда объект вращается вокруг своей оси), то образуется некое относительное движение точек поверхности этого объекта. Сие относительное движение дает информацию о том, какие точки поверхности ближе, какие дальше от наблюдателя. То есть, относительную глубину видимой поверхности можно восстановить из движения объекта и тем воспринять форму этого объекта.

Способность воспринимать объем предмета через его движение была впервые продемонстрирована американцами Г. Валлахом и Д. О'Коннелом (1953) на примере феномена, именуемого *кинетическим эффектом глубины (КЭГ)*. Как показано на рис. 5.4.3, исследователи сконструировали трехмерную проволоочную фигуру неправильной формы и дали проекцию ее тени на полупрозрачный экран. Пока



**Рис. 5.4.3** Кинетический эффект глубины. Трехмерная фигурка из гнутой проволоки подсвечивается сзади, а ее тень падает на полупрозрачный экран, за которым наблюдает испытуемый. В покое тень выглядит обычным двухмерным контуром, но когда проволока начинает движение вокруг своей оси, то сразу начинает восприниматься трехмерным вращающимся объектом.

проволока оставалась неподвижной, объем фигуры никак не ощущался — воспринимался лишь изломанный теневой контур на плоском двухмерном экране. Но как только проволочная фигура начинала вращаться вокруг своей вертикальной оси, наблюдатель, находившийся позади экрана моментально воспринимал ее трехмерность. Когда же фигурка останавливалась, давешний трехмерный перцепт сразу схлопывался до двухмерной линии на плоскости экрана. Сей эффектный опыт наглядно демонстрирует нам, что в данном случае информация о глубине извлечена зрением сугубо из динамической стимуляции, как минимум по ряду параметров. Как же сие возможно и как зрительная система пользуется динамической информацией в данном ключе?

Во-первых, мы должны понимать, почему реконструкция размерности глубины из вращения объекта, равно как и любой другой трехмерной информации, не предопределена геометрически и, следовательно, логически недоопределена. В данном случае двухмерное ретинальное движение, возникшее в результате движения жесткого трехмерного объекта *может* восприняться строго по номиналу, то есть как двухмерная гибкая ритмично изменяющаяся фигура. В случае КЭГ стимул в буквальном смысле является такой фигурой — это циклично деформирующаяся двухмерная картинка, поскольку является лишь тенью трехмерного объекта на полупрозрачном экране. Однако же зрением наблюдателя такая стимуляция толкуется только как тень жесткого трехмерного объекта определенной формы, вращающегося вокруг своей оси.

Как всегда для достижения однозначной интерпретации стимуляции в недоопределенных ее условиях, зрительной системе приходится в неявном (или даже в явном) виде строить дополнительные предположения. В данном случае зрение работает как система сугубо эвристическая — идет смещение в сторону восприятия именно движений некоего жесткого предмета, нежели пластических деформаций. К примеру, исключительно мала вероятность того, что некая двухмерная фигура стала бы деформироваться строго тем же образом, какой согласуется с движением твердого предмета (если только движение это не вызывает двухмерной деформации как в случае эксперимента Валлаха — О'Коннелля).<sup>1</sup>

Отчего же зрительная система селективно «настроена» на интерпретацию стимуляции как движения твердого предмета, нежели как пластических метаморфоз? Наиболее очевидный ответ — это ответ эмпирический: движение твердого предмета много вероятнее, нежели пластические изменения. Гештальтисты, однако, скажут, что сие объяснимо принципом Прягнанца, то есть, что зрительная система дает простейшую интерпретацию в данных условиях стимуляции. Движение твердого тела — это *простейший вариант*, поскольку не предполагает динамических изменений формы предмета и потому может быть представлен и обработан наиболее экономично. Сие еще один пример того, что принцип подобия и принцип Прягнанца часто прогнозируют одно и то же следствие у самых разных причин.

#### 5.4.4 Динамика текстуры

Еще одним источником зрительной информации о глубине пространства — это динамика текстуры позади движущегося края (Gibson, Kaplan, Reynolds & Wheeler, 1969). Край всегда принадлежит ближней поверхности, а динамика текстуры — даль-

<sup>1</sup> Однако в гл. 10 мы покажем, что существуют некие условия в которых вращение твердого тела воспринимается как пластическая деформация. Однако общая тенденция остается неизменной, то есть — в направлении восприятия стимуляции как движения твердого предмета.



ней. Мы можем убедиться в этом следующим образом: в двух экземплярах распечатаем стохастическую стереограмму рис. 5.3.8, а затем вырежем из первого экземпляра небольшой квадратик и положим его на второй. Если копия сделана аккуратно и ее края плотно легли на плоскость, то различить квадрат будет трудно. Теперь подвигаем квадратик пальцем туда-сюда или, скажем, по кругу — перцепт моментально изменится: из гомогенной текстурированной поверхности активно выделится текстурный квадратик, который окажется перед текстурным фоном; стоит остановиться — как в течение нескольких секунд квадратик исчезнет, слившись с фоном.

Динамика текстуры, возникающая в результате движения наблюдателя, тем же образом относится к стереопсису Леонардо, каким параллакс движения — к бинокулярной диспаратности. То есть, динамика текстуры — это динамический источник информации о глубине пространства, обнаруживающий себя во временном домене, равно как статическая бинокулярная информация (за счет разности полей зрения правого и левого глаз — стереопсис Леонардо) — в домене пространственном. Они также относятся и к окклюзионной информации, которую мы будем обсуждать в следующем разделе. Наибольшее отличие состоит в том, что информация, поступающая от динамики текстуры требует временной интеграции.

## 5.5 ПИКТОРИАЛЬНАЯ ИНФОРМАЦИЯ

Несмотря на то, что стереопсис и движение дают полноценный перцепт объема — это далеко не единственные источники информации о глубине пространства. Наиболее очевидный аргумент в пользу этого вывода таков, что если мы закроем один глаз и будем держать голову неподвижной, то мир вокруг все равно будет восприниматься трехмерным. С другой стороны, высококлассные фотографии и компьютерные анимации могут на удивление точно передавать глубину пространства. Сие видно даже при мимолетном взгляде на фотографию рис. 5.5.1, на котором изображена трехмерная сцена. Поэтому все остальные статические источники визуальной информации о глубине пространства относят к категории т.н. *пикториальных факторов*, поскольку они потенциально доступны в статике и при монокулярном рассматривании изображений.

Пикториальная информация может оказаться на удивление добротной и часто ведет к полноценному ощущению глубины при рассматривании двухмерных картин, то есть тогда, когда и стереоинформация, и информация о движении предметов указывают на то, что источник визуальной стимуляции — совершенно плоский. Пикториальная информация может порой даже превозмочь стереоинформацию (то есть, подаваемую специальными оптическими устройствами отдельно в правый и левый глаз, что мы обсудим в разделе 5.5.10, когда поведем разговор об интеграции информации о глубине пространства).

### 5.5.1 Перспективные проекции

Свет, отраженный от объектов сцены, по прямому падает на сетчатку наблюдателя. Геометрия, описывающая этот процесс, именуется *перспективной проекцией* и демонстрирует принципиальное отличие между предметами реального мира и их оптическими изображениями. Наиболее очевидные из этих отличий касаются пространственной мерности: *трехмерные* объекты дают лишь *двухмерные* проекции. Утраченная (третья) размерность — глубина. Однако было показано, что, к счастью, двухмерные изображения, сформированные перспективной проекцией, все же несут в себе оп-



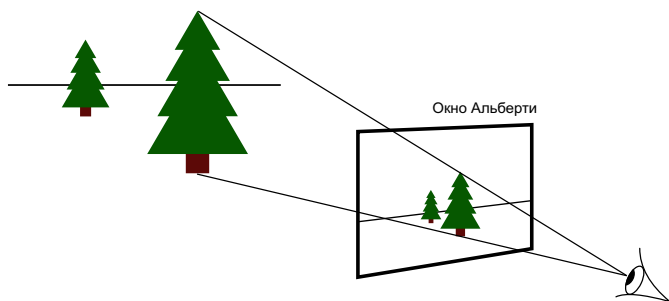
**Рис. 5.5.1** Пикториальные источники визуальной информации о глубине пространства. В данном снимке есть практически вся визуальная информация, которую зрение в состоянии извлечь из стационарной монокулярной структуры изображения.

ределенную визуальную информацию об этой «потерянной» размерности, что позволяет хотя бы частично восстановить ее.

Значимость перспективной проекции в зрительном восприятии глубины пространства осознали несколько веков назад, когда живописцы стали пытаться как можно реалистичнее передать глубину пространства на плоских холстах. В 1436 итальянский живописец Леон Альберти<sup>1</sup> описал технику письма в т.н. *собственной перспективе*, что произвело революцию в живописном представлении глубины. Основная идея довольно проста (рис. 5.5.2): чтобы точно перенести трехмерную сцену на двухмерный холст, нужно организовать изображение так, как будто оно проходит через плоское окно, расположенное на холсте. Если мы отнесемся к этой идее буквально, то можем воссоздать вполне реалистичный контурный рисунок сцены, рассматривая ее через оконное стекло одним глазом (в фиксированной позиции) и отрисовывая контуры восковым карандашом или мелком непосредственно на стекле. Если затем положить это стекло на равномерную белую поверхность и рассматривать монокулярно примерно с той же позиции, то наш рисунок выдаст всю пикториальную информацию о глубине, заключенную в контурах. Сей метод достижения собственной перспективы в рисунке именуется *окном Альберти*. Спустя полвека Леонардо, время от времени занимавшийся вопросами перспективы в живописи, отстаивал аналогичные идеи.

Наиболее интересной особенностью перспективной проекции является то, что с по-

<sup>1</sup> Леон Баттиста Альберти (1404 — 1472) — итальянский ученый, гуманист, писатель, один из зачинателей новой европейской архитектуры и ведущий теоретик искусства эпохи Возрождения. — *Прим. пер.*

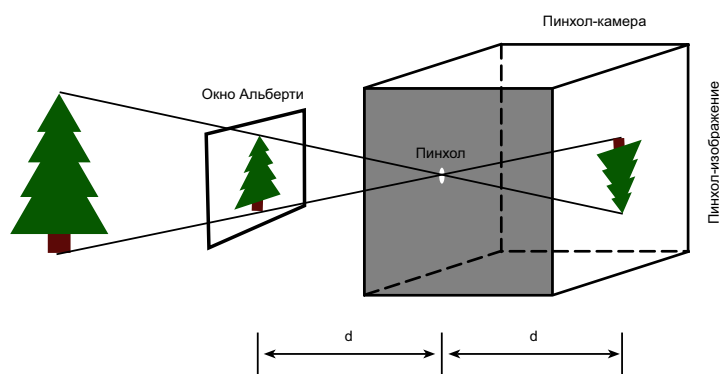


**Рис. 5.5.2.** Окно Альберти. Перспективный фактор глубины, заложенный в трехмерную сцену, можно сохранить на двухмерной поверхности, если отрисовать контуры предметов на стекле, рассматривая сцену через это стекло с определенной точки.

ределенных условиях иллюзия глубины (при монокулярном просмотре) окажется весьма убедительной. При этом, разумеется: фотографическое качество должно быть безупречным; глаза должны оставаться фиксированными в определенной позиции (то есть не должно возникать потока какой-либо информации о движении); края фотографии должны быть закрыты так, чтобы мы не могли видеть ее границ, проходящих по сцене; просмотрное расстояние должно быть не меньше, чем 1.5 – 2 м, дабы информация от монокулярной аккомодации не повредила иллюзию глубины (Pirenne, 1970). Ощущение глубины при этом все равно не окажется столь явным, как при стереоскопии, даже не таким сильным, как при восприятии движения -- оно может быть лишь таким, какое возникает при рассматривании реальной сцены через окно одним глазом в неизменном положении.

Перспективная проекция, реализуемая высококачественной фотографией, фактически несет в себе все пикториальные факторы глубины пространства (о чем мы поговорим ниже), поскольку захватывает всю информацию из стационарного ретинального оптического изображения. В нескольких следующих разделах мы разберем некоторые из наиболее серьезных источников пикториальной информации о глубине пространства — поговорим о том, каковы они и как работают. Однако прежде

мощью окна Альберти мы можем вызвать у наблюдателя весьма убедительную иллюзию глубины. Подобное происходит в камере обскура при проекции экспонирующего изображения на плоскость листовой фотопленки (рис. 5.5.3): проекция сцены на плоскость окна Альберти — это всего лишь прямая версия перевернутого изображения в камере обскуры, в которой плоскость изображения расположена на таком же расстоянии от отверстия, как глаз наблюдателя от окна Альберти. Сказанное означает, что если мы старательно отрисуем картинку на окне Альберти, она явится полным эквивалентом перевернутому изображению, созданному пинхолом. В оп-



**Рис. 5.5.3.** Окно Альберти и пинхол-камера. Изображение в окне Альберти — это прямая версия перевернутого изображения сцены, сфокусированного на задней поверхности камеры обскуры. Отверстие камеры расположено в точке наблюдения, а расстояние до задней поверхности равно расстоянию от точки наблюдения до окна Альберти.



важно отметить, что все эти источники исключительно эвристические, что порой ведет к возникновению иллюзий глубины. Как мы уже говорили ранее, перцепт глубины, возникающий при рассматривании фотографий и живописных полотен — это одна из форм иллюзии, поскольку изображения эти на самом деле совершенно плоские.

### 5.5.2 Конвергенция параллельных линий

Одной очень важной особенностью перспективных проекций является то, что параллельные линии сцены проецируются на двухмерную плоскость непараллельными и пересекаются в т.н. *точке схода* на линии горизонта. На рис. 5.5.4 показано, что все параллельные линии, лежащие в сцене на одной плоскости, в изображении этой сцены сходятся в одной общей точке. Конвергенция линий возникает из-за того, что зрительный угол между двумя ближними к наблюдателю отрезками этих линий оказывается очень большим, а по мере удаления — падает, сходя в ноль по бесконечно удаленным отрезкам.

Сие свойство перспективной проекции часто именуют *линейной перспективой*, но мы назовем его *конвергенцией параллелей*, подчеркивая тем особый характер заключенной в нем визуальной информации. Возможно, наиболее эффектным примером конвергенции параллелей является уменьшение визуального расстояния между рельсами железнодорожного полотна или обочинами дороги, уходящей вдаль.

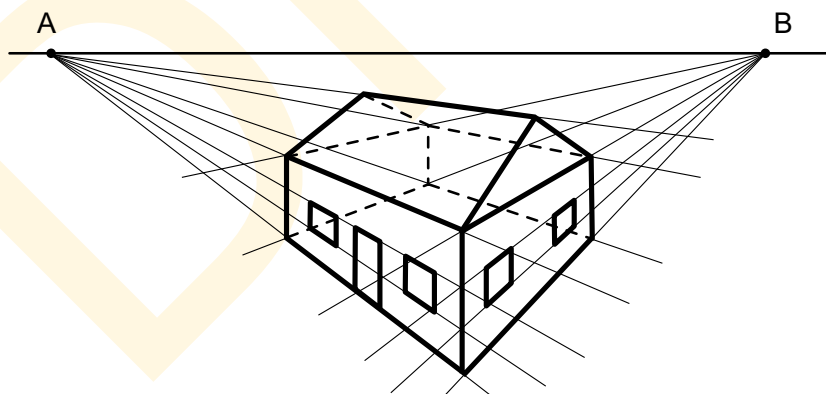


Рис. 5.5.5. Пример двухточечной перспективной проекции. Линии параллельных краев домика пересекаются в двух точках на линии горизонта.

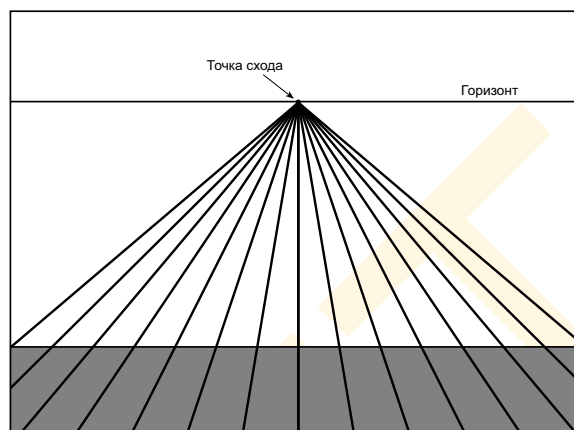


Рис. 5.5.4. Конвергенция параллелей (линейная перспектива). В перспективной проекции трехмерной сцены на двухмерную плоскость линии, параллельные на исходной плоскости сцены, пересекаются в т.н. *точке схода*, расположенной на линии горизонта.

Давайте теперь внимательно посмотрим на рис. 5.5.5 и прикинем точки схода для передачи реалистичной глубины. В данном рисунке две таких точки, поскольку имеются два отдельных множества параллельных линий, лежащих на т.н. *исходной плоскости*. Убедиться в существовании точек схода на линии горизонта (лежащей, соответственно, на исходной

плоскости), достаточно легко, однако существует бесконечное множество точек схода, лежащих на «линиях горизонта» в других плоскостях трехмерного пространства. К примеру, вертикальная линия, проходящая через точку В на рис. 5, — это линия горизонта для вертикальной плоскости, несущей на себе правую фронтальную поверхность изображенного предмета. Все параллельные линии этой плоскости будут конвергировать в точке схода, расположенной где-то вдоль этой линии, также как в случае «истинного» горизонта на плоскости земли.

### 5.5.3 Склонение от горизонта

Еще одним пикториальным источником информации о глубине пространства, образующимся из перспективной проекции, является высота объектов в картинной плоскости относительно горизонта. На рис. 5.5.6 дан пример сцены, в которой нет прямых, сходящихся в одну точку, но при этом все равно имеется устойчивое ощущение глубины пространства. Деревья внизу выглядят как стоящие ближе тех, что сверху, а облака, которые расположены возле верха изображения, смотря ближе тех, что возле линии горизонта. Базой такого перцепта является определенное расположение элементов изображения на плоскости относительно линии горизонта.

Рассмотрим первый случай, когда предметы находятся на уровне земли, например, как камни, разбросанные по ней и наблюдаемые наискось с высоты 1.5 – 2 м (типичная ситуация). Предметы, расположенные на картинной плоскости выше, то есть, ближе к горизонту, воспринимаются как более удаленные от точки наблюдения. Теперь вообразим, что те же камни рассеяны на прозрачной поверхности над нашей головой. В этом случае камни, расположенные на картинной плоскости ниже, но по-прежнему ближе к горизонту, опять же будут восприниматься как более удаленные. То есть, общее правило таково, что все объекты на горизонтальной поверхности, чем ближе окажутся на изображении этой поверхности к линии горизонта, тем более удаленными будут казаться.

Сие описание звучит так, как будто позиция относительно горизонта на плоскости земли это единственный штатный источник информации о глубине пространства. На деле же информация оказывается доступной если горизонт виден или определен каким-то образом из зрительной информации (Sedwick, 1986). Можно доказать геометрически, что расстояние ( $d$ ) от наблюдателя до любой точки ( $P$ ) на поверхности можно рассчитать из угла горизонта ( $A$ )<sup>1</sup> и длины перпендикуляра к поверхно-

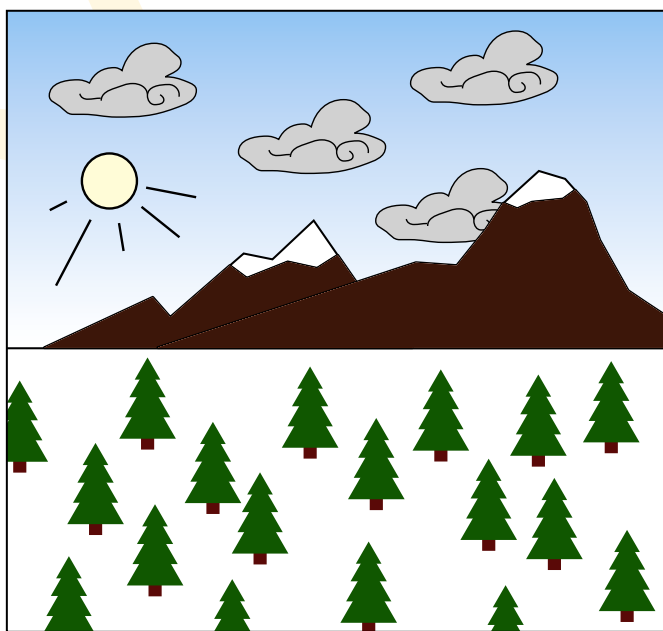
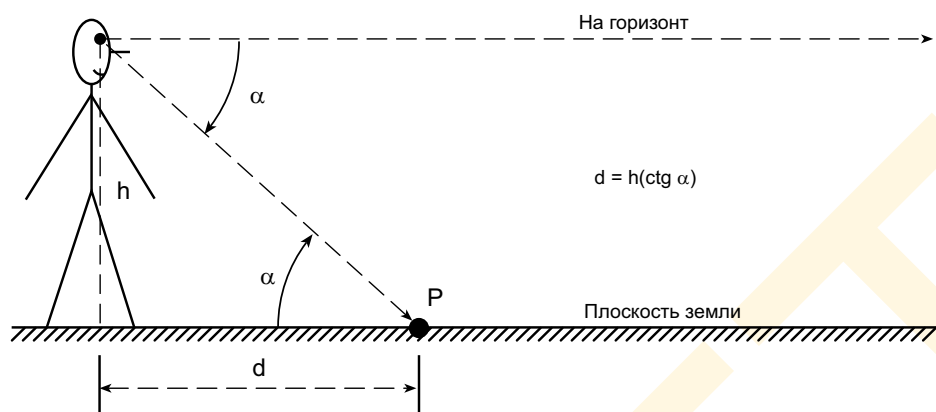


Рис. 5.5.6. Положение относительно горизонта. В перспективной проекции трехмерной сцены объекты, расположенные в одной горизонтальной плоскости ближе к горизонту, будут восприниматься более удаленными от наблюдателя.





**Рис. 5.5.7** Расстояние как функция от горизонтального угла точки на плоскости. Расстояние ( $d$ ) до точки на плоскости (в данном случае — плоскости земли) есть произведение длины перпендикуляра к плоскости ( $h$ ) на котангенс горизонтального угла ( $\alpha$ ), т.е. угла между зрительными осями на горизонт и на точку.

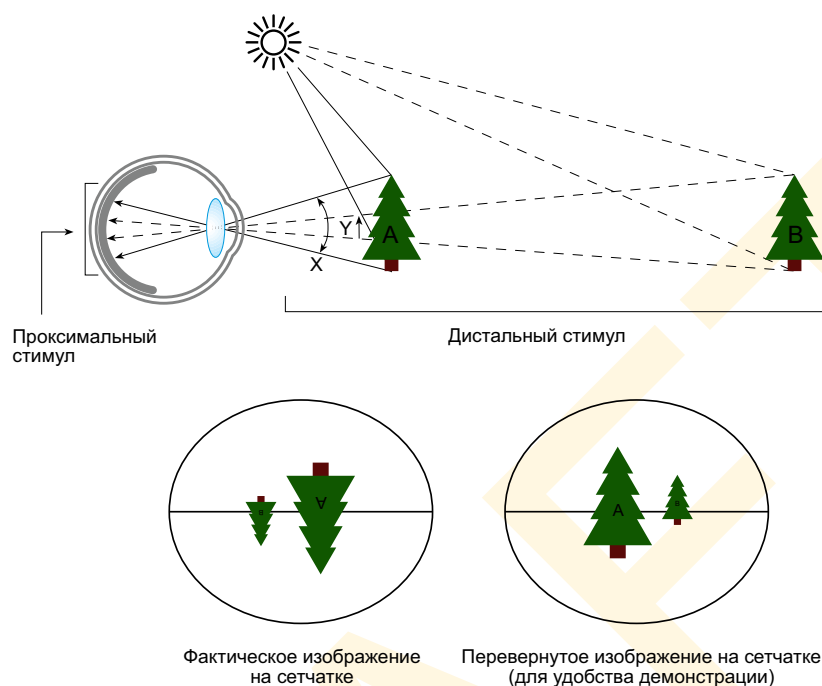
сти ( $h$ ), т.е. роста наблюдателя (см. рис. 5.5.7). Поскольку наблюдатель обычно четко знает о высоте собственных глаз над поверхностью земли из величины роста и позы, он может, в принципе определить эгоцентрическое расстояние до низа объектов на этой плоскости. Поскольку те же геометрические отношения касаются любой поверхности, чья линия горизонта видна, расстояние до точек других поверхностей можно также определять до тех пор, пока их линии горизонта имеют место в ретинальном изображении и пока известно перпендикулярное расстояние от глаза до земли. Длину перпендикуляров, не доходящих до земли, не так легко узнать, но даже без этого параметра, относительные расстояния до двух объектов на общей поверхности можно рассчитать из углов их горизонтов. Сие возможно, поскольку перпендикуляр к поверхности ( $h$ ) аннулируется, если взять отношение расстояний до двух объектов на одной плоскости. Таким образом, угловые расстояния до горизонта на плоскости поставляют зрению весьма важную количественную информацию о расстоянии от наблюдателя до объектов на этой плоскости.

#### 5.5.4 Относительный размер

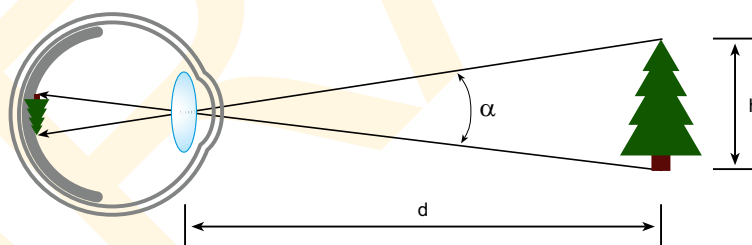
Давайте вновь посмотрим на рис. 5.5.6 и обратим внимание на некоторую странность: «ближние» елки на нем того же размера, что и «дальние». Действительно, деревья на рисунке не соответствуют очевидному факту: если два предмета равны по размеру, то проекция дальнего предмета на сетчатку окажется мельче проекции ближнего. По рис. 5.5.8 четко видно, что, когда свет от двух объектов одинакового размера, расположенных на разных расстояниях от наблюдателя, падает на сетчатку этого наблюдателя, *зрительный угол* (угловой размер) ближнего предмета оказывается больше, чем дальнего.

Простейшая тригонометрия поможет нам определить угловой размер изображения предмета на сетчатке как функцию от расстояния до самого предмета в сцене. На рис. 5.5.9 дан прямоугольный треугольник, образованный зрительными осями на

<sup>1</sup> Угол между осью взгляда на горизонт и осью взгляда на данную точку поверхности.



**Рис. 5.5.8.** Формирование на сетчатке оптического изображения сцены. Свет, отраженный от объемных (трехмерных) объектов (дистальный стимул), попадает в глаз и формирует двухмерное изображение на сетчатке (проксимальный стимул). Пространство между объектами сцены в ее двухмерной проекции на сетчатку напрямую не отражено, но разноудаленные от наблюдателя объекты обретают на этой проекции разный размер.



**Рис. 5.5.9.** Соотношение размера и расстояния. Геометрия перспективной проекции демонстрирует то, что расстояние до объекта можно определить из высоты этого объекта ( $h$ ) и тангенса его зрительного угла ( $\alpha$ ).

верх и низ объекта справа (елка), расположенного своей нижней частью на уровне глаз. Если высоту объекта обозначить  $h$ , а охватываемый им зрительный угол —  $\alpha$ , то расстояние до объекта ( $d$ ) можно выразить через тангенс этого угла:

$$\tan(\alpha) = \frac{h}{d},$$

то есть:

$$d = \frac{h}{\tan(\alpha)}.$$

Отношение высоты объекта к расстоянию до него, то есть угловой размер ретинальной проекции этого объекта ( $\alpha$ ), позволяет нам оценить его удаленность ( $d$ ) от наблюдателя.

Однако имеется и проблема. Чтобы решить вышеприведенное уравнение, должна быть известна фактическая высота объекта ( $h$ ), однако эту информацию невозможно извлечь из изображения на сетчатке, поскольку размер объекта и его удаленность как раз и задают размер проекционного изображения. Следовательно, по данной проекции мы никак не можем определить, почему ее размер именно таков: то ли это проекция мелкого, но близкорасположенного предмета, то ли крупного, но далекого. Существует бесконечное число логически возможных габаритов, которые дадут проекцию данного размера. Если, к примеру, мы заранее знаем, что какие-либо два объекта одинаковы, то, исходя из относительных размеров их проекций, мы можем сказать, что один предмет ближе, а другой дальше. Но, чтобы правило это сработало, объекты обязательно должны быть одинакового размера. Однако мы не можем этого знать, не зная их удаления от наблюдателя. Образуется заколдованный круг. Эвристически зрительная система в силах разорвать его, предположив, что два идентичных по всем остальным параметрам объекта одинаковы и по размеру. Тогда по относительным размерам проекций объектов на сетчатку можно «вычислить» и относительные расстояния до них.

Описанная ситуация — пример того, как бессознательная инференция может помочь определить глубину пространства из относительных размеров изображений на сетчатке. Информация, несомая стимулом, логически недоопределена, но, присовокупив к ней дополнительные эвристические предположения (в данном случае то, что изображения на сетчатке сходны во всем за исключением размеров их проекций, т.е. что это фактически два одинаковых по размеру объекта сцены), приходим к выводу, что объекты эти расположены на разном удалении от наблюдателя. То есть, допустив, что размер объектов одинаковый и взяв за основу разность их проекций на сетчатку, зрительная система делает верное заключение. Не приняв тезис о равенстве габаритов исходных объектов, мы получили бы зрительную иллюзию расстояния.

Блестящий пример того, как размер проекции объекта на сетчатку может послужить источником информации о его удалении от наблюдателя, дан на рис. 5.5.10: у спортсменов примерно один и тот же рост, однако их изображения уменьшаются в размере, потому что стоят ныряльщики на разных расстояниях от точки съемки. Наша зрительная система интерпретирует их рост как более или менее одинаковый, поэтому прогрессивное уменьшение размера мы воспринимаем как прогрессивное удаление.

Еще одним примером того, как из относительных размеров ретинальной проекции зрение извлекает информацию об относительном размере объектов, являются *текстурные градиенты* (о чем вскоре пойдет речь).

### 5.5.5 Сходство размеров

Относительный размер проекции на сетчатку как параметр зрительной стимуляции вполне самостоятелен. Однако габариты многих объектов, относимых нами к той или иной категории, лежат в таком диапазоне, что ощущения наблюдателя оказываются в целом весьма сходны: рост взрослых людей, разумеется, вариативен, но у большинства он лежит между 1.5 и 2 м; столы возвышаются над полом примерно на 60 — 80 см, автомобили примерно на 1.5 м, потолки на 2.5 — 3 м и т.д. Значимость сего обстоятельства в том, что ежели размер объекта заранее известен наблюдателю, то зре-



**Рис. 5.5.10.** Относительный размер как фактор зрительной глубины пространства. Изображение интересно тем, что прочие пикториальные факторы в нем практически отсутствуют, но наше зрение однозначно воспринимает глубину пространства сцены: мы сходу и бессознательно предполагаем, что рост троих ныряльщиков примерно одинаков, но коль скоро их проекции на плоскость изображения (и, соответственно, на сетчатку глаза) уменьшаются в размере по мере удаления от фотокамеры, нашему зрению удастся восстановить соотношения расстояний в сцене.

так, что все три монеты имели одинаковый диаметр и были равноудалены от наблюдателей. Однако же испытуемые воспринимали десятицентовую монету как ближнюю, а полудолларовую — как дальнюю, поскольку «опирались» на их оригинальный размер, хорошо известный им из повседневной жизни.

### 5.5.6 Текстурные градиенты

Еще одним важным свойством перспективной проекции в контексте восприятия глубины пространства является то, что в свое время выдающийся американский исследователь физиологии зрительного восприятия Джеймс Джером Гибсон (1904—1979) окрестил *текстурными градиентами* — это систематически меняющиеся в размере и по форме мелкие текстурные элементы, коих на поверхностях естест-

ние в силах решить приведенное ранее уравнение, соотносящее размер проекции объекта на сетчатку с расстоянием до этого объекта и вывести величину этого расстояния. При этом подчеркнем, что осведомленность о размере объекта — это не осознанное знание, а решение уравнения — не осознанная символическая манипуляция. Напротив — это скоростные неосознаваемые процессы, проходящие автоматически, без какого-либо обдумывания.

Экспериментально эффект сходства размеров в восприятии глубины пространства впервые продемонстрировал американский исследователь Вильям Иттлесон (1951), показавший, что человек может довольно точно оценить расстояние до игровой карты, чья высота (3.5 дюйма) заранее известна зрению — хотя, вероятно, и неосознанно.

Еще один вполне убедительный экспериментальный результат был получен В. Эпштайном (1965), который предъявлял своим испытуемым высококачественные фотографии десятицентовой, двадцатипятицентовой и полудолларовой монет в условиях монокулярного показа в затемненном помещении, где специальным точечным источником были освещены лишь эти изображения. Испытуемых просили оценить расстояние до каждой из монет. Показ был организован



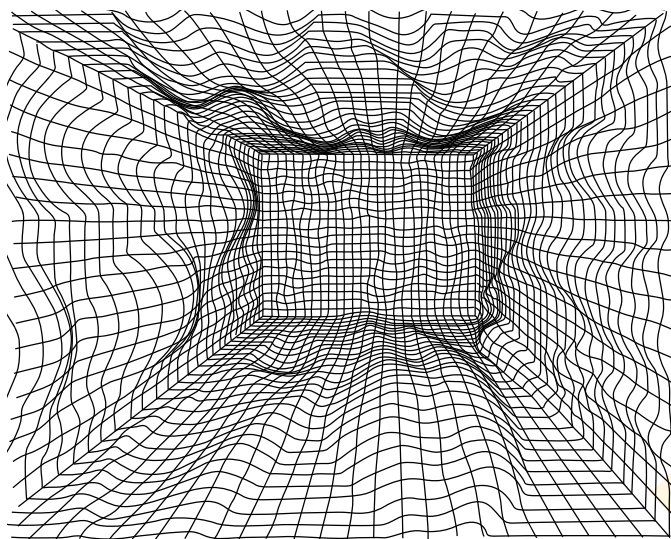


**Рис. 5.5.11.** Естественные текстурные градиенты. Многие поверхности естественного происхождения состоят из текстурных элементов, сходных по габаритам и форме, что позволяет зрительной системе реконструировать глубину пространства из систематических изменений размера и формы проекций этих элементов на сетчатку. В данном примере глубина пространства читается устойчиво, несмотря на то, что кроме текстурных градиентов прочие факторы глубины почти не представлены (особенно в верхнем примере).

венных объектов великое множество. Примеров естественных текстур предостаточно — это и травинки на лугу, галька, ковровая пряжа, узор кафеля в ванной и пр. Три примера даны на рис. 5.5.11, и, глядя на них, мы уже догадываемся, что помимо информации о глубине пространства текстурные градиенты дают наблюдателю сведения об ориентации поверхности в глубину и о ее кривизне. Отметим, что текстурными вариациями можно весьма реалистично изобразить поверхности замысловатых, сложных форм.

Американец Кент Стивенс (1979) показал, что два аспекта текстурных вариаций — размер элементов и их форма — являются независимыми источниками информации об ориентации поверхностей во внешнем мире. Общий размер текстурных элементов





**Рис. 5.5.12.** Искусственные текстурные градиенты, выполненные с помощью специальной компьютерной программы.

По материалам Marr, 1982.

если размеры этих элементов более или менее одинаковы. Если же нет, то исходом станет *иллюзия* глубины и ориентации. К примеру, глубина и изменения в ориентации поверхности, воспринимаемые нами в рис. 5.5.10 и 5.5.11 фактически иллюзорны, поскольку текстурные элементы в самих картинках на самом деле меняются. Когда мы смотрим на реальные сцены, изображениями которых являются эти снимки, то аналогичное предположение (т.е. что все текстурные элементы примерно одного размера) в целом дает веридикальный перцепт. Однако же при рассматривании *изображения* эти же догадки ведут к ощущению глубины, которой фактически нет. То есть, тот факт, что мы воспринимаем глубину пространства на изображениях, несущих в себе текстурированные поверхности, — это следствие того, что, несмотря на ложность предположения о равенстве размеров элементов самой картинки, в целом такое допущение верно в отношении изображенной сцены.

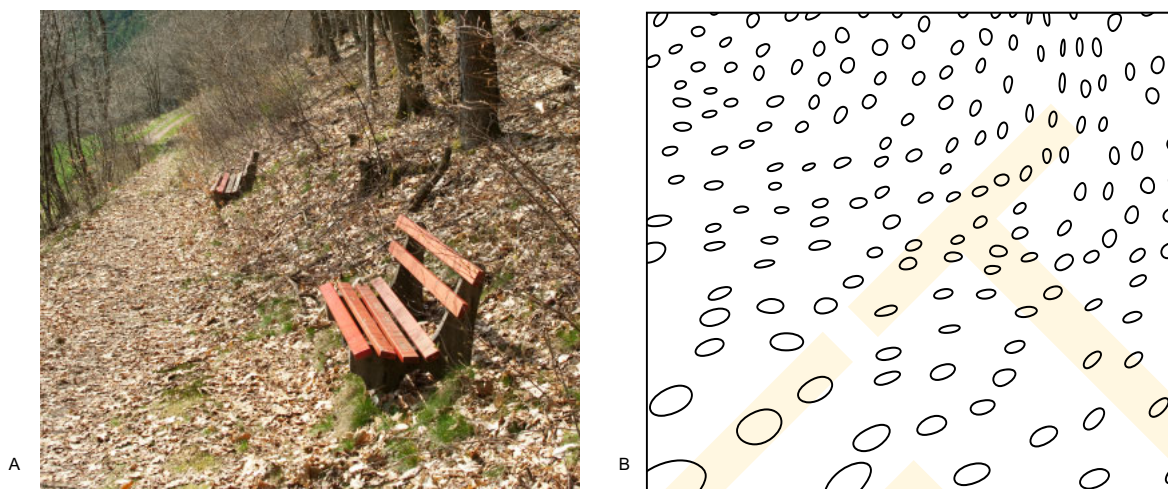
Проекционная форма текстурных элементов может также нести информацию об ориентации поверхности (рис. 5.5.12). Однако вновь эту информацию можно реконструировать только путем дополнительных предположений о фактической форме текстурных элементов в сцене. Чтобы определить ориентацию как самих текстурных элементов, так и поверхности, на которой они расположены, Стивенс (1979) использовал их *пропорции* (*форматное соотношение*, т.е. соотношение длинной стороны к узкой). Его анализ основывался на предположении, что стороны текстурных элементов реальных объектов примерно одинаковы по всем направлениям.

Американский исследователь зрения Джон Кендер (1979) предложил несколько иные алгоритмы диагностики ориентации поверхностей из текстурных форм, сделав слегка иные допущения, а именно, что текстурные элементы максимально регулярны, гомогенны и симметричны.

Эндрю Виткин (1981) озвучил еще один вариант, основанный том предположении, что края текстурных элементов стремятся к т.н. *изотропическому распределению*, то

уменьшается по мере увеличения расстояния благодаря падению по всем трем пространственным размерностям. Следовательно, размер элементов можно использовать в определении относительного расстояния до различных частей текстурированной поверхности и тем самым реконструировать ориентацию этой поверхности. (Однако отметим, что сказанное справедливо только в том случае, если размер текстурных элементов более или менее постоянен.)

Сие, разумеется, пример эвристического допущения в восприятии глубины, поскольку перцепционный вывод о расстоянии до текстурных элементов (основанный на данных, полученных из размеров их проекций на сетчатку глаза) окажется верным, только



**Рис. 5.5.13.** Определение локальной ориентации поверхностей из текстуры. А — естественная сцена, содержащая значимые текстурные данные. В — результат работы виткиновой программы, где форма и размер овалов передают информацию о глубине отдельных участков поверхности.

По материалам Witkin, 1981.

есть к тому, что их контур в различных направлениях примерно один и тот же или как минимум равно распределен по этим направлениям. Сие вполне эвристично, поскольку когда изотропические текстурные элементы рассматриваются под наклоном, их края становятся неизотропичны в плоскости их проекционного изображения. Т.е. они оказываются как бы сплюснуты в определенном направлении (см. рис. 5.5.13).

Однако же эвристические допущения — это палка о двух концах, поскольку алгоритмы на них основанные, дают тяжелые сбои, когда лежащие в их основе допущения, оказываются ложными. То есть, программы Стивенса, Кендера и Виткина начинают ошибаться в имитации перцептов поверхностей с вытянутыми текстурами, такими как, например, трава или древесина. Алгоритмы эти «предполагают», что вытянутая структура элементов изображения возникает только в случае их наклона в глубину, но не потому, что элементы эти сами по себе могут оказаться вытянутыми. Если бы зрительной системе человека была свойственна та же ошибка, это было бы очко в пользу вышеописанной концепции, но очевидно, что в работе зрения человека крайне редко возникают подобные сбои.

Американки Джинтендра Малик и Рут Розенхольц (MIT, 1994) сравнительно недавно предложили весьма любопытное количественное решение текстурной проблемы: они придумали алгоритм восстановления наклона и поворота мелких фрагментов поверхности, основанный лишь на том простом допущении, что на коротких участках поверхности текстурные элементы более или менее инвариантны. То есть, исследователи предположили у всех текстурных элементов примерно одни размер и форма, а отличаются они лишь *смещением* внутри поверхности. Такое допущение хорошо тем, что позволяет без каких-либо ограничений (наподобие изотропии, симметрии и пр.) пользоваться всей визуальной информацией, заключенной в текстуре. Алгоритм основан на поиске оптимальной подгонки параметров ориентации поверхности и кривизны, что позволяет учитывать трансформации в форме и размере у соседних текстур-

ных элементов. Оказалось, что алгоритм Малик — Розенхольц отлично работает в ситуациях, когда описанные выше схемы (Стивенса, Кендера, Виткин) дают сбои.

### 5.5.7 Интерпретация краев

Еще один важный вид пикториальной информации, на основе которого зрение в том числе делает вывод о глубине пространства, — это информация, извлекаемая посредством краевой интерпретации контуров. Простейший пример мы рассмотрели ранее — это *окклюзия*, или, как еще ее называют, — *интерпозиция*: блок световых потоков, отраженных от одного объекта другим — непрозрачным объектом, расположенным ближе к наблюдателю. К примеру, рис. 5.5.14 мы однозначно воспринимаем как квадрат позади круга, частично этим кругом закрытый. Все, что фактически представлено в стимуле, — это некая конфигурация двухмерных областей, ограниченных контурами. Но контуры эти мы воспринимаем еще и как указание на отношения по глубине: круг перед квадратом.

Как источник информации о глубине край имеют сильные и слабые стороны.

Слабая сторона заключается в том, что краевая информация — относительна, а не абсолютна; а также качественна, но не количественна. То есть края информируют наше зрение только о *порядковых* отношениях по глубине: мы видим, что квадрат дальше круга, но при этом не знаем насколько.

Сильная же сторона в том, что краевая информация доступна зрению практически в неограниченном диапазоне расстояний. Единственное требование при этом, чтобы предметы находились в поле зрения и чтобы ближний (кроющий) был непрозрачен.

Итак, вопрос, который мы задаем себе в этом разделе, звучит так: «Откуда в контурах двухмерного изображения на сетчатке берется информация о глубине сцены и каким образом зрение извлекает ее?».

Вспомним концепцию краевой интерпретации в рамках *блочного мира* (*мира из кубиков*), применяемой в машинном зрении. Цель этой концепции — найти оптимальную интерпретацию всех контуров в изображении блочной сцены. Коль скоро контуры формируются на сетчатке несколькими путями, выходной продукт алгоритма контурной интерпретации — это сортировка краев в логике их источников во внешней среде: некоторые образуются в результате окклюзии одной поверхности другой поверхностью, некоторые — в результате стыка поверхностей, третьи — от тени, падающей на поверхность, и т.д. Эти отличия весьма важны в контексте трактовки пикториальной информации в размерность глубины пространства, поскольку некоторые варианты этой трактовки могут оказаться весьма специфичными. К приме-

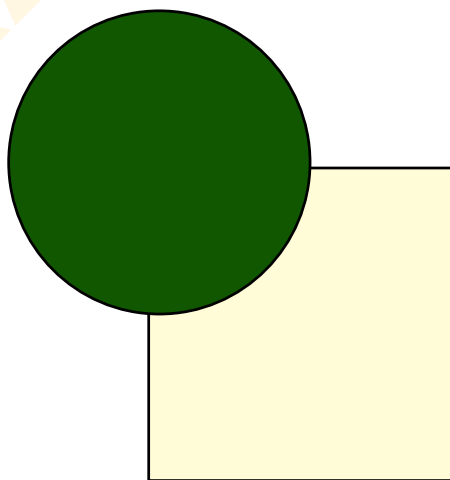
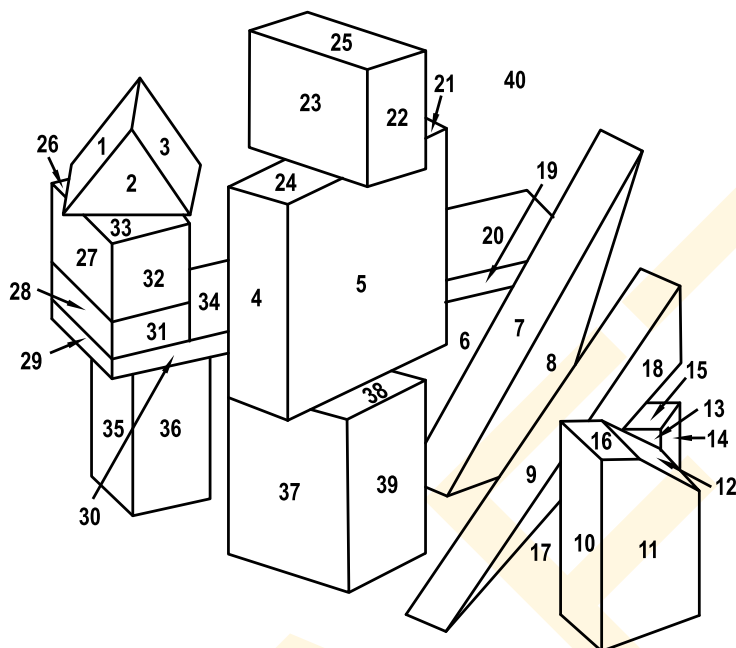


Рис. 5.5.14. Частичная окклюзия (перекрывание) как фактор визуальной глубины пространства. Когда один объект частично закрывает другой, кроющий объект воспринимается как расположенный ближе к наблюдателю.



**Рис. 5.5.15.** Контурный рисунок блочной сцены. На базе информации о линиях и вертексах гуцманова программа группирует участки по принципу подобия. Результаты в целом сходны с биологической перцепцией (например, участки 26, 27, 32 и 33 были сгруппированы в один объект, а участки 19, 20, 29, 30 и 34 — в другой.)

По материалам А. Guzman, 1968.

ру, если край А частично закрывает край В, то А неизбежно окажется ближе к наблюдателю, чем В.

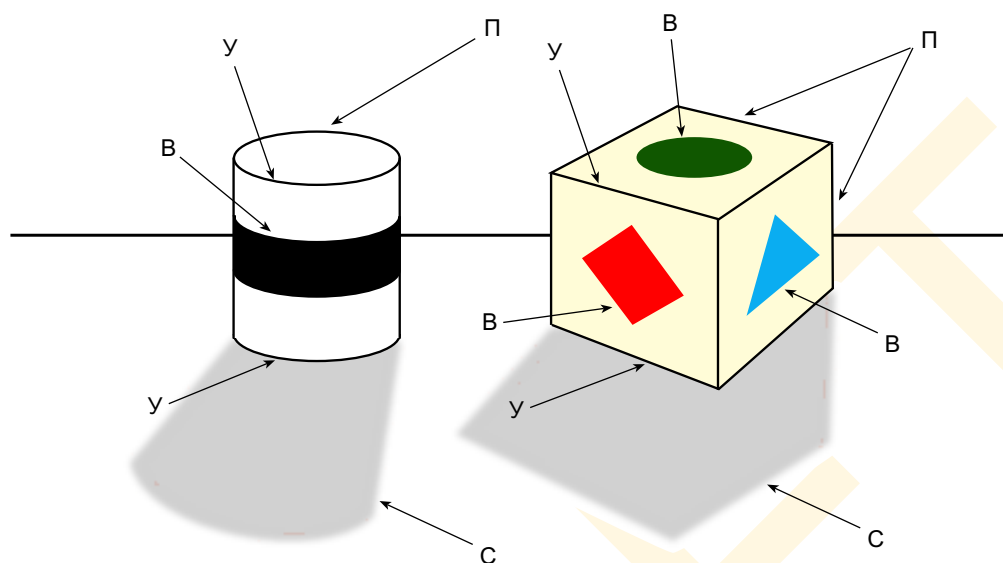
#### 5.5.7.1 Вертекс-классификация

Количественный подход к зрительной интерпретации краев начинается с SEE-программы Адольфо Гуцмана (Эдибнург, 1968, 1969), с помощью которой исследователи зрения пытались интерпретировать простые штриховые рисунки различных блоков, в частности тех, что показаны на рис. 5.5.15. Гуцман исходил из того, что точки пересечения краев критичны в диагностике того, какой край каким закрыт. Ученый разработал схему стандартной классификации краевых пересечений — т.н. *вертексов*: Т, Y, К, X, L и т.д. К примеру, Т-вертексы относятся к наиболее очевидным ситуациям окклюзии: верхушка литеры «Т» соответствует кроющему контуру (краю), а ножка — кроемому. И т.п.

Гуцманова SEE-программа предлагала весьма хитроумную эвристическую интерпретацию краевой информации и при своей более чем скромной теоретической базе работала на удивление эффективно. Впрочем, начало формальному анализу краев в простейших случаях было положено еще до Гуцмана — Дэвидом Хаффманом и М. Клавсом (1971), разработавшими обширный каталог вертексов разных типов, возникающих при рассматривании простых трехгранных углов<sup>1</sup> со всех возможных точек на-

<sup>1</sup> Трехгранные углы — это прямые углы, образованные тремя планарными поверхностями.





**Рис. 5.5.16.** Четыре типа краев. В данной сцене наличествуют яркие края четырех типов: угловые края (У), возникающие из-за смены направления поверхности; пространственные края (П), образующиеся благодаря зазору между поверхностями; вещественные края (В), происходящие из смены коэффициента отражения поверхности; световые края (С) — плод светотени.

блюдения. Они также продемонстрировали то, как локальные ограничения по каждому вертексу и их взаимосвязи между ними сокращают число возможных интерпретаций каждого края.

Британский исследователь А. Вальтц (1972) углубил данные Хаффмана — Клавса, включив в список теневые края, а Малик (1987) применила все наработанное к искривленным объектам. В итоге стало ясно, что для корректной интерпретации краев в любых контурных изображениях простых геометрических тел необходимы намного более сложные алгоритмы.

### 5.5.7.2 Четыре типа краев

Итак, мы начали с предположения о том, что на ранних этапах зрительного акта на сетчатке формируется оптическое изображение, в котором контуры — то есть выраженные перепады фотометрической яркости — четко идентифицированы. Проще говоря, что входные данные алгоритма краевой интерпретации — это идеальное контурное представление всех краев сцены. Наша задача состоит в том, чтобы интерпретировать их в логике экологической ситуации, породившей эти края.

Вначале выделим четыре основных варианта возможной краевой интерпретации:

**1. Угловые края.** Относятся к участкам сцены, в которых имеет место смена положения (ориентации) поверхности. Такие края образуются, когда две поверхности разных ориентаций стыкуются в пространстве. Обычно это собственные края каких-то объектов (например, кубика) или линия соприкосновения предметов (например, кубика, стоящего на столе). Примеры угловых краев даны на рис. 5.5.16 (отмечены «У»).

**2. Пространственные края.** Относятся к тем участкам сцены, где имеется разнос поверхностей по глубине, то есть когда одна поверхность закрывает другую, но при



этом находится перед нею на некоем расстоянии. Если поверхности соприкоснутся, то пространственный край превратится в край угловой. Пример пространственных краев на рис. 5.5.16 с меткой «П».<sup>1</sup>

**3. Световые края.** Образуются там, где имеется отличие в количестве света, падающего на однородную (равномерно окрашенную) поверхность. Световые края могут появиться там, где на поверхность падает тень, где есть рефлекс от глянцевых поверхностей, направленные световые пятна (скажем, на театральной сцене) либо же при смене ориентации поверхности. Примеры световых краев даны на рис. 5.5.16 с меткой «С».

**4. Вещественные края.** Результат изменения в отражательной способности поверхности. Вещественные края образуются тогда, когда поверхности состоят из разных материалов или по-разному покрашены. Наиболее простой пример — полоса, проведенная маркером на листе бумаги. Примеры вещественных краев даны на рис. 5.5.16 с меткой «В».

Напомним, что мы уже обсуждали световые и вещественные края в гл. 3, когда говорили о хроматической константности и константности светлоты. Нынче же мы сосредоточимся на краях угловых и пространственных, поскольку края эти вводят строгие ограничения в интерпретацию глубины воспринимаемой сцены.

### 5.5.7.3 Подвиды краев

Угловые и пространственные края в объектах с плоскими поверхностями — суть взаимное исключение. Если контур на сетчатке — это плод соприкосновения двух поверхностей разной ориентации в реальной сцене, — сие угловой край; если край образован за счет того, что одна поверхность закрывает другую, но с неким пространственным зазором, — это край пространственный. Следовательно, в контурном представлении каждый край явится либо угловым краем, либо пространственным и может быть однозначно отнесен либо к первому, либо ко второму. Цель концепции краевой интерпретации — исследовать процесс, приводящий каждый контур оптического изображения на сетчатке в соответствие фактическим краям реальной сцены.

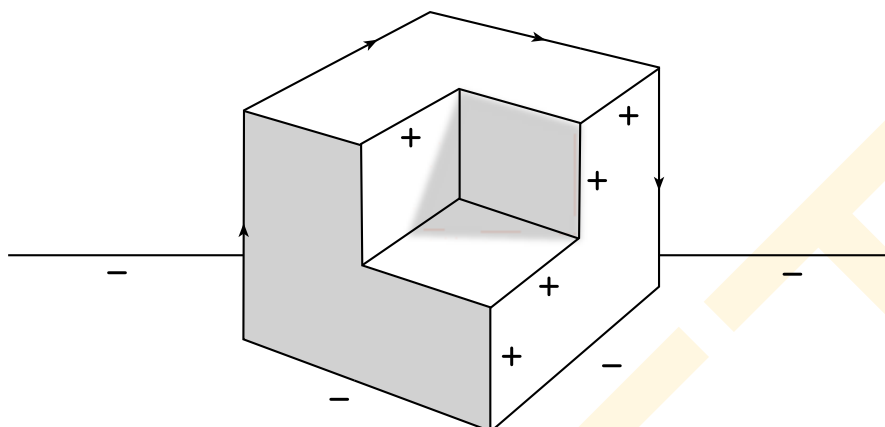
При этом нам понадобится деление угловых и пространственных краев на два подвиды.

**Выпуклые и вогнутые угловые края.** Угловые края делятся на два подвиды — *выпуклые* и *вогнутые* — и несут важную информацию о глубине края относительно поверхностей.

*Выпуклые угловые края* появляются тогда, когда две поверхности, соприкасаясь, образуют заполненный объем, соответствующий двугранному углу меньшему  $180^\circ$ . Краевой угол направлен *к* наблюдателю, как, скажем, угол куба снаружи. Выпуклые углы показаны на рис. 19 линиями, отмеченными «+».

*Вогнутые угловые края* появляются тогда, когда две поверхности, соприкасаясь, образуют заполненный объем, соответствующий двугранному углу большему  $180^\circ$ . Краевой угол направлен *от* наблюдателя, как внутренние края полого куба при рассмотрении его изнутри. Вогнутые края отмечены на рис. 5.5.17 знаком «—».

<sup>1</sup> Отметим, что угол между гранью куба и поверхностью, на которой он находится, — это угловой край, но не пространственный, даже несмотря на то, что поверхности принадлежат разным объектам и что поверхность стола частично закрыта этим кубом. Причина в том, что между кубом и столом нет пространства — они *соприкасаются*, образуя угол.



**Рис. 5.5.17.** Выпуклые и вогнутые угловые края. Выпуклые края (помечены «+») появляются, когда две поверхности образуют внутренний угол меньше  $180^\circ$ . Вогнутые края (помечены «-») возникают, когда поверхности образуют угол больше  $180^\circ$ . Стрелки указывают на то, что ближняя к наблюдателю поверхность, образующая пространственный край, — справа.

В простых трехгранных планарных объектах каждый угловой край — либо выпуклый, либо вогнутый. Но никогда вогнуто-выпуклый.

**Правые и левые пространственные края.** Пространственные края также делятся на две подвиды: у одних кроющая поверхность находится на одной стороне угла, а у вторых — на другой. Пространственные края помечены на рис. 5.5.17 стрелками, направление которых подчиняется правилу правой руки: кроющая (ближняя поверхность) всегда расположена справа, а кроемая (дальняя) — слева. Эти два возможных варианта деления пространственных краев — взаимоисключающие, поскольку кроющий контур может располагаться только с одной стороны. Отметим, что правильное деление несет очень важную информацию о глубине, т.к. устанавливает, какая поверхность ближе, какая дальше.

Итак, до сих пор мы пользовались четырьмя возможными вариантами обозначения краев в контурном представлении, содержащем только угловые и пространственные края. Это значит, что если в контурном представлении имеется  $n$  краев, то существует  $4^n$  логически возможных их обозначения, что соответствует  $4^n$  качественно отличным интерпретациям.<sup>1</sup> Как видим, получается астрономически большое число даже в отношении простейших сцен. Например: на рис. 5.5.17 всего 20 краев, но сие простейшее контурное представление предполагает 1048576 логически возможных вариантов их интерпретации! В данном случае, однако — и во многих других случаях — наше восприятие выбирает лишь один. Как же сие возможно и за счет чего множество логически возможных интерпретаций сокращается до одной единственно верной?

#### 5.5.7.4 Физические рамки

Хаффман и Клавс построили свой анализ на критично важном предположении, что не все логически возможные интерпретации возможны физически. Ученые оценили ло-

<sup>1</sup> Каждый край имеет четыре варианта интерпретации. Поэтому, когда в наличии один край, то имеют место четыре варианта его интерпретации; когда два — 16 вариантов и т.д.

кальные ограничения на вертексах трехгранных объектов — объектов, углы которых образованы тремя поверхностями — и обнаружили, что лишь малая доля логически возможных интерпретатов реализуема физически.

Для примера рассмотрим набор возможных «стрелочных» вертексов, т.е. таких, в которых три края сходятся в плоскости изображения под углом меньшим  $180^\circ$ . Поскольку существуют четыре варианта интерпретации края (т.е. как выпуклый или вогнутый угловой, или как правый или левый кроющий), имеют место  $4^3 (=64)$  логически возможных варианта интерпретации стрелочного вертекса. Однако Хаффман и Клавс показали, что физически возможны лишь три варианта интерпретации. Аналогичное сокращение (с 64-х до 3-х) касается Y-вертексов. Хаффманов каталог возможных физических источников трехгранных углов в контурном ретинальном представлении дан на рис. 5.5.18.

Хаффман и Клавс показали также, что имеют место и дальнейшие ограничения краевой интерпретации, но на более общем уровне. Исследователи пришли к такому выводу на том основании, что в рамках ограничений, свойственных блочному миру (миру многогранников с планарными плоскостями) интерпретация каждого края по всей его длине постоянна: выпуклые углы не могут стать вогнутыми, а правокроющие — левокроющими (если только вертексы между краями не разрешат смену интерпретации). Приняв такой принцип, мы еще сократим количество возможных интерпретатов.

На рис. 5.5.19 дан простой пример того, как сие логическое ограничение работает в отношении тетраэдра (четырёхгранника):

1. Рассмотрим все физически возможные варианты стрелочного вертекса А. Из хаффманова каталога на рис. 5.5.18 мы знаем, что таковых всего три.

2. Затем рассмотрим то, как эти три варианта взаимодействуют со всеми возмож-

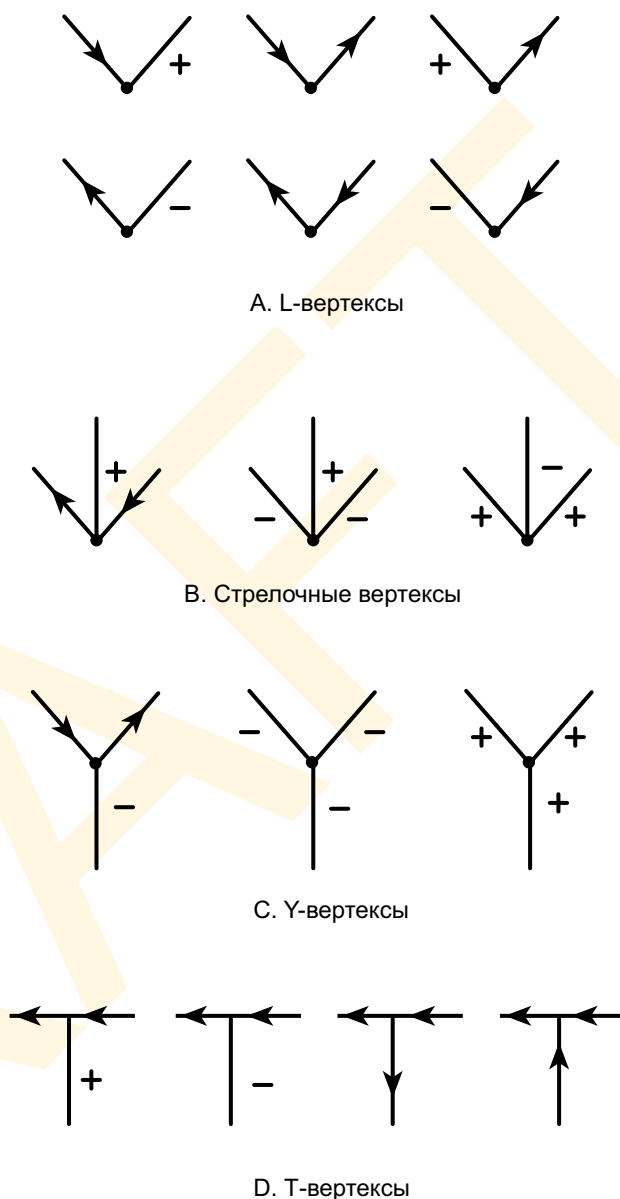
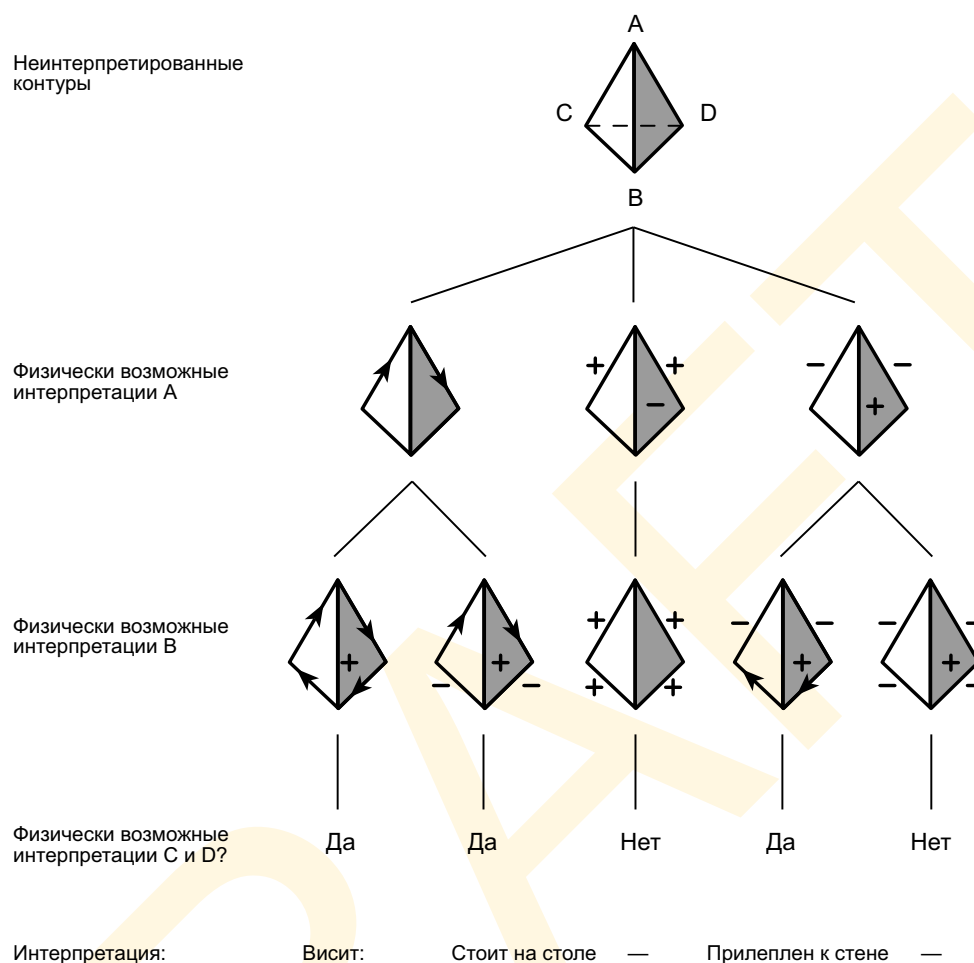


Рис. 5.5.18. Типы физически возможных вертексов на трехгранных углах.



**Рис. 5.5.19.** Интерпретация краев четырехгранника. Простой четырехгранник (вверху) вначале можно интерпретировать по вертексу A во все физически возможные варианты, затем по вертексу B и, наконец, исключив невозможные варианты по вертексам C и D. Остаются лишь три физически возможных варианта интерпретации.

ными вариантами интерпретации стрелок вертекса B: из девяти логических возможных вариантов четверо исключаются по физическим причинам, поскольку потребуют двух отличных интерпретаций срединного края.

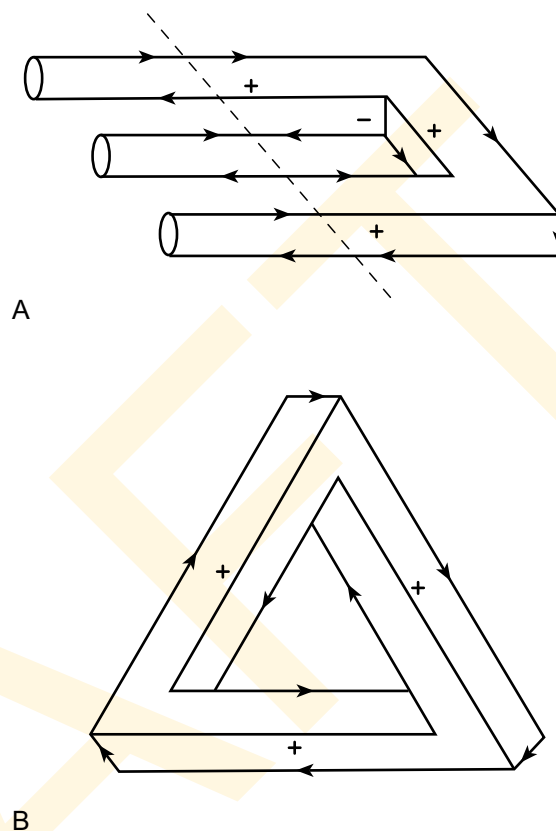
3. Наконец убедимся, действительно ли результирующие пять вариантов интерпретации вертексов A и B дают физически возможные варианты по L на вертексах C и D. Эти ограничения исключают еще два варианта, оставляя лишь три физически возможных варианта интерпретации.

Каковы же эти три варианта? Перцепционно приоритетной будет интерпретация нижних краев как вогнутых, то есть, как если бы тетраэдр стоял своей нижней гранью (BCD) на поверхности стола. Альтернативный вариант интерпретации верхних краев как вогнутых соответствует перцепту тетраэдра, как бы прилепленного к стенке своей задней гранью (ACD). Как кроющийся пространственный край периметр фигуры воспринимался бы, если тетраэдр висел бы в воздухе.

Отметим, как из-за введения физических ограничений сокращается множество

возможных интерпретаций: мы начали с рисунка из пяти контуров, который мог иметь  $4^5$  ( $=1024$ ) логически возможных интерпретата. А после введения в анализ ряда физических ограничений их число упало до трех! Подчеркнем, что анализ сей не позволяет сделать выбор между тремя возможными вариантами интерпретации, и поэтому чтобы выбрать окончательный вариант, придется и далее вводить ограничения.

Отметим также, что возможных интерпретаций может и не существовать. Сие обычно касается т.н. *невозможных фигур*, в частности таких, как изображения, о которых шла речь в разделе 1.1. На первый взгляд эти рисунки выглядят как изображения реальных объектов, но при этом физически они совершенно нереальны (Penrose & Penrose, 1958). К примеру, невозможную фигуру, показанную на рис. 5.5.20 А, нельзя однозначно интерпретировать, поскольку она требует того, чтобы кроющие края одной стороны выступали в роли угловых краев на другой стороне. Однако существует ряд невозможных фигур с вполне однозначной интерпретацией краев, в частности треугольник Пенроуза, показанный на рис. 5.5.20 В.



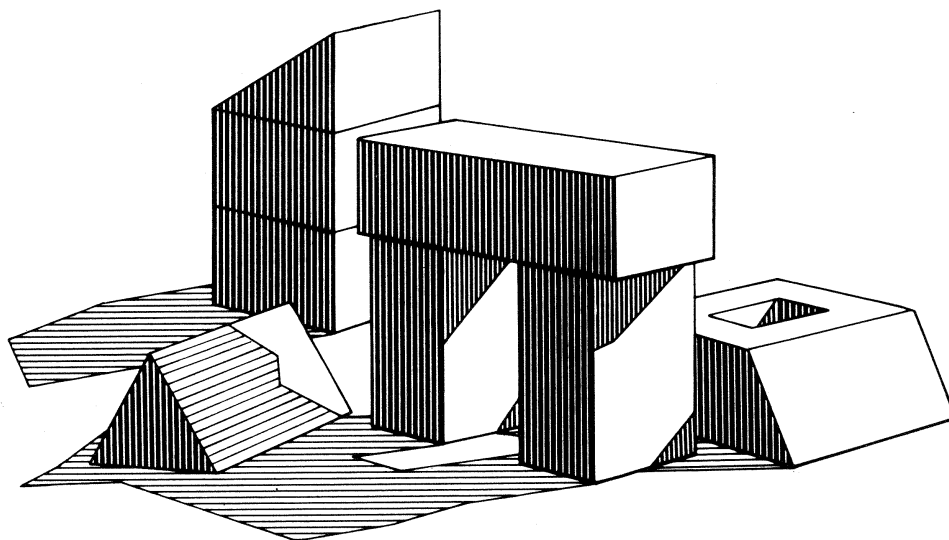
**Рис. 5.5.20.** Интерпретация краев невозможных фигур. «Трезубец» (А) не имеет последовательной интерпретации, но у треугольника Пенроуза (В) она есть.

#### 5.5.7.5 Расширения и обобщения

Исследование Хаффмана — Клавса явилось концептуальным прорывом в теории краевой интерпретации. Однако Дэвид Вальтц (1975) — аспирант Массачусетского технологического института — углубил и дополнил данные Хаффмана — Клавса, включив в них 11 типов краев, в том числе тени и «трещины» (линии, которые являются результатом точной пригонки копланарных поверхностей: главным образом угловые края на  $180^\circ$ ). Сия экспансия пополнила каталог физически возможных вертексов до нескольких тысяч наименований, но, как мы уже догадываемся — и его придется сократить. К примеру, сложная контурная картина, показанная на рис. 5.5.21, имеет лишь одну физически возможную интерпретацию, которую человек, собственно, и воспринимает. Как видим, дополнительные ограничения в вальтцевом анализе идут от углов, отбрасывающих тени.

Однако, несмотря на явный прогресс, даже вальтцев алгоритм не дотягивает до уровня зрительной системы, бо применение его ограничено планарными многогранниками. Алгоритм не работает по искривленным плоскостям или по объектам, состоя-





**Рис. 5.5.21.** Блочная сцена с тенями. Вальтцев алгоритм дает только один вариант ее интерпретации — тот же, что и зрительная система человека. Добавка теней повышает точность интерпретата, поскольку вводит дополнительные ограничения.

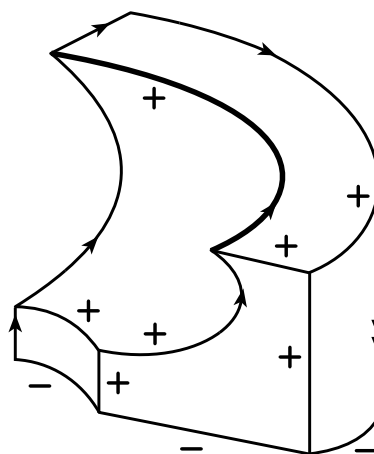
По материалам Waltz, 1975.

щим из тонких листов (таким, как пачка писчей бумаги). При этом ни что из перечисленного не представляет серьезной проблемы для зрительной системы человека, способной интерпретировать объекты любой сложности и любого типа. С другой стороны, вальтцева программа поражает своей способностью давать физически возможные интерпретаты, некоторые из которых даже человеку не под силу без явных подсказок.

До уровня искривленных объектов наше понимание краевой интерпретации продвинула американка Джинтендра Малик (университет Беркли, США). Ее анализ потребовал выделения нового вида пространственных краев, именуемого *сверхкраем* или *лимбом*. Лимб образуется тогда, когда искривленная поверхность частично закрывает сама себя (рис. 5.5.22). Лимбы мы отмечаем двойными стрелками, опять же по правилу правой руки.

Искривленные объекты весьма усложняют алгоритм краевой интерпретации, поскольку физически один и тот же край может по-разному толковаться на различных своих участках. Пространственный край, к примеру, может перейти в выпуклый угловой край, что четко видно по верхнему контуру объекта на рис. 5.5.22.

Однако зрение человека совершеннее и малик-анализа: мы легко и вполне адекватно воспринимаем листья комнатного растения,



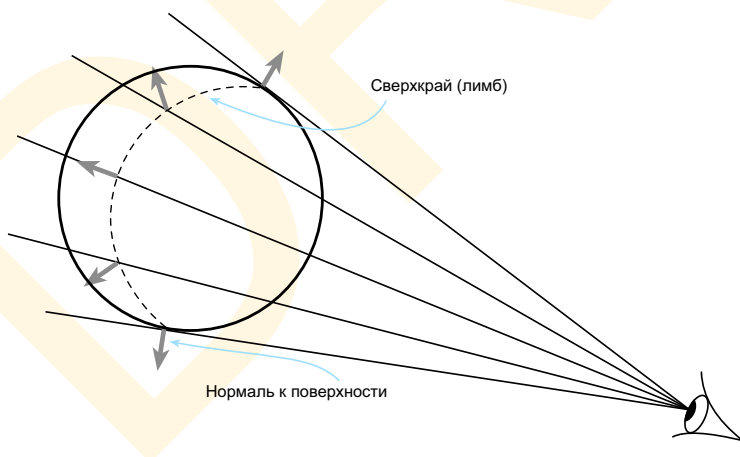
**Рис. 5.5.22.** Интерпретация краев искривленных объектов. Переход пространственного края в выпуклый угловой отмечен утолщенной линией; лимб (сверхкрай), образуемый искривленными поверхностями, кроющими сами себя, — двойной стрелкой.

изображенного на рис. 5.5.23, как имеющие весьма специфическую форму в трехмерном пространстве. Каким же образом можно восстановить информацию из такого изображения?

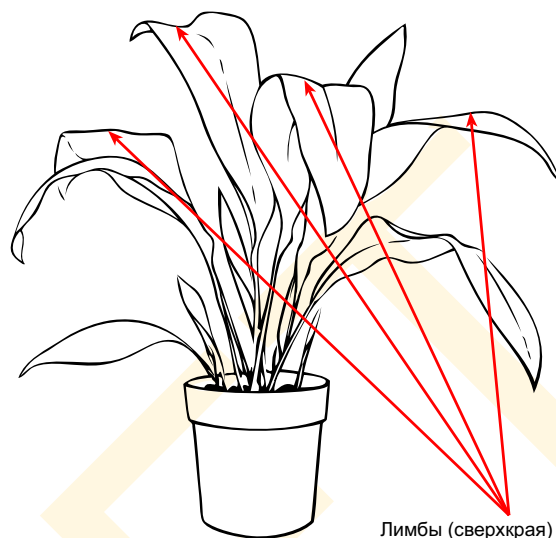
Американцы Гарри Барроу и Джей Тенненбаум (1978) обнаружили любопытный факт, что когда кроющийся край образован гладкой, самокроющей поверхностью, то есть представляет собой лимб (сверхкрай) — ориентацию поверхности вдоль края можно восстановить очень точно: в каждой точке лимба плоскость поверхности перпендикулярна как зрительной оси, так и касательно к этой поверхности (рис. 5.5.24). Эти два момента описывают ориентацию поверхности на всем протяжении сверхкрая (что отмечено нормальными к поверхности, т.е. векторами, нанесенными перпендикулярно к поверхности по краю).

По другим кроющим контурам ограничения слабее, и поэтому дабы найти однозначное решение, требуются дополнительные ограничения. К таковым, работающим на идею Барроу — Тенненбаума, относятся *плавность* и *общая точка наблюдения*.

1. *Гипотеза плавности.* Предполагает то, что ежели кроющийся контур в ретинальном изображении гладок, то гладок и контур поверхности, образовавшей этот край. Сие не всегда верно однако, поскольку это — эвристическое допущение. То есть, порой случается так, что поверхность с прерывистым краем мы рассматриваем под определенным углом, под которым ретинальная проекция этого края оказывается плавной.



**Рис. 5.5.24.** Реконструкция ориентации поверхности из сверхкрая. Нормали к поверхности по сверхкраю перпендикулярны зрительной оси, равно как и касательной к этой поверхности. Сие однозначно характеризует ориентацию поверхности повсему краю.



**Рис. 5.5.23.** Пример естественных лимбов (отмечены красными стрелками). Зрение человека в данном рисунке легко воспринимает форму листов фикуса, тогда как блочные алгоритмы не в силах это сделать.

По материалам Barrow & Tennenbaum, 1981.

2. *Гипотеза общей точки наблюдения.* Предполагает то, что точка наблюдения является общей в том смысле, что небольшие изменения в ней не вызывают качественных отличий в ретинальном изображении, в частности, таких как появление скрытых неравномерностей. Предположение об общей точке наблюдения блокирует

возможность возникновения «замаскированных» неравномерностей, в частности тех, о которых шла речь чуть выше.

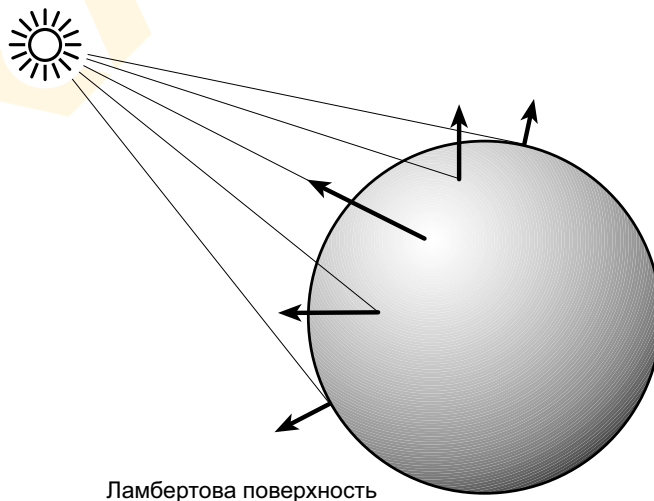
Приняв сии условия, Барроу и Тенненбаум показали, что можно реконструировать трехмерную форму кроющей поверхности с наибольшей равномерной кривизной и минимальным перекосом.

Таким образом, сверх- и прочие кроющие контуры дают весьма специфическую информацию об ориентации искривленных поверхностей на их краях, но что же происходит между ними? Барроу и Тенненбаум предположили, что имеет место процесс интерполяции, реализуемый за счет «оптимально стыкуемой» поверхности, проходящий через известные точки и минимальный по энергозатратам. Такой процесс характеризуется метафорой поиска тонкой мембраны, идущей через известные точки и требующей минимума усилий для удержания своей формы. И вновь допущения, имеющие место в такой концепции, во-первых, максимально приближают ситуацию к реальной (о чем и толковал Гельмгольц и прочие конструктивисты), во-вторых — дают максимум вероятности в некотором вполне определенном смысле (позиция гештальтистов).

Наконец отметим то, что количественный анализ, проводимый исследователями машинного зрения, хоть и идентифицирует информацию, извлекаемую из двухмерных изображений (с целью интерпретации краев в трехмерной сцене), но при этом не дает ответа на вопрос о том, как фактически работает зрительная система человека? Удивительно и печально то, что прогнозы, построенные на этом анализе, почти никто не проверял в контексте зрения человека — сосредоточились в основном на формальной структуре алгоритмов и их эффективности. Безусловно, в смысле собственно *количественного* описания информации, на которой должна основываться интерпретация краевой информации — это гигантский шаг вперед, но как алгоритмическое описание *фактических процессов*, происходящих в зрительной системе — «детский лепет».

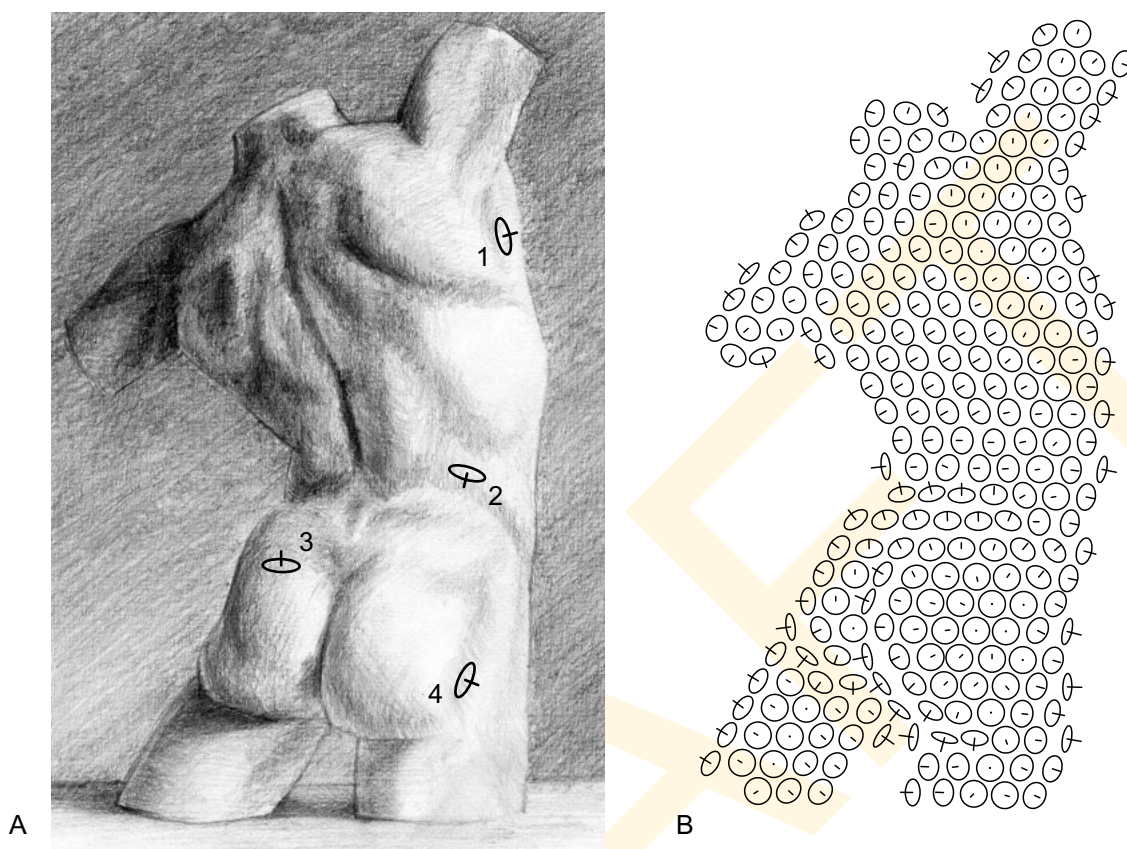
### 5.5.8 Светотень

Еще одним весьма эффективным источником информации о расположении поверхностей в глубину пространства является т.н. *светотень (шейдинг)* — колебания количества света, отраженного от поверхности, как результат вариаций положения поверхности относительно источника этого света. Чтобы интуитивно оценить характер такой информации, рассмотрим шар (рис. 5.5.25), «изготовленный» из однородного матового материала, равномерно отражающего свет по всем направлениям (т.н. *ламбертова поверхность*) и освещенного единичным



Ламбертова поверхность

**Рис. 5.5.25.** Реконструкция ориентации поверхности из светотени. Паттерн яркостей световых потоков, отраженных от матовой поверхности, дает зрению информацию о локальном наклоне и локальном повороте этой поверхности.



**Рис. 5.5.26.** Изучение восприятия ориентации поверхности из светотени. А — испытуемые смотрели на изображение мужского торса и меняли форму и ориентацию курсора так, чтобы «положить» ее на поверхность (1 и 2). 3 и 4 демонстрируют явное «непопадание». В — результат работы одного из испытуемых.

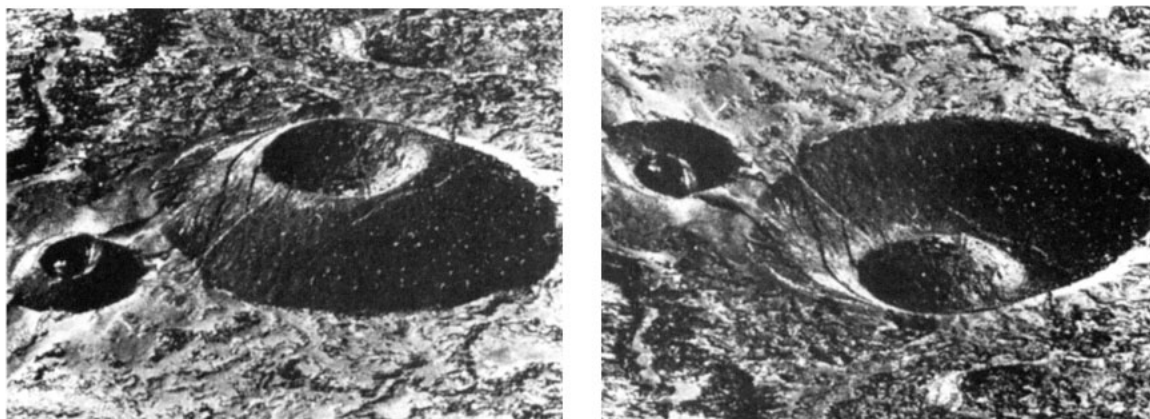
По материалам Koenderink, van Doorn & Kappers, 1996.

точечным источником, удаленным на некоторое расстояние. Мы видим, что наиболее освещенная часть поверхности шара — это место, на которое свет падает по нормали к этой поверхности. По мере отклонения от перпендикуляра количество отраженного света снижается, создавая некий паттерн яркостей. Последний и поставляет зрению информацию о структуре и положении освещенной поверхности.

#### 5.5.8.1 Оценка ориентации

Способность человеческого зрения к реконструкции из яркостных перепадов ориентации поверхности и ее удаления от наблюдателя активно изучалась в конце прошлого века голландцами Жаном Кёндеринком, Андреа ван Дорн и Астрид Капперс (Утрехтский университет, 1992, 1996). Исследователи предъявляли испытуемым изображение человеческого торса (рис. 5.5.26), а индикация направлений осуществлялась за счет довольно плотного сэмпинга позиций на этой поверхности. Наблюдатели могли указывать направление с помощью специального экранного курсора, напоминавшего перевернутую канцелярскую кнопку — овальную "плоскость" с коротким шипом (рис. 5.5.26). Курсор использовался для указания направления поверхности, потому что легко воспринимается как круг, направленный в глубину с определенными





**Рис. 5.5.27.** Направление освещения и восприятие выпуклости/вогнутости. Фигуры на левой картинке смотрятся бугорками, на правой — выемками, поскольку зрительная система всегда предполагает, что источник света сверху. Если мы перевернем страницу, картина изменится на противоположную: выемки справа (переехавшие налево) «превратятся» в бугорки и наоборот.

ми наклоном и поворотом (аналогичная кнопка присутствует на рис. 5.0.1 для демонстрации отличий между наклоном и поворотом). Испытуемым объяснили, как управляться с этим курсором, чтобы он выглядел кружком, лежащим на поверхности объекта с перпендикуляром к этому кружку (рис. 5.5.26 А и В).<sup>1</sup> Рис. 5.5.27 продемонстрирует нам усредненный пример нескольких наблюдательских сессий. Любопытно, что такой способ представления данных, т.е. в виде текстурного градиента из множества кружков аппроксимирует внешний вид оригинального объекта (торса).

Опыт позволил экспериментаторам реконструировать восприятие человеком глубины поверхности и ее ориентации с минимумом допущений. Исследователи обнаружили, что у разных наблюдателей на удивление схожи качественные визуальные оценки поверхностей, но существенно отличаются количественные, т.е. *степень* воспринимаемой ими глубины.

Другим важным выводом явилось то, что зрительная система в своем отклике не использует строго локальной информации, но интегрирует данные по крупным областям поверхности объекта.

Оказалось также, что результаты поставленных экспериментов не зависят от узнаваемости предмета, потому что ровно те же реакции испытуемые выдали по абстрактным, неидентифицируемым предметам.

Подобно анализу прочих факторов визуальной глубины пространства и ориентации поверхностей, зрительный анализ светотени в основном строится на эвристических предположениях. И, вероятно, в первую очередь наш мозг имплицитно (неявно) исходит из того, что свет всегда падает сверху. На рис. 5.5.27 дан пример, иллюстрирующий этот тезис: изображение слева воспринимается как два конуса с ямками вместо вершущи; изображение справа — как воронки с выпуклым дном. Фактически такое восприятие веридикально только тогда, когда свет падает сверху (услов-

<sup>1</sup> Варианты С и D даны как пример того, когда кнопка не воспринимается кружком, лежащим на поверхности.



ный источник расположен сверху, но не изображен). Убедиться в этом очень легко, перевернув страницу вверх ногами: сразу же «перевернется» и наше предположение о направлении «освещения», инвертировав воспринимаемую выпуклость/вогнутость. Версия о свете, падающем всегда сверху, критична, потому что естественные сцены всегда освещены именно сверху. Следовательно, перед нами еще один пример использования зрительной системой скрытых догадок.

### 5.5.8.2 Количественный анализ Хорна

Несмотря на то, что качественно отношения между светотенью и ориентацией поверхности описаны очень давно и веками служили живописцам, только лишь в 1975 г. Бертхольд Хорн (Массачусетский технологический институт, 1975, 1977) первым подключил их к решению задач машинного зрения. Исследователь показал, что колебания фотометрической яркости в ретинальном изображении прямо пропорциональны изменениям в ориентации поверхности.<sup>1</sup> Информация эта недостаточна для реконструкции ориентации поверхности, но лишь позволяет определить угол между потоком падающего света и нормалью к поверхности. Чтобы определить абсолютную ориентацию поверхности, следует позиционировать т.н. *опорные точки* — точки известной ориентации.

Вспомним, что мы как-то уже обсуждали условия, в которых можно определить абсолютное положение поверхности — по свехкраям, как в случае границ шара, где ориентация той или иной локали поверхности перпендикулярна как зрительной оси, так и касательной к поверхности. Второе условие касается т.н. *теневого границы* между освещенными и теневыми участками (рис. 5.5.28). В данном случае заранее известно, что нормаль к поверхности перпендикулярна направлению падающего света. Хорн показал, что при условии использования неких дополнительных ограничений, положение ламбертовых поверхностей можно восстановить. Позже профессор MIT Алекс Пентлэнд (1989) продемонстрировал, что сие выполнимо силами биологически подобных рецептивных полей, наподобие тех что обнаружены в V2 (см. раздел 4.3.3).

Отметим однако, что текущие алгоритмы хорошо подходят к ситуациям четко определенных типов, в частности таких, как рассматривание ламбертовых поверхностей при точечном источнике. Однако и в этом случае способность зрительной системы человека извлекать информацию о глубине и ориентации поверхностей из колебаний яркости многократно превосходит таковую у машинных алгоритмов. Сие очевидно пото-

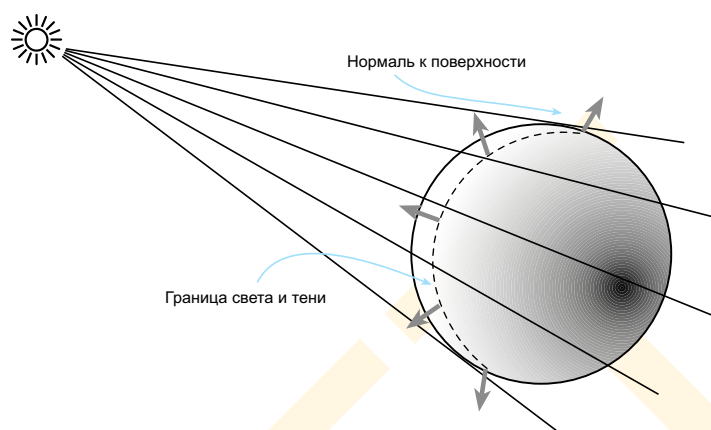
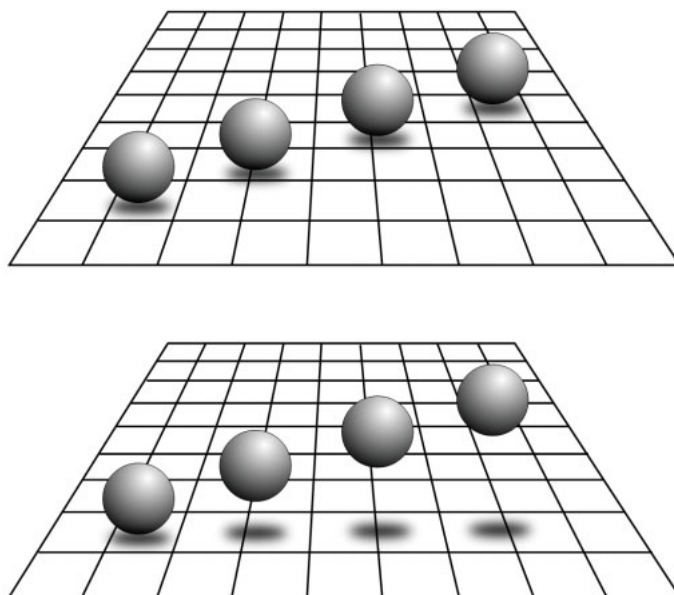


Рис. 5.5.28. Восстановление ориентации поверхности из теневых границ. Анализ Хорна показал, что нормаль к поверхности однозначно восстанавливается вдоль границы света и тени, поскольку известно, что нормаль в этих точках всегда перпендикулярна направлению падающего света.

<sup>1</sup> Вернее сказать, колебания яркости пропорциональны косинусу наклона поверхности.



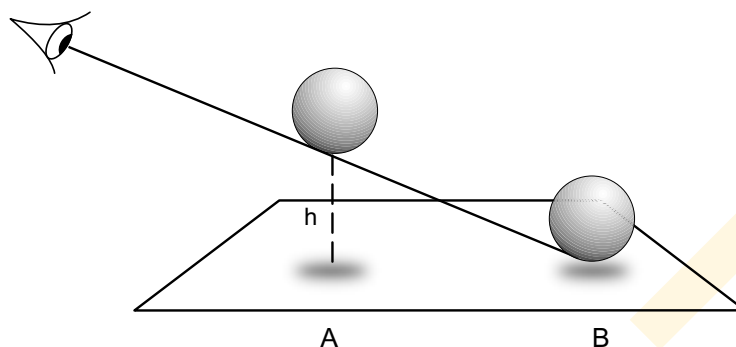
**Рис. 5.5.29.** Демонстрация влияния характера тени на ощущение глубины: верхний и нижний рисунки отличаются только положением теней от шариков.

му, что зрение человека эффективно справляется с такими задачами в несравнимо большем диапазоне условий, например, в условиях нормального освещения, когда имеется множество источников, в том числе рефлекс от соседних поверхностей, немислимо усложняющие математический анализ. Зрение человека свободно обходится с неламбертовыми поверхностями, имеющими в том числе зеркальные характеристики, в частности глянцевые и полуглянцевые, отражающие свет много когерентнее матовых. Наука до сих пор не знает, за счет чего зрение человека способно использовать такую несоизмеримо более сложную информацию при реконструкции формы поверхностей из картины яркостных колебаний, но понято и доказано Кёндеринком и Ван Дорн, что использует.

### 5.5.8.3 Характер теней

Очередным источником информации о глубине пространства является характер наведения тени одного объекта на поверхность другого. На рис. 5.5.29 дан пример: пространственные позиции четырех шариков и текстурированной поверхности на первом и втором полях совершенно одинаковы — отличие состоит лишь в положении теней (Kersten, Knill, Mamassian & Bulthoff, 1996). На рис. 5.5.29 А тени «приклеены» к низу шариков, указывая на то, что последние лежат на поверхности. В этом случае они смотрятся диагонально уходящими в глубину — крайний правый явно дальше, чем крайний левый. Однако на рис. 5.5.29 В тени по нарастающей убегают от низа шариков, указывая на разную их высоту над поверхностью. Теперь они воспринимаются диагонально по высоте, но при этом в одной вертикальной (фронтальной) плоскости: крайний правый выше всех, левый — ниже.

Таким образом, характер теней, отбрасываемых одними объектами на поверхность других, несет в себе информацию об удалении этих объектов от наблюдателя: позиция проекции шарика на сетчатку (ретинальная позиция) задает направление от наблюда-



**Рис. 5.5.30.** Визуальная геометрия сцены из рис. 15. Высота шарика над его тенью ( $h$ ) дает зрению информацию о его высоте над плоскостью, на которую падает тень, характеризуя тем самым позицию шарика на зрительной оси.

теля (зрительная ось на рис. 5.5.30), а расстояние между шариком и его тенью на поверхности дает зрению информацию о высоте низа шарика над поверхностью (которое равно нулю, когда шарик лежит на поверхности и не равно, когда он приподнят над нею). Этих данных достаточно для зрительной диагностики разного расположения предметов относительно наблюдателя.

### 5.5.9 Воздушная перспектива

Воздушная (атмосферная) перспектива относится к прогрессивным изменениям в яркостном контрасте и хроматике объектов, когда те оказываются удаленными от наблюдателя на большое расстояние. Каждый из нас замечал, что здания и прочие крупные предметы становятся все более «замутненными» и «выбеленными» по мере удаления от них. Действительно, их контраст снижается за счет увеличения слоя атмосферы между нами и этими объектами, поскольку воздух содержит рассеивающие свет частички пыли и воды (туман, дымка), к которым порой примешиваются и различные загрязняющие агенты (например, смог от промышленных выбросов и автомобилей). Частички эти добавляют шум в процесс формирования изображения на сетчатке, эффективно размывая последнее. В результате контуры изображения становятся менее проработанными («мутными») и с меньшим контрастом («выбеленными»). Пример находим на рис. 5.5.31.

Мы также знаем, что, например, поросшие лесом горы на большом удалении выглядят синеватыми или даже слегка фиолетовыми, хотя вблизи те же самые горы смотрятся совершенно зелеными. Дело в том, что толстый слой атмосферы, сквозь который мы смотрим на далекий объект, слегка подкрашивает этот объект синим, и это трактуется зрением как факт большой удаленности (рис. 5.5.31). Причина появления синевы в том, что атмосферные частицы больше рассеивают фотоны длинных волн, нежели коротких, вызывая систематический спектральный сдвиг в сторону высоких частот. А это значит, что наше зрение стимулируется с относительным перевесом «коротких» фотонов (воспринимаемых синими).

Живописцы с незапамятных времен пользовались воздушной перспективой, имитируя удаленность изображенных ландшафтов. При этом отметим, что падение контраста и хроматики само по себе не является сакральным носителем информации о



**Рис. 5.5.31.** Воздушная перспектива как фактор восприятия глубины пространства. Понижение контраста и нарастание синевы заднего плана вызваны толстым слоем атмосферы, через который свет пробивается к наблюдателю. Наряду с прочими данный фактор усиливает ощущение глубины, которое, разумеется, сугубо иллюзорно.

глубине пространства, но в связке с прочими факторами восприятия глубины усиливает впечатление.

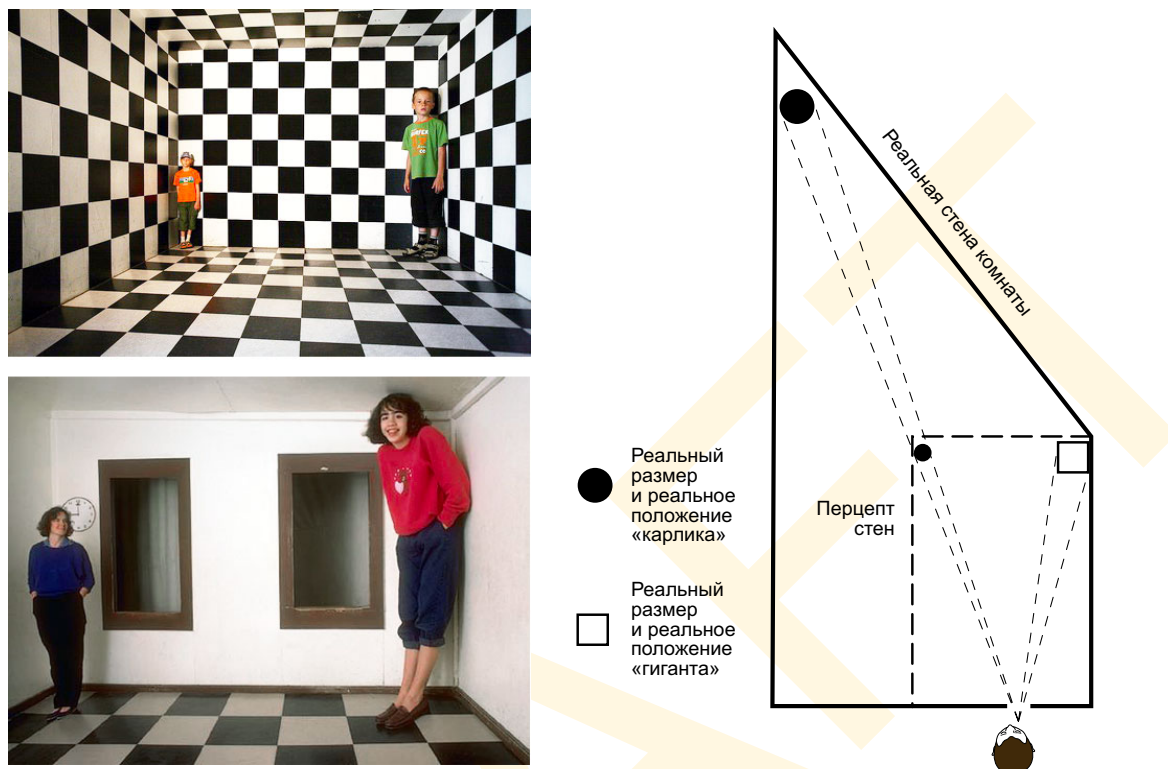
### 5.5.10 Интеграция источников информации

Итак, мы разобрали источники информации, из которых зрение черпает информацию, необходимую ему в оценке глубины пространства. Поскольку все эти источники работают на единую общую интерпретацию пространственного положения поверхностей, источники эти следует каким-то манером объединить в общее логически обоснованное представление. Каким же образом зрительная система выполняет такую интеграцию?

В естественных условиях наблюдения интеграция факторов глубины пространства не представляет трудности для зрительной системы, поскольку факторы эти плотно коррелируют в единое, согласованное и высокоточное представление расстояний и ориентаций поверхностей относительно наблюдателя. В лабораторных условиях, однако, разными факторами можно управлять независимо, что часто является причиной конфликта между ними. Что происходит в таких случаях и что это может дать в плане общего принципа интеграции различных видов информации? Рассмотрим три возможных варианта:

- один из конфликтующих источников *доминирует* над другим;
- между двумя конфликтующими источниками имеется *компромисс*;
- два источника *взаимодействуют* в стремлении достичь оптимального решения.





**Рис. 5.5.32.** Конфликт факторов глубины в комнате Эймса. Перспектива и сходство размеров конфликтуют, когда предъявлены зрению в искривленном помещении, воспринимаемом прямоугольным с определенной точки. Перспективный фактор берет верх (доминирует), в результате чего персонажи воспринимаются как сильно отличные по росту.

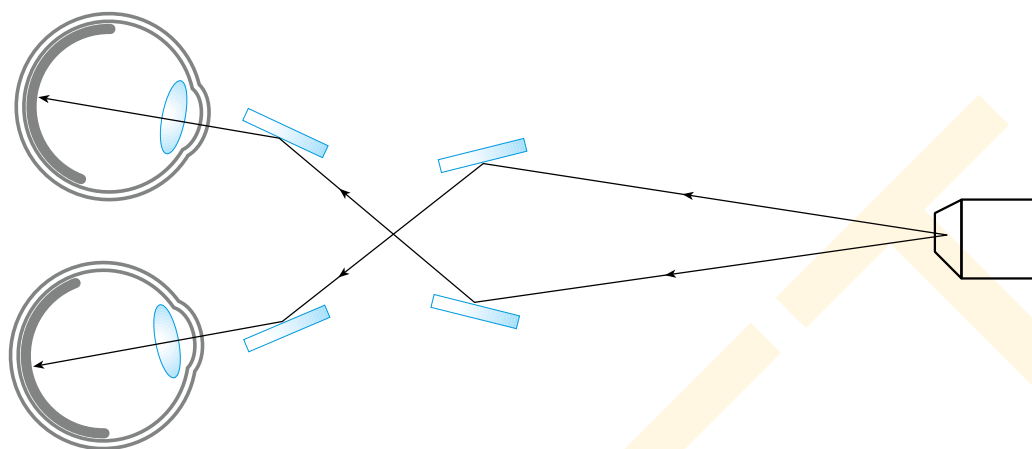
#### 5.5.10.1 Доминанта

Вероятно, простейшей является ситуация, когда один источник информации жестко доминирует над другим, конкурирующим с ним источником, в результате чего последний полностью игнорируется зрением. Такая форма интеграции предполагает некую иерархию факторов глубины, в которой те факторы, что стоят выше, доминируют над нижестоящими (хотя не совсем понятно, что может произойти в этой иерархии, если несколько источников «набросятся» на кого-то одного).

Широко известным примером явной доминанты является *комната Эймса*,<sup>1</sup> усугубляющая перспективную информацию об объектах сходных размеров. Эймсова комната (рис. 5.5.32) — это весьма необычное помещение, но которое выглядит самым обычным образом с определенной точки. Несмотря на то, что помещение на самом деле трапециевидно, оно с этой точки смотрится прямоугольным. Когда объекты, примерно равные по размеру расставлены вдоль задней стены этой комнаты, испытуемые однозначно докладывают о двух иллюзиях: во-первых, им кажется, что персонажи находятся на одном удалении от наблюдателя; во-вторых, что они очень сильно разнятся по

<sup>1</sup> Уникальный опыт, разработанный в 1946 году выдающимся американским исследователем в сфере физиологической оптики, директором Дартмутского института глаза Адельбертом Эймсом младшим. — *Прим. пер.*





**Рис. 5.5.33.** Псевдоскоп. Световые проекции в левый и правый глаз инвертированы, чем инвертирована стереоинформация. Однако сцена воспринимается нормально, поскольку обычные пикториальные факторы глубины не искажены и берут верх над факторами стереоскопическими.

размеру. Перспективная информация о глубине пространства в двухмерном ретинальном изображении эймсовой комнаты (напомним — лишь при наблюдении с определенной точки) оказывается такова, что помещение воспринимается правильной четырехугольной формы, с равными расстояниями между углами. Если принять сие за достоверный факт, то люди должны в этом случае неестественно разниться по росту согласно отличию в их проекциях на сетчатку — что, собственно, и наблюдаем. Фактор примерного сходства габаритов у человеческих особей *подавлен* перспективным фактором, абсолютно *доминирующим* в данном случае.

Несколько неожиданный пример доминанты в восприятии глубины пространства показывает нам, что даже если источник информации высоко значим и, казалось бы, необорим (как, скажем, бинокулярная диспаратность), его все равно можно подавить. Простейший оптический прибор — *псевдоскоп* (рис. 5.5.33) выполняет инверсию бинокулярной диспаратности, меняя местами оптические потоки, то есть, образуя на сетчатке левого глаза изображение, предназначенное для правого, а на сетчатке правого — для левого. Поскольку аппарат обращает бинокулярную диспаратность всех элементов изображения, то те объекты, что ближе к наблюдателю должны, по идее, восприниматься как удаленные и наоборот; вогнутые предметы должны выглядеть выпуклыми и т.д. Однако ежели взглянуть на обычную сцену через псевдоскоп, то ее глубина, вопреки прогнозу по диспаратности, воспримется вполне адекватно, то есть *не* обратится. Причина в том, что диспаратная информация может оказаться подавленной монокулярной пикториальной информацией, такой, в частности, что идет от кроющих контуров, текстурных градиентов, перспективы и исходной осведомленности наблюдателя о выпуклости объектов — то есть, той информацией, что не была искажена псевдоскопом. Таким образом, мы говорим, что по многим участкам сцены имеет место конфликт между диспаратностью и прочими факторами глубины пространства и что пикториальные источники всегда и полностью берут верх.

### 5.5.10.2 Компромисс

Второй вариант исхода конфликта факторов визуальной глубины таков, что зрительная система интегрирует информацию из разных источников, отыскивая компромиссные решения, не согласуемые ни с одним из факторов в отдельности, но расположенные где-то между. Наиболее очевидной тактикой является независимая оценка глубины отдельно по каждому источнику и последующая интеграция результатов по некоей математической схеме.

Никола Бруно и Джеймс Каттинг (Корнельский университет, 1988) провели опыт, в котором стимул нес в себе четыре фактора глубины пространства, управляемых независимо: относительный размер, положение относительно горизонта, окклюзию и параллакс движения. Исследователи просили испытуемых прошкалировать воспринятую глубину от 0 (удаление равно нулю) до 99 (максимальное удаление) по экранному показу, в котором различные комбинации факторов конфликтовали с другими комбинациями. В итоге, ученые предъявили убедительные доказательства того, что в своей оценке глубины пространства наблюдатели интегрируют множество источников, а не просто полагаются на какой-то один (чего следовало бы ожидать, целиком приняв концепцию иерархической доминанты). Они обнаружили также, что процессу интеграции информации свойственна аддитивность: испытуемые фактически суммировали независимые эффекты от четырех факторов, и чем больше было источников информации — тем большей представлялась им глубина пространства.

Доминик Массаро (Калифорнийский университет, 1988) перепроверил эти же данные и пришел к выводу, что лучшие результаты дает т.н. мультипликативная модель интеграции, то есть модель, построенная на умножении.

Ряд других исследователей (например, Doshier, Sperling & Wurst, 1986) также доложили о мультипликативном характере функций интеграции. Однако же возможны и другие математические подходы, и какие именно использует зрительная система человека — еще предстоит выяснить.

Отметим, наконец, что идею компромисса можно назвать «*легкой фузией*» факторов глубины (Clark & Yuille, 1990; Landy, Maloney, Johnston & Young, 1995), поскольку не предполагается взаимодействия между отдельными источниками информации.<sup>1</sup> Легкая фузия вполне совместима со строго модулярной позицией, которую мы коротко рассмотрели в разделе 2.3.3 и чуть ранее в текущей главе (раздел 5.1.2). Основная идея такова, что множество различных версий о глубине можно проработать независимо и параллельно. Каждая даст свой вариант, который можно нанести на некую карту — *карту глубин*. Карты интегрируются затем путем усреднения, сложения, перемножения и прочих математических способов комбинации по каждому участку. Результат обычно и являет собой некий компромисс.

### 5.5.10.3 Взаимодействие

Данный выше разбор вариантов интеграции факторов восприятия глубины пространства может показаться самодостаточным, однако за кадром остались несколько вопросов, суть которых становится понятной лишь при дальнейшем исследовании. Во-первых отметим, что только абсолютные источники информации о глубине про-

<sup>1</sup> Фузию множественных источников информации нельзя путать со стереоскопической фузией левого и правого ретинальных изображений: первое и второе — суть совершенно разные вещи.

странства способны дать т.н. *метрические карты глубины*, т.е. такие, которые можно количественно объединить для получения представительных результатов. Даже жестко количественные источники информации об относительной глубине (т.е. бинокулярная диспаратность и параллакс движения) для извлечения абсолютной информации о глубине пространства требуют дополнительных данных. При этом совершенно непонятно, как чисто качественные источники (окклюзия и динамика текстур) вообще могут приниматься во внимание. Более того, весьма вероятно, что различные виды информации о глубине, поступают от различных «модулей» отнюдь не отдельно, но *взаимодействуют* (как минимум, до определенной степени), давая единое согласованное представление расстояния до видимых объектов.

Рассмотрим, к примеру, синергичное взаимодействие бинокулярной диспаратности и конвергенции. Вспомним, что сама по себе бинокулярная диспаратность не характеризует абсолютные расстояния от наблюдателя до объектов, но лишь *соотношение* этих расстояний. Конвергенция *характеризует* абсолютное расстояние, но лишь до одного выбранного объекта, на котором сосредоточен взгляд. Однако же вместе они детектируют абсолютное расстояние до каждого из объектов в поле зрения. Если конвергенция дает информацию об абсолютном расстоянии до одного выбранного объекта, сие измерение можно затем использовать при шкалировании диспаратных отношений относительной глубины, дабы получить величину абсолютных расстояний до всех объектов поля зрения. Таким образом, полную карту глубин можно рассчитать, если оба источника информации взаимодействуют.

Сказанное справедливо в отношении взаимодействий между диспаратностью и многими другими источниками информации о глубине, в частности такими, как сходство размеров и аккомодация: зная любое одно абсолютное расстояние, можно рассчитать абсолютные расстояния до любых объектов в сцене.

Факт взаимодействия между различными источниками самоочевиден, то есть, совершенно невозможно представить, что его нет. Теоретически, однако, взаимодействия эти сложны и беспорядочны в сравнении с модулярной простотой легкой фузии. Майкл Лэнди (Беркли, 1995) сформулировал версию легкой фузии, которую окрестил «*модифицированной*», что позволяет четко выделить и классифицировать виды взаимодействий между источниками. Один из наиболее значимых вариантов взаимодействия — это взаимодействие между бинокулярной диспаратностью и любым другим источником абсолютной информации о глубине. Взаимодействие, при котором один из источников дает информацию, достаточную для того, чтобы поднять другой на уровень абсолютной глубины по всему полю зрения. Лэнди назвал сие метрическое представление *картой глубин*, а процессы подъема информации от ее источника до уровня карты — *продвижкой*. Таким образом, конвергенцию, аккомодацию или сходство размеров можно посредством шкалирования использовать для продвижки диспаратной информации до карты глубин. Сию «продвинутую» карту можно затем, как и предполагалось, объединить путем количественной интеграции в общий концепт легкой фузии.

Вопрос о том, как различные источники информации о глубине объединяются в единое согласованное представление трехмерного взаиморасположения видимых поверхностей (такое как марров 2.5D-скетч) — многогранен и сложен. Перцепционное доказательство в целом скудно и бессистемно, количественные исследования еще только-только разворачиваются, а четкого физиологического доказательства идеи и вовсе нет. Каждый исследователь допускает, что определенная форма интеграции имеет место, но никто не знает какова она. Мы вынуждены признать, что сие наименее

изученный аспект зрительного восприятия взаиморасположения поверхностей в естественных сценах.

### 5.6 ОНТОГЕНЕЗ

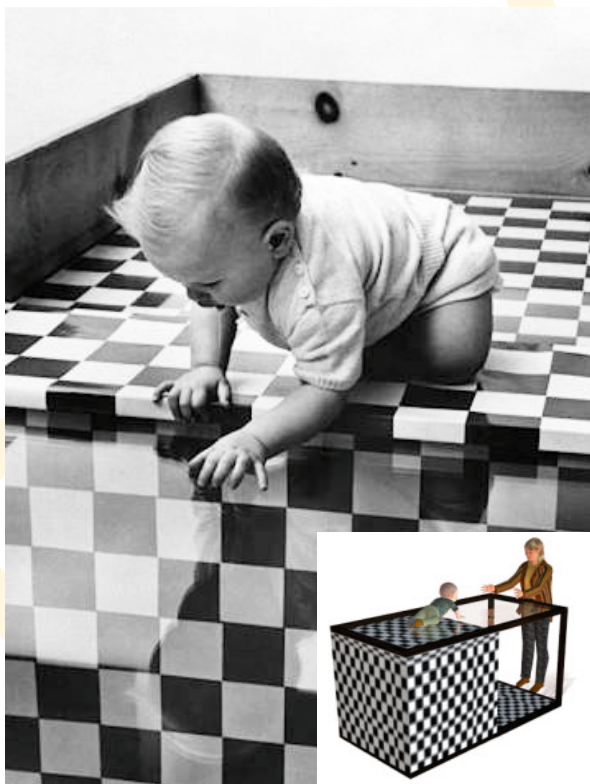
Взрослыми мы без видимых усилий воспринимаем расстояния до поверхностей и их ориентацию в пространстве. Однако же родились мы с этой способностью, или она развилась самостоятельно после рождения? Или, может быть со временем мы обучились ей и натренировались до автоматизма? Ответ на эти интригующие вопросы дают наблюдения за младенцами.

Самые первые и наиболее эффектные исследования, проведенные при помощи аппарата, названного *зрительным уступом*, демонстрируют то, что дети обладают способностью к восприятию глубины пространства в самом раннем возрасте (Е. Gibson & Walk, 1960). Зрительный уступ представляет собой некий стеклянный «аквариум» (рис. 5.6.1). Под стеклом на любой из сторон и на разной глубине можно расположить текстурированную поверхность, а подсветку организовать так, что стекло станет совершенно невидимым. В классическом варианте эксперимента американцы Р. Уок и Е. Гибсон положили текстурированную шашечками ткань прямо под стекло на «мелкой» стороне и идентичную текстуру на полу — на 1 м ниже стеклянной столешницы на «глубокой» стороне.

Затем экспериментаторы просили маму ребенка положить малыша на центральную вставку и попытаться заставить его ползти либо по «мелкой» стороне, либо по «глубокой». Если ребенок способен воспринимать расстояние до двух текстурированных поверхностей, то он наверняка поползет к матери через «мелкую» сторону, но никак не через провал уступа.

Таким манером были протестированы 36 младенцев в возрасте от 6 до 14 месяцев: 27 из них поползли к матери и все 27 прошли по «мелкой» стороне хотя бы один раз, но только трое попытались преодолеть метровый «провал». Результат однозначно показал, что когда ребенок уже ползает, то есть в период с 6 до 12 месяцев, он способен к восприятию глубины пространства.

Чтобы ответить на вопрос, могут ли младенцы еще неспособные ползать, воспринимать глубину пространства, Дж. Кампос, А. Лангер и А. Кровитц (1970) записывали частоту сердцебиений у двухмесячных младенцев на зрительном уступе. Они обнаружили, что малыши в возрасте от двух до пяти ме-



**Рис. 5.6.1.** Младенец на зрительном уступе. Когда мать зовет, большинство детей смело проползают по мелкому участку уступа, но далеко не все — по глубокому.



сяцев, положенные на «мелкую» сторону прибора, выдают небольшое, но статистически вполне устойчивое изменение в частоте сердцебиения. Когда их клали на «глубокую» сторону, частота их сердцебиения заметно... падала. Тот факт, что частота пульса снижалась, а не повышалась, указывает на то, что детей не пугала глубокая сторона (что привело бы к учащению сердцебиения), напротив — они с любопытством рассматривали предметы на глубине. Следовательно, в двухмесячном возрасте малыши уже способны воспринимать глубину, но еще не боятся высоты.

При исследованиях на зрительном уступе имеется одна сложность, состоящая в том, что никак не определить, каким источником информации о глубине пользуется малыш: вполне вероятно, что способность извлекать окулярную информацию, информацию стереоскопическую, о движении, а также пикториальную развивается у ребенка в один и тот же возрастной период (внося примерно равновеликий вклад в суммарный отклик). Но при этом столь же вероятно, что и в разные, селективно определяя реакцию младенца в зависимости от возраста и опыта. Чтобы решить эти вопросы потребовались новые методики. Рассмотрим некоторые из них.

### 5.6.1 Окулярная информация

Окулярная информация о глубине пространства идет из мышечного сигнала и у новорожденных присутствует лишь в рудиментарной форме. Следовательно, аккомодации и конвергенции требуется определенное время для развития. М. Банкс (1980) изучал аккомодацию у детей первого, второго и третьего месяцев жизни, используя т.н. *ретиноскоп* — прибор, позволявший определять, насколько четко глаз фокусирует свет. Исследователь обнаружил, что одномесячные дети едва способны аккомодировать, тогда как аккомодация трехмесячных младенцев практически такая же, как у взрослого. Однако последнее вовсе не значит, что трехмесячные дети при восприятии глубины пространства пользуются этим источником. Мы можем себе лишь предположить. Собственно аккомодация нужна ведь еще и для эффективного стереопсиса, и восприятия деталей. И первое, и второе тоже играют немаловажную роль в зрительном восприятии пространства.

Также предпринимались попытки выяснить, представлена ли у младенцев способность к бинокулярной конвергенции. Ричард Эслин (1977) напрямую измерял характер глазных движений у детей, предъявляя им некий образец-мишень, который постепенно отдалялся от малыша. Если ребенок способен к бинокулярной конвергенции, его глаза должны постепенно расходиться, когда мишень удаляется, и сходиться — когда приближается. Эслин обнаружил, что хотя способность к адекватной окулярной конвергенции появляется весьма рано, устойчиво следить за перемещением предметов малыши могут лишь с трех месяцев.

И вновь исследования не дают ответа на вопрос — извлекает ли малыш полезную информацию о глубине пространства из этой способности или только лишь *может* извлечь?

С. Хофстен (1977) показал, что пятимесячные дети все-таки воспринимают глубину пространства на основе окулярной конвергенции, надев на них специальные очки: ученый обнаружил, что ребенок начинает путаться в расстояниях согласно мере искажения, даваемой очками.

Однако непрямые методы исследования свидетельствуют о том, что способность к визуальной эффективной конвергенции дана ребенку с рождения. Исследования новорожденных показали, что тенденция к константности формы и размера (см. разделы 7.11



и 7.2.1) фиксируются исследователями уже на первой неделе жизни ребенка (Stater, Mattlock & Brow, 1990; Stater & Morison, 1985). Следовательно, новорожденные каким-то образом извлекают визуальную информацию о глубине пространства, вероятнее всего — бинокулярную (хотя в этом и нет полной уверенности). Подробнее мы поговорим об этом в гл. 7, когда речь пойдет об онтогенезе зрительной константности.

### 5.6.2 Стереоскопическая информация

Итак, коль скоро дети способны к бинокулярной конвергенции, есть все предпосылки к развитию стереоскопического зрения. Ряд процедур указывает на то, что восприятие глубины пространства, основанное на бинокулярной диспаратности, развивается примерно в 3.5 месяца, и лишь спустя некоторое время — способность к конвергенции.

Один из подходов к исследованию данной проблемы позволяет ответить на вопрос, как рано дети обретают способность воспринимать стохастическую стереограмму. Вспомним (раздел 5.3.2), что пара стохастических стереограмм не содержит в себе монокулярной информации о какой-либо форме глубины, но при бинокулярном просмотре наблюдатель с нормальным стереоскопическим зрением воспринимает некую фигуру перед фоном из точек или позади него. В экспериментах, проведенных Р. Фоксом, Р. Эслином, С. Ши и С. Думэйсом (1980), детям в возрасте от 2.5 до 6 месяцев с помощью специальных очков предъявляли стохастические стереограммы. Фигура, «отстоящая» от фона, представляла собой прямоугольник, который смещался либо вправо, либо влево, а экспериментатор оценивал наличие перцепта глубины, отмечая протяженность слежения ребенком за движением прямоугольника в том или ином направлении. Исследователь обнаружил, что трехмесячный младенец не обращает внимания на движение фигуры, а вот в три с половиной месяца смещение прямоугольника уже интересует малыша. Отметим, что результаты эксперимента указывают на то, что дети фактически воспринимают глубину пространства, а не только лишь способны к такому восприятию — потому как будь иначе, они не следили бы за движением фигуры.

Однако способность к восприятию стохастических стереограмм — это довольно грубый критерий способности к стереоскопическому восприятию глубины. При этом прочие методы указывают на то, что стереовосприятие проявляет себя раньше. Американцы Р. Хельд, Е. Бирч и Дж. Гвизда (1980) на основе т.н. парадигмы предпочтений отыскивали *стереопороги* — минимальную диспаратность, на которую малыши способны отреагировать. Использовался стереопоказ (рис. 5.6.2) — одна его сторона представляла собой плоскую решетку (без диспаратности), другая — аналогичную решетку, в которой прутья (наблюдателем с нормальным стереозрением) воспринимались как расположенные на разной глубине (поскольку предъявлены диспаратно). Предварительные результаты указали на то, что дети предпочитают смотреть на трехмерное изображение, нежели двухмерное (Fantz, 1965). Экспериментаторы пришли к выводу, что дети предпочитают стереоскопический показ только тогда, когда способны к восприятию глубины пространства.

Используя специальные стереоочки для предъявления слегка отличных стимулов правому и левому глазам, Хельд обнаружил устойчивое предпочтение стереорешеток в возрасте 3.5 месяца при диспаратности в  $1^\circ$  зрительного угла. Предпочтение на диспаратности меньше 1 минуты ( $1/60^\circ$ ) обнаружилось в пятимесячном возрасте. Сие заметно меньше, чем способность взрослого человека к детекции диспаратности<sup>1</sup>, но

<sup>1</sup> Взрослый способен воспринять диспаратность в секунды зрительного угла ( $1/60$  минуты).



**Рис. 5.6.2.** Объемный показ при исследовании онтогенеза зрения. Детям предъявляли две решетки: одну — плоскую, другую, воспринимаемую взрослыми объемной за счет бинокулярной диспаратности. Малыши моложе 3.5 месяцев не отдавали предпочтения ни первому, ни второму вариантам показа. Тогда как дети старшего возраста предпочитали разглядывать объемный стимул.

По материалам Held, Birch & Gwiazda, 1980,

четко указывает на то, что дети в 3.5 месяца обладают сравнительно эффективным стереозрением.

### 5.6.3 Динамическая информация

Как уже отмечалось ранее, существует несколько динамических источников информации о глубине пространства: параллакс движения, градиент движения, люминг, кинетический эффект глубины, а также динамика текстур. Востребованы ли какие-то из перечисленных факторов при рождении или развиваются лишь со временем?

Возможно главным кандидатом на врожденно воспринимаемый фактор глубины пространства является *люминг* (градиент движения приближающейся поверхности), поскольку применим к объектам, движущимся в сторону неподвижного наблюдателя. Новорожденные не могут двигаться сами, создавая тем самым экстенсивный самообразующийся оптический поток, но биологически значимые объекты — мать и собственные конечности — могут двигаться к ребенку и от него, сразу после рождения. Дети от второго месяца жизни уже откликаются на люминг соответствующими оборонительными реакциями (Bower, Broughton & Moore, 1970, 1971) — они откидывают голову назад и подносят ручки к лицу, когда контуры предмета увеличиваются в их поле зрения. Исследователи интерпретировали эти результаты как указующие на то, что информация о глубине пространства, извлекаемая из люминга объектов, воспринимается в самом раннем возрасте, возможно даже, что сразу после рождения (Bower, 1974).

Авторская интерпретация результатов — открытый вопрос, однако: возможно, дети откидывают голову и поднимают ручки просто потому, что они очерчивают верхний контур предмета, когда контур этот, к примеру, пошел вверх. В пользу такой версии говорит то, что впоследствии обнаружилось — младенцы на единичный контур,

движущийся верх (равно и на растущий в размере объект), выдали один и тот же отклик (Yonas, 1981). Интересно однако, что месячные младенцы чаще моргали глазками именно в ответ на люминг (Yonas, 1981).

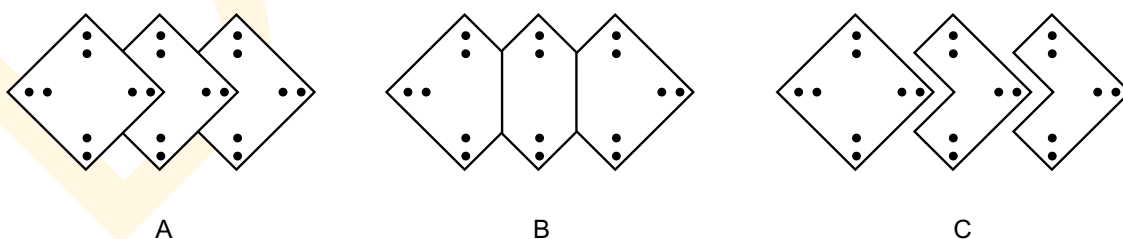
Таким образом, оптическая экспансия вполне может врожденно стать фактором визуальной глубины пространства. Однако даже если сие и не так, то, по крайней мере воспринимаемым на самых ранних этапах внеутробной жизни.

Было обнаружено также, что и прочие динамические источники информации о глубине — градиенты движения, динамика текстур, а также кинетические эффекты глубины оказывают влияние на психофизическую реакцию ребенка в раннем младенческом возрасте (см. Kellman, 1995, 1996). К примеру, П. Кельман (1984) установил, что после трех месяцев дети сенситивны к форме предметов, что проявляет себя кинетическими эффектами глубины: малышей вначале приучали к тому, что объект вращается в пространстве по двум различным осям; затем вновь предъявляли вращающийся предмет — либо тот же самый, либо другой. Дети реагировали только на другой предмет, что свидетельствует об их способности распознавать форму объектов. Дальнейшие изменения в условиях эксперимента показали, что реакция основана на кинетическом (а не на статическом) восприятии формы, поскольку отсутствует, когда последовательно предъявляются неподвижные предметы.

#### 5.6.4 Пикториальная информация

Существует множество пикториальных источников информации о глубине пространства, но не все еще изучены в контексте онтогенеза организма человека. Имеющиеся сегодня данные, свидетельствуют о том, что такие факторы глубины как размер, окклюзия (интерпозиция), светотень (шейдинг), линейная перспектива и текстурные градиенты начинают оказывать влияние и поставлять визуально-значимую информацию о глубине где-то между пятью и семью месяцами внеутробной жизни. Сие важно в том плане, что перечисленные факторы становятся источниками полезной информации много позже, чем источники окулярные, стереоскопия и движение.

Здесь однако мы не станем давать полный обзор исследований, выполненных по данному вопросу (лучше см. Yonas, Arterberry & Granrud, 1987), но рассмотрим одно из них в качестве примера. Ц. Гранруд и А. Йонас (1984) попытались выявить возраст, в котором малыши становятся чувствительными к окклюзионной информации. Детям предъявлялись стимулы, показанные на рис. 5.6.3, но только одному глазу, в результате чего бинокулярно диспаратная информация никак не могла выдать их двух-



**Рис. 5.6.3.** Стимулы, предназначение для изучения онтогенеза значимости окклюзии в восприятии глубины пространства. Дети старше 7 месяцев больше всего разглядывали А-показ, несущий фактор окклюзии и существенно меньше обращали внимания на варианты В и С.

По материалам Granrud & Yonas, 1984.

мерности. По стимулам рис. 5.6.3 А для выявления кроющих и кроемых краев использовались Т-вертексы, дающие вполне добротную информацию об относительной глубине различных объектов. Стимулы рис. 5.6.3 В и 5.6.3 С использовались для поверхностного аналогичного контроля условий, в которых информация о глубине не представлена и, соответственно, не читается даже взрослыми. Результаты показали, что семимесячные дети намного чаще обращали внимание на часть А, тогда как пятимесячные одинаково реагировали на все три стимула. Исследователи пришли к выводу, что окклюзия как фактор глубины пространства становится доступной в период между пятью и семью месяцами внеутробной жизни.

\* \* \*

Итак, результаты весьма искусных экспериментов по изучению онтогенеза зрения, в частности, описанных выше, показали, что большинство источников информации о глубине пространства не используются новорожденными, но постепенно обретают значимость в течение первого года жизни. Однако вполне возможно, что более тонкие методы заставят нас пересмотреть имеющиеся данные, сместив их к более раннему возрасту.

Реальные сцены заполнены по большей части крупными *объектами* — людьми, деревьями, домами, автомобилями и пр., но никак не краями, штрихами или пятнами (см. гл. 4) и даже не собственно поверхностями (гл. 5). То есть, до сего момента в наших рассуждениях о пространственном зрении отсутствовал некий очень важный компонент — мы не говорили еще о вопросах т.н. *перцептивной организации* зрения на низких пространственных частотах. Проще говоря, мы никак не касались того, как биты и пакеты зрительной информации структурируются в *перцептивные единицы* — перцепты самостоятельных объектов и взаимосвязи между этими перцептами.

Перцептивная организация неразрывно связана со зрительным опытом и определяет его. Когда мы смотрим на реальную трехмерную сцену, мы однозначно воспринимаем всю картину взаимосвязей между объектами этой сцены. Нам это дается сходу и легко, и почти невозможно вообразить, что здесь могут возникнуть какие-то сложности. Но давайте на секунду представим, что перцептивной организации в крупные поверхности и объекты не существует: вероятнее всего, в этом случае в нашем сознании вообще не возникало бы цельного зрительного образа. Неорганизованный зрительный образ, пожалуй, правильнее всего уподобить вихрю разноцветного конфетти — какофонии сигналов, пришедших от миллионов никак не связанных между собой ретинальных рецепторов.

Исследователи — приверженцы эмпирической традиции — говорят о том, что таковы зрительные ощущения новорожденных. Эмпиристы полагают, что вся зрительная структура — это плод обучения, опыта, и что мир глазами новорожденного малыша совершенно деструктурирован. Известный философ-психолог Вильям Джеймс даже придумал остроумный термин — «клубящаяся муть». Но на самом деле мы попросту не знаем, каковы зрительные впечатления детей раннего возраста, поскольку малыши не в силах рассказать о них — по крайней мере напрямую. Впрочем, сравнительно недавно были разработаны весьма эффективные методы, позволяющие по крайней мере предположить, каков мир глазами младенцев. И чем больше мы узнаём — тем яснее становится то, что у малышей наличествуют базовые элементы перцептивной организации. Оставшееся развивается где-то между шестью и восьмью месяцами внеутробной жизни и, действительно, благодаря обучению в процессе общения с окружающим миром. То есть, перцептивная организация строится как на врожденных, так и на приобретенных механизмах (о чем пойдет речь в конце главы).

Некоторые данные о значимости перцептивной организации мы можем получить из слов человека, чье зрение оказалось дезорганизовано (или, как минимум, плохо организовано) вследствие черепно-мозговой травмы. Впрочем, жалобы, им озвученные, были далеки от «вихря конфетти», т.е. пострадал лишь самый высокий уровень восприятия. Но даже при таком раскладе у нас есть возможность заглянуть внутрь ситуации, в которой перцептивная организация нарушена. Показания больного подкреплены и тем ценным фактом, что еще до того, как был поставлен неврологический диагноз, состоялась консультация медицинского психолога, которому удалось добыть кое-какие интроспективные данные о зрительных ощущениях пациента:



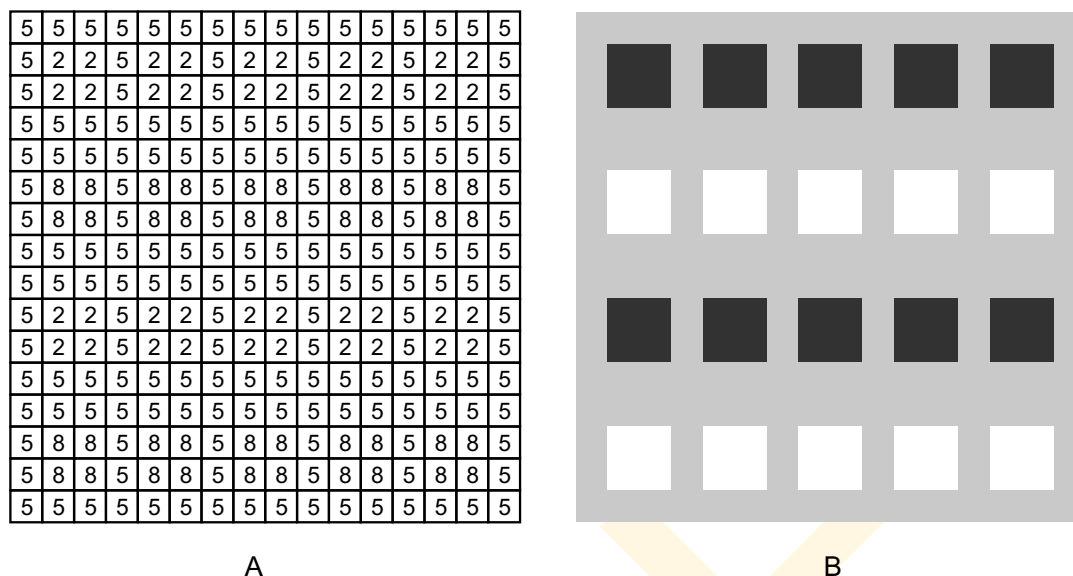
«Если я смотрю на сложный объект, скажем, человека или группу людей, я порой вижу, что не хватает отдельных фрагментов, в смысле, они не связаны между собой, эти фрагменты, хотя... если человек движется в каком-то одном направлении и все части его тоже в этом направлении, то... ну, тогда он будет как один. Впрочем, у меня со многими вещами возникала эта же проблема: каждая вещь была в отдельности, но только сама по себе — они никак не собирались вместе... А в некоторых случаях то, что я видел, было совсем абсурдным, например, я хорошо помню случай, когда видел предмет, состоящий частично из капота машины, частично — из дерева и частично — из человека в бейсболке.» (Marcel, 1983 а).

Концепцию перцептивной организации предложили гештальтисты в начале XX века, причем в роли градобойного орудия атомистических посылов структуралистов. Вспомним, что последние рассматривали зрительный перцепт как простую совокупность сенсорных «атомов», как множество точечных цветовых ощущений. Такая позиция сугубо локалистична: каждый «атом» определен конкретной ретиальной локалью и независим от других «атомов» (по крайней мере до тех пор, пока путем ассоциативного обучения они не объединятся в крупные пространственные комплексы). В противоположность сему гештальтисты полагают, что зрительное восприятие рождается из двух источников: из общей структуры зрительной стимуляции и из глобальных взаимодействий внутри зрительной системы. Термином «перцептивная организация» они объединяют и первое, и второе. Множество исследованных ими зрительных феноменов подтвердили эту точку зрения.

Первым заявил проблему перцептивной организации Макс Вертхаймер (напомним, один из отцов — основателей гештальт-психологии) в своей знаменитой статье «*Untersuchungen zur Lehre von der Gestalt*»<sup>1</sup> (Wertheimer, 1923). Он озаботился вопросом: За счет чего мы воспринимаем мир цельным, состоящим из *объектов*, а не хаотически разбросанных цветных точек? Понять вертхаймерову мысль, наверное, проще всего, положив выходной сигнал сетчатки простым, но при этом высокоструктурированным изображением. На рис. 6.11 А этот сигнал дан в виде матрицы чисел, в которой каждое значение представляет величину нейронного отклика единичного фоторецептора. Чтобы в сей числовой форме разглядеть хоть какую-то структурность и организацию, нужны неимоверные усилия: стимул очень тяжел для восприятия, поскольку зрительная информация не представлена им в явном и пригодном для нормального зрения виде. Проще говоря, любой зрительный стимул несет в себе бесчисленное число возможных вариантов организации изображения, но воспримем мы лишь какой-то один. Каким будет наше зрительное ощущение и почему оно будет именно таким, а не другим? — это главный вопрос текущей главы.

Структура числового изображения становится понятной и очевидной, когда информация перекодируется в колебания яркости (рис. 6.1.1 В). Это картина нескольких черных и белых квадратов на сером фоне, организованных в четыре горизонтальных ряда. Почему же тогда так трудно воспринимать то же, но в виде цифр? Разумеется, потому что зрительная система умеет извлекать информацию о краях, участках, предметах, группах и сценах лишь из яркостной и хроматической структуры ретиальных изображений, но никак не из цифр. Яркостное ахроматическое изображение на рис. 6.1.1 В почти полностью задействует этот механизм, тогда как в числовом представлении — едва ли. В обоих изображениях представлена одна и та же информация, но числовое представление зрительная система не в состоянии воспринимать непо-

<sup>1</sup> Исследования по изучению гештальтов — *Прим. перев.*



**Рис. 6.1.1** Вопрос перцептивной организации. Когда на сетчатке появляется оптическое изображение, зрительная система начинает искать структуру в паттерне выходных сигналов сетчатки (матрица цифр в части А — чем выше число, тем ярче стимул). Когда наблюдателю предъявляют соответствующее яркостное изображение (В), он моментально и без усилий организует его в две пары рядов из черных и белых квадратов на сером фоне.

средственно. Когда исследователи пытаются истолковать механизмы зрительного восприятия, они оказываются примерно в той же ситуации, что и мы, когда хотим отыскать структуру в числовом изображении: невозможно представить себе ни одного варианта естественной зрительной организации, который логика наша могла бы объяснить столь же легко и непринужденно, как это делает наша зрительная система. Слишком высок уровень структуры.

Почему у зрительного перцепта структура именно такая, а не иная? Самый очевидный ответ таков, что зрительный образ должен отражать устройство внешнего мира. То есть, коль скоро мир этот фактически состоит из пространственно организованных поверхностей и объектов, а не окрашенных точек, то восприятие наше организовано сходным образом. Сие простейшая реалистическая позиция, но в ней несомненно есть здоровое зерно: понятно, что по эволюционным причинам перцептивная организация как раз и должна отражать структуру внешнего мира, или по крайней мере ту его часть, что значима в плане выживания особи. Вообразим, к примеру, насколько мало было бы толку от зрения, ежели б оно характеризовало мир бессистемно (как, скажем, в описанной выше ситуации с поражением мозга). И хотя реалистическая позиция может помочь истолковать перцептивную организацию в эволюционном ключе (то есть дать ответ, *почему* структура перцептов такова, а не иная), позиция эта не может объяснить *механизм* этой организации, то есть то, как она формируется во время зрительного акта. Цель текущей главы — пролить свет на эти механизмы и факторы стимуляции, запускающие их.

Наконец, два слова о том, что принято называть «перцептивной обманкой». Вечная «головная боль» позиции наивного реализма в том, что зрительная система не имеет прямого доступа к объектам внешнего мира, но только к проекциям этих объектов на сетчатку. То есть, особь не может сделать вывод о структуре внешнего мира на основе

одной лишь сенсорной информации. Гештальтисты окрестили наивно-реалистический подход к проблеме перцептивной организации *перцептивной обманкой*, поскольку возникает она из ошибочного (и часто неявного) предположения, будто структура перцепта напрямую выводится из светового потока, простимулировавшего сетчатку (Kohler, 1947). Оптический массив фактически несет в себе бесконечное число вариантов организации, из которых зрительная система выбирает единственно верный.

Заблуждение, лежащее в основе перцептивной обманки, происходит обычно из того, что отправной точкой зрительного акта полагают не проксимальный стимул, но дистальный. И попасть в эту ловушку очень легко, поскольку, безусловно, дистальный стимул — это важнейший компонент в причинно-следственной цепочке событий, ведущих к возникновению зрительного ощущения. Более того — он четко согласуется с интерпретацией, к которой стремится зрение. Однако принять дистальный стимул в качестве стартовой позиции зрения — это грубо недооценить сложность зрительного акта, это допустить мысль, будто необходимая информация поступает из сцены «даром».

Итак, какой бы очевидной и логичной не выглядела идея примата дистальной стимуляции, мы видим теперь, какие трудности возникают, например, при создании компьютеров, способных «видеть», и что масштаб проблемы перцептивной организации толком не осознавался до появления в 1923 г. эпохальной статьи Вертхаймера. Действительно, несмотря на существенный прогресс прошлых лет, исследователи зрения до сих пор открывают новые грани этой важнейшей и всеобъемлющей проблемы.

## 6.1. ПЕРЦЕПТИВНАЯ ГРУППИРОВКА

Исследовать вопрос перцептивной организации Вертхаймер начал с изучения факторов стимуляции, определяющих т.н. *перцептивную группировку*, то есть то, как различные элементы сложного стимула воспринимаются «сочетанно», т.е. объединяются в единый, цельный перцепт. Ученый начал с того, что изготовил несколько простейших геометрических фигурок и стал менять отношения между ними, стараясь отыскать причину, по которой фигуры объединяются в перцептивные группы. Логически множество элементов можно было разбить по группам, исходя из принципа деления на взаимоисключающие подмножества. Количество таковых росло очень быстро: по 10 элементам имелось 42 варианта группировки, тогда как по 100 — уже 190569292. Число логически возможных группировок оказалось бы еще бóльшим, если бы одно подмножество вошло в состав другого, то есть, если бы появились общие члены подмножеств. Психологически, однако, в данный момент времени воспринимается лишь один единственный вариант. Как сие возможно? И какие параметры *проксимальной* стимуляции влияют на перцептивную группировку?

### 6.1.1 Классические принципы группировки

Возможно, сказанное ниже, станет некоторым и весьма удачным исключением из методологии гештальтистов, которых частенько критикуют за регулярные отступления от строгой экспериментальной процедуры и склонность к психофизическим, поведенческим исследованиям (например, Pomerantz & Kubovoy, 1986). На самом деле гештальтисты вполне могут обойтись и без экспериментов, поскольку изучаемые ими феномены столь выразительны и явны, что в собственно экспериментах нет нужды. Например, если сотни или даже тысячи людей в ответ на стимуляцию дружно сообща-

ют об одних и тех же феноменологических эффектах, одних и тех же зрительных впечатлениях, то к чему здесь лабораторные формальности?

Как часто отмечал Ирвин Рок, демонстрации гештальтистов (в частности, показанную на рис. 6.1.2) вполне можно рассматривать как бессрочный эксперимент с неопределенно большим числом испытуемых. Но в ситуациях, когда факты не столь очевидны, гештальт-психологи ставят тщательно продуманные опыты и записывают объективные данные, например такие, как количество наблюдателей, сообщивших об одном варианте зрительного ощущения, и количество сообщивших о противоположном (например, Goldmeier, 1936/1972). Таким образом, феноменологические методы гештальт-психологии имеют полное право значиться в списке современных реактивных исследований.

### 6.1.1.1 Сближенность

Свои опыты Вертхаймер начал с демонстрации линии, состоящей из равноотстоящих друг от друга точек (рис. 6.1.2 А), то есть, не сгруппированных в какую-то самостоятельную перцептивную единицу (ежели не считать самой линии). После чего ученый уменьшил просвет в каждой паре точек, то есть так, что он оказался ощутимо меньше, чем между парами (рис. 6.1.2 В). Сей фактор, который исследователь окрестил *сближенностью*, стал первым среди знаменитых вертхаймеровых «законов» группировки.<sup>1</sup> Доказательство того, что сближенность есть фактор потенциальной группировки, было сугубо феноменологическим: ученый попросту предъявлял массив точек с рис. 6.1.2 В своим студентам и просил их непосредственно указать, какие из точек по их мнению «сочетаются». Поскольку никто и не пытался всерьез обсуждать вертхаймеров опыт, и никому в голову не приходило ставить под сомнение его результаты, фактор сближенности закрепился в науке сугубо феноменологически, то есть без какого-либо формального эксперимента.

### 6.1.1.2 Подобие и общий ток

Убедившись в эффекте сближенности, Вертхаймер приступил к поиску других факторов группировки (см. рис. 6.1.2). На рис. 6.1.2 С, D и E, к примеру, показана работа принципа *подобия*:

⇒ При прочих равных все элементы, подобные друг другу<sup>2</sup> — группируются.

Мы говорим при этом, что подобие — это наиболее общий принцип перцептивной группировки, поскольку охватывает множество различных свойств стимула.

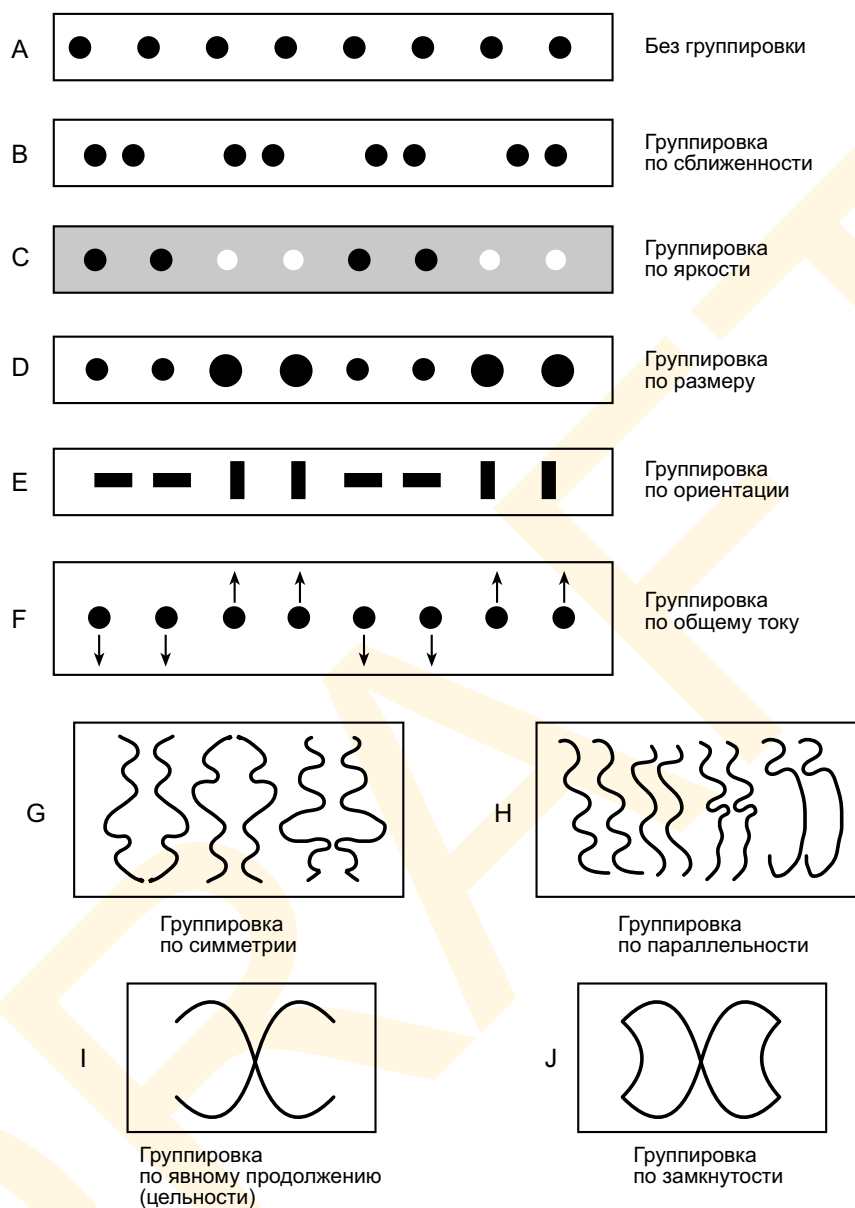
Следующий вертхаймеров фактор звучит как *общий ток*:

⇒ При прочих равных все элементы общего направления движения — группируются.

Эффект невозможно продемонстрировать в статике, поэтому рис. 6.1.2 F отражает его сугубо символически. Отметим, что общий ток фактически можно рассматри-

<sup>1</sup> Сегодня мы говорим о них все же именно как о принципах или факторах группировки, поскольку, как будет показано далее, они не отвечают требованиям, предъявляемым к научным законам.

<sup>2</sup> В данном примере подобие по яркости (С), размеру (D) и ориентации (E).

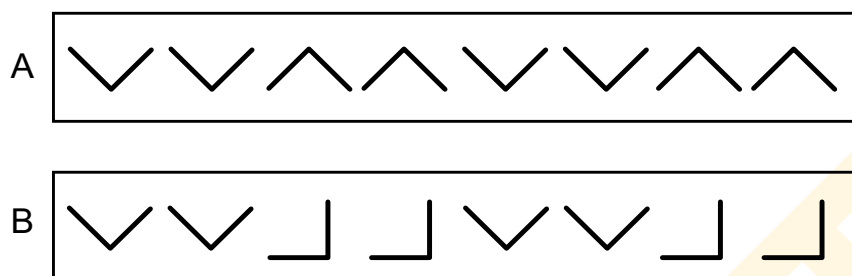


**Рис. 6.1.2** Классические факторы группировки. Гештальт-психология выделяет множество различных факторов, влияющих на то, какие из визуальных элементов станут перцептивно собираться в крупные группы (подробнее см. в тексте).

вать как частный случай группировки по подобию, в котором подобными оказываются векторы движения.

При этом не все возможные варианты подобия одинаково эффективны, а некоторые даже не способны привести к группировке. Рассмотрим, к примеру, V-образные фигурки на рис. 6.1.3 А: соседние пары отличаются друг от друга на  $180^\circ$  по ориентации, но тенденция к группировке остается весьма слабой. На рис. 6.1.3 В даны те же фигурки, пары которых разнятся поворотом всего лишь на  $45^\circ$ , но теперь стремление к визуальной группировке очевидно. То есть, мы начинаем понимать, что зрительная система





**Рис. 6.1.3** Степени группировки. Не все факторы одинаково эффективны в данном смысле: вверху (А) элементы повернуты друг относительно друга на  $180^\circ$ , но группировка весьма вялая, тогда как внизу (В) отличие всего лишь на  $45^\circ$ , но при этом тенденция к визуальной группировке очевидна.

чувствительнее к одним видам отличий, нежели к другим. Даже если присмотревшись к парам рис. 6.1.3 А, мы и выявим отличия, сие потребует усилий в сравнении с моментальным и легким восприятием групп на рис. 6.1.3 В. Ниже мы разберем, почему так происходит.

#### 6.1.1.3 Симметрия, параллельность, цельность и замкнутость

Гештальт-психология описывает еще ряд факторов перцептивной группировки вытянутых элементов. К примеру, *симметрия* (рис. 6.1.2 Г) и *параллельность* (рис. 6.1.2 Н) — это факторы, определяющие перцептивную группировку отдельных линий и кривых.

Рис. 6.1.2 I иллюстрирует весьма важный фактор — т.н. *явное продолжение (цельность)* линий и краев:

⇒ Все элементы, что при прочих равных, цельны — склонны к визуальной группировке.

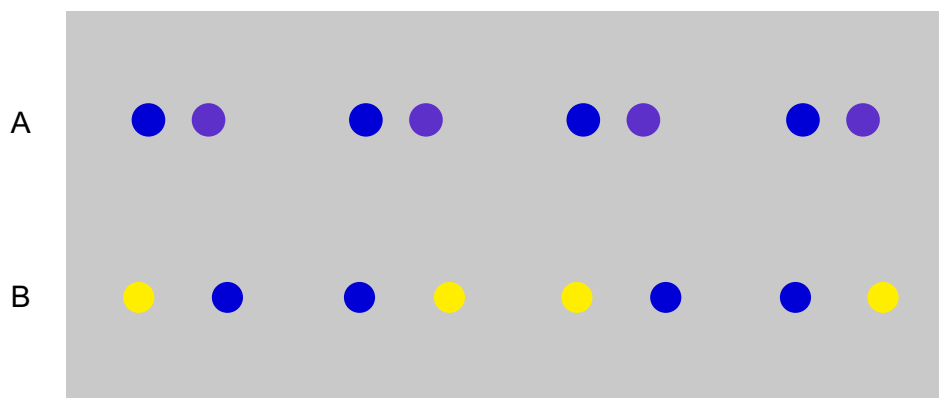
Эффект весьма выражен — мы явно видим две пересекающихся кривые, но никак не два угла, соприкасающихся вершинами.

На рис. 6.1.2 J дан еще один фактор — *замкнутость*.

⇒ Все элементы, при прочих равных образующие закрытую фигуру — группируются.

Обратим внимание на то, что замкнутость может доминировать над цельностью: те же линии, что смотрелись как линии, теперь образуют два угла, сходящихся вершинами. Согласно вертхаймеровому рассуждению так происходит от того, что нецельные сегменты теперь являются составными частями одной фигуры.

Стоит сказать, что цельность и замкнутость (рис. 6.1.2 I и 6.1.2 J соответственно) — это наглядная иллюстрация несовершенства нашего сегодняшнего представления о принципах визуальной группировки. Поясним: гештальт-психологи вынуждены рассуждать в рамках принципа *ceteris paribus* — что значит «при прочих равных», то есть прогнозировать визуальный результат только при условии, что никакие прочие факторы группировки не повлияют на этот результат. Мы наблюдаем, к примеру, что когда элементы не образуют закрытой фигуры, цельность управляет группиров-



**Рис. 6.1.4** Дуализм группировки по окраске и сближенности. Сильные отличия по сближенности и малые отличия по окраске ведут к группировке по сближенности (А), тогда как сильные отличия по окраске и малые по сближенности — ведут к группировке по окраске (В).

кой. Но как только появляется замкнутость, так сразу она берет верх над остальными факторами, в том числе и над цельностью.

Несовершенство критерия *ceteris paribus* состоит в том, что он исключает интеграцию потенциально конфликтующих факторов. В результате отсутствует алгоритм прогнозирования силы суммарного эффекта этих факторов. Проблема касается всех ранее описанных принципов группировки. Если, к примеру, сближенность подталкивает группировку к одному варианту, а окраска — к другому, результат окажется контекстно зависим. На рис. 6.1.4 А показана ситуация, в которой сближенности вполне достаточно для устойчивого доминирования над окраской, тогда как на рис. 6.1.4 В дан вариант, когда сходство по окраске берет верх. Очевидно, зрительная система интегрирует множество факторов, но мы до сих пор не понимаем, как именно. Ниже мы познакомимся с некоторыми современными методиками, позволяющими интегрировать ряд аспектов группировки по подобию (на примере процессов текстурной сегрегации), но по-прежнему не сможем ничего сказать об иных принципах, таких как общий ток, цельность и замкнутость.

## 6.1.2 Дополнительные факторы группировки

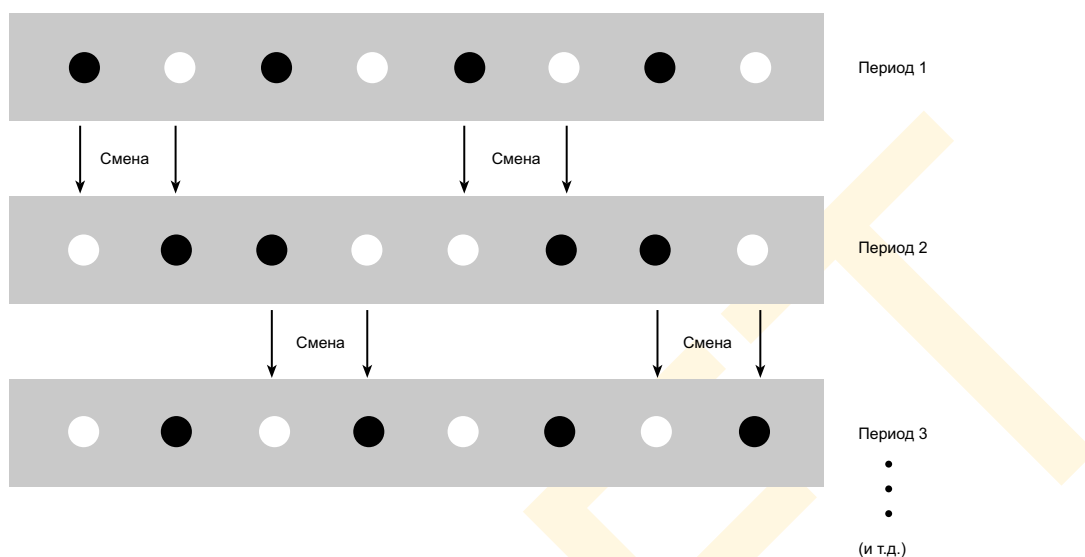
Современных работ, посвященных исследованию вопроса зрительной группировки, в частности ее принципов, на удивление мало. Тем не менее, недавно были предложены еще три фактора: *синхронность* (Palmer & Levitin, 1999), *общий участок* (Palmer, 1992) и *связность* (Palmer & Rock, 1994 a).

### 6.1.2.1 Синхронность

Принцип синхронности основан на том, что:

⇒ Зрительные события, при прочих равных возникающие в одно и то же время — группируются.

И, хотя в сфере слухового восприятия этот фактор и ранее признавался весьма значимым (к примеру, Bregman, 1978), в науке о зрении на него до последнего времени не



**Рис. 6.1.5** Группировка по синхронности. Элементы, при прочих равных синхронно меняющие свои свойства — визуально группируются.

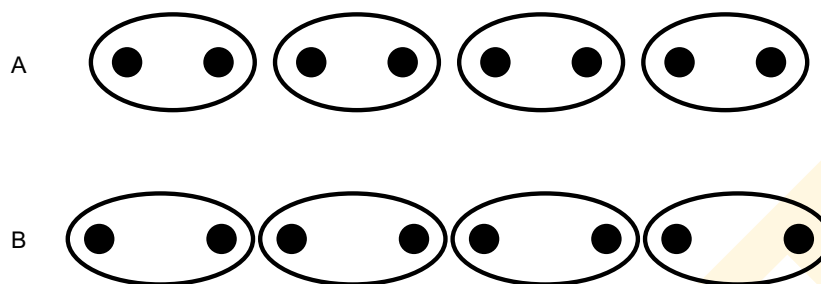
обращали особого внимания (Palmer & Levitin, 1999). Итак, на рис. 6.1.5 дан пример, в котором каждый элемент в ряду равноотстоящих друг от друга черных и белых точек с определенной частотой меняет окраску: черные делаются белыми и наоборот. Стрелками показано то, что половина кружков переходит от черных к белым или от белых к черным в один момент времени, а половина — в другой. Когда частота перемигивания составляет 25 изменений в секунду (или немного ниже), зрение испытуемого по критерию синхронности четко делит точки на пары. На больших частотах группы распадаются, а наблюдатель видит лишь хаотически мигающие точки. На очень низких частотах происходит молниеносная группировка в момент смены окраски, но распадается в промежутках. Отметим, что синхронность относится к классическому принципу общего тока (то есть, мы говорим, что сие — динамический фактор), но, как видно из приведенного примера, «ток» элементов нельзя назвать «общим» — в один и тот же момент времени некоторые точки ярче, другие — темнее.

### 6.1.2.2 Общий участок

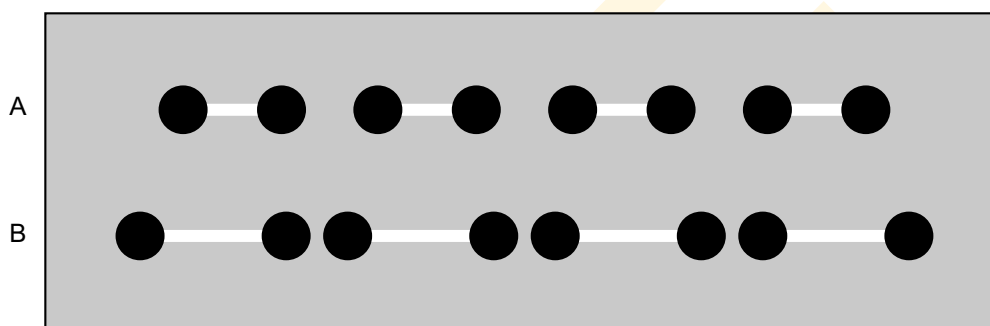
Еще один, недавно позиционированный принцип зрительной группировки — *общий участок* (Palmer, 1992):

⇒ Все элементы, что при прочих равных расположены внутри одного и того же замкнутого участка — группируются.

На рис. 6.1.6 А дан пример аналогичный классической вертхаймеровой демонстрации (рис. 6.1.2 В — 6.1.2 Е): при прочих равных, равноотставленные точки жестко группируются в пары, когда оказываются внутри замкнутого окружающего их контура. По рис. 6.1.6 В четко видно, что общности участка полностью достаточно, чтобы взять верх над сближенностью, которая в данном примере несет противоположный вектор группировки.



**Рис. 6.1.6** Группировка по общему участку. При прочих равных элементы внутри одного пространственного участка — группируются (А), даже если они расположены дальше друг от друга, чем элементы разных участков (В).



**Рис. 6.1.7** Группировка по связности элементов. При прочих равных элементы, связанные друг с другом через другие элементы — группируются (А), даже если они расположены дальше друг от друга, чем элементы разных участков (В).

### 6.1.2.3 Связность элементов

Третий принцип — *связность элементов*:

⇒ Все элементы, что при прочих равных соединены с другими элементами — группируются.

Пальмер и Рок (1994) несколькими примерами продемонстрировали, что связность элементов обладает определенным группирующим потенциалом. Пример, аналогичный классической вертхаймеровской демонстрации, дан на рис. 6.1.7: равноотставленные друг от друга точки однозначно группируются, когда соединены друг с другом дополнительными элементами (такими, в частности, как короткие горизонтальные отрезки). На рис. 6.1.7 В показано как связность берет верх даже над сближенностью.

Отметим, что Вертхаймер толковал связность элементов как частный случай максимальной сближенности. Однако Пальмер и Рок показали отличие связности от сближенности по ряду моментов:

- во-первых, имеет место важная качественная разница между фактической связностью и полной сближенностью;

- во-вторых исследователи показали, что «связность» в строго физическом смысле — это фактическая соединенность, но отнюдь не просто тесное соседство. Части объектов, что физически соединены друг с другом, перцептивно намного сильнее спаре-

ны, чем просто близкорасположенные предметы (как бы близко друг к другу они не располагались). Следовательно, зрительной системе есть прямой смысл быть особо чувствительной к связности как к фактору прогнозирования событий внешнего мира;

— в-третьих, существует и феноменологическое отличие между связностью и простым соседством — связанные элементы обычно воспринимаются единым целым (т.е. комплексным объектом, состоящим из разных частей), тогда как простая сближенность перцептивно дает рыхлый конгломерат самостоятельных, хотя и родственных объектов.

Отличие между простой сближенностью и фактической связностью указывает на то, что в некоторых случаях может возникнуть т.н. *агрегация элементов* — рыхлая их совокупность как результат процессов перцептивной группировки. Сближенность, подобие, общий участок и некоторые случаи общего тока часто дают агрегацию при которой элементы удерживают высокую степень перцептивной независимости и оказываются вне групповых взаимоотношений. Прочие принципы группировки могут дать *единичное образование* — перцепт отдельного, визуально связанного объекта, происходящего из нескольких исходных элементов. Связность элементов, их цельность и общий ток дают более когерентную организацию в единые объекты.

Читая о принципах группировки, можно подумать, что они — очередные книжные сущности, имеющая лишь отдаленное отношение к реальному восприятию. Вертхаймер показал однако, что сущности эти заполняют собой весь наш перцептивный опыт, поскольку отвечают за идентификацию объектов и участков воспринимаемого внешнего мира.

Несколько эффектных примеров того, как перцептивная организация дает сбой, предлагает нам природная маскировка (рис. 6.1.8). Ее цель в том, чтобы помешать процессам группировки в зрительной системе хищника, выделяющим жертву из ее окружения (удачно замаскировавшаяся особь группируется с окружением в основном за счет различных вариантов подобия). Если окраска животного и рисунок его тела сильно схожи с окружением по яркости, хроматике, ориентации, размеру и форме, особь перцептивно группируется с фоном и становится условно невидимой. Эффект может оставаться близким к идеальному до тех пор, пока животное остается неподвижным, но даже безупречный камуфляж не держит критерий общего тока: как только начинается движение, общий ток пятен и контуров вызывает сильную степень группировки, давая хищнику достаточно информации, чтобы моментально выделить жертву как самостоятельный объект из ее окружения.

### 6.1.3 Степень группировки

Итак, коль скоро гештальт-демонстрации группировки работают лишь в рамках критерия *ceteris paribus*, они не годятся для разработки количественных концепций, описывающих интеграцию множественных факторов, и нам нужна отдельная методика измерения *степени* группировки. Сравнительно недавно были предложены две таких методики — одна, основанная непосредственно на словесных сообщениях наблюдателей, другая — на решении не прямой, но объективно определенной задачи.

#### 6.1.3.1 Метод Кубового — Вагеманса

Американский психолог Майкл Кубовой и бельгиец Йохан Вагеманс (1995) измерили весовые доли различных вариантов группировки, демонстрируя испытуемым неопределенную точечную сетку (рис. 6.1.9 А). Исследователи оценили вероятность с



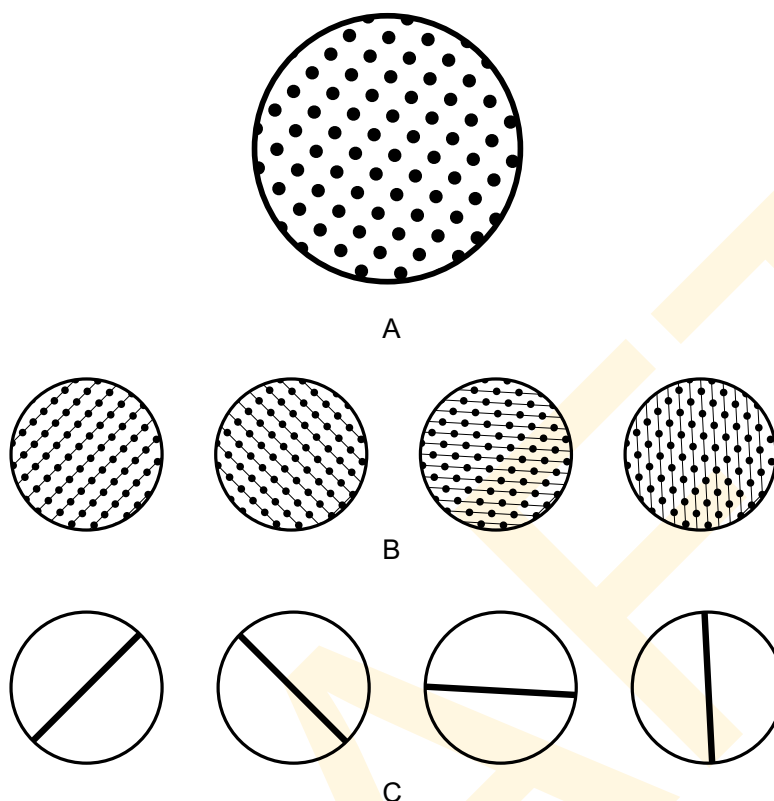


**Рис. 6.1.8** Пример естественной маскировки. Многие звери, птицы и насекомые обладают удивительной способностью визуально сливаться со своим естественным окружением, активно эксплуатируя гештальт-факторы группировки. Но стоит только животному двинуться с места, как весь камуфляж немедленно и полностью пропадает.

которой наблюдатели докладывали о том, что видят тот или иной вариант организации группировки. Сетка была неопределенна в том смысле, что ее точки могли группироваться в линии четырех ориентаций (рис. 6.1.9 В). Стимул предъявляли наблюдателю в течение 300 мс (миллисекунд), после чего просили показать «свой» вариант (рис. 6.1.9 С). После серии попыток вычислялась вероятность появления тех или иных перцептов. Как и ожидалось (согласно гештальт-принципу сближенности), результаты показали то, что наиболее вероятный вариант организации тот, что определен ближайшим расположением точек — прочие варианты теряли вероятность по мере роста расстояния между точками. Более того, данные четко укладывались в математическую модель, в которой перцептивное стремление точек друг к другу падало экспоненциально по мере увеличения расстояния.

#### 6.1.3.2 Метод Пальмера — Бек

Еще один, недавно предложенный Стивеном Пальмером и Дианой Бек (Принстонский университет), способ оценки степени группировки был назван *методикой детекции повтора*. Подобно алгоритму Кубового — Вагеманса, метод Пальмера — Бек предполагает объективно корректные ответы наблюдателя. Испытуемым предъявлялись стимулы, аналогичные показанным на рис. 6.1.10: каждый представлял собой ряд из перемежающихся квадратиков и кружочков за исключением одной пары, где элемент был просто раздвоен. Задача наблюдателя заключалась в том, чтобы определить, из чего состоит такая пара — из квадратиков или кружков. Ответ регистрировался возможно быстрым нажатием одной из двух кнопок — одной для квадратиков, другой — для кружков. Время отклика измерялось в трех возможных условиях: при *внутри-*



**Рис. 6.1.9** Неопределенность в группировке элементов точечной решетки (А): точки могут группироваться в линии разных ориентаций (В). Кубовой и Вагеманс просили испытуемых показать, какую ориентацию они увидели, выбрав соответствующий символ (С).

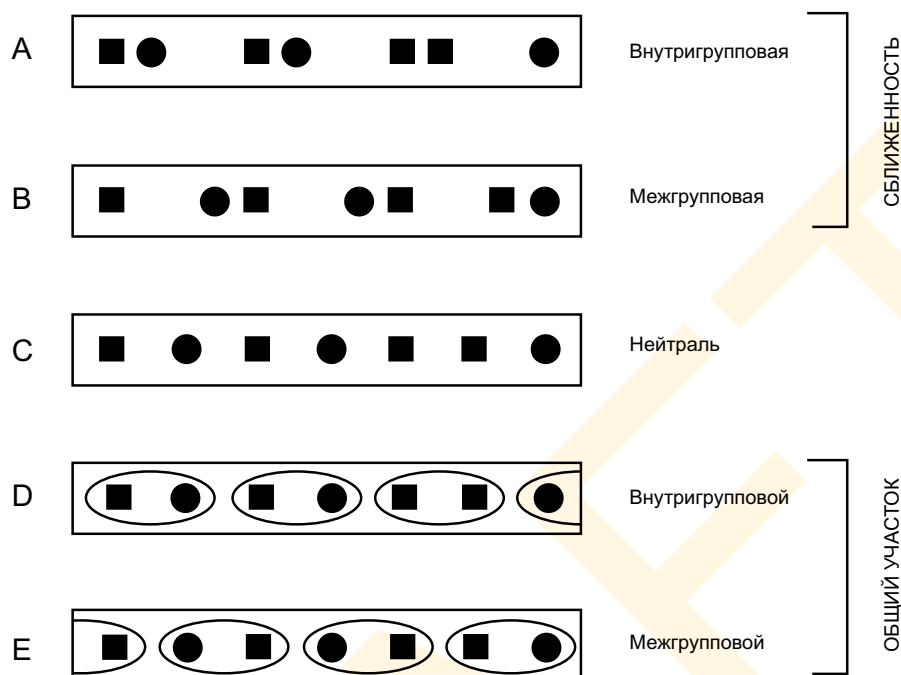
По материалам Kubovy & Wagemans, 1995.

*групповом* режиме — когда группирующий фактор (сближенность на рис. 6.1.10 А) сдвигал целевую пару к организации в той же группе; при *междугрупповом* режиме — когда тот же самый фактор сдвигал целевую пару к варианту организации как части двух разных групп (рис. 6.1.10 В); при *нейтральном* режиме, когда сдвига не возникало (рис. 6.1.10 С). Ожидаемый результат был таков, что целевую пару наблюдатель должен распознать быстрее, если она входит в состав одной группы.

Результаты дали устойчивую корреляцию группирующих факторов с временем реакции: отклики на сближенные стимулы оказывались более скорыми при внутригрупповом режиме показа (719 мс), нежели при междугрупповом (1144 мс) — сравним рис. 6.1.10 А и 6.1.10 В. При этом отклики при внутригрупповом режиме оказались столь же скоры, как и при нейтральном (730 мс), в основном потому что подобие по форме заставляло их группироваться в пару даже при отсутствии прочих факторов. Сходные результаты были получены при детекции пар, когда те формировались по сходству окраски, общему участку и связности элементов.

На рис. 6.1.10 С, 6.1.10 D и 6.1.10 Е даны нейтральный, внутригрупповой и междугрупповой режимы показа, в которых группирующим фактором являлся общий участок. Результаты эксперимента оказались сходными с первым вариантом опыта (по сближенности).

\* \* \*

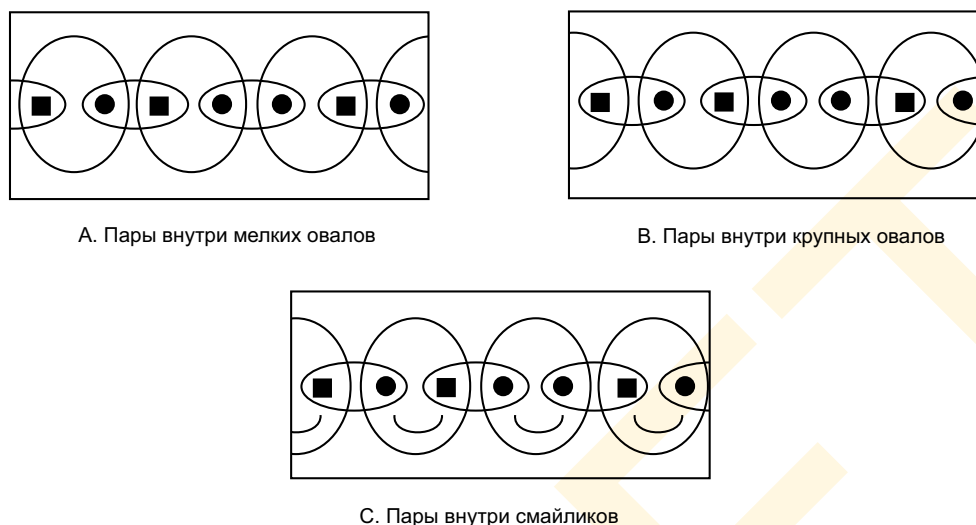


**Рис. 6.1.10** Примеры стимулов, использованных в эксперименте по детекции повтора: наблюдатели должны были распознать повторяющуюся пару кругов, либо пару квадратов. При внутригрупповой попытке (A и D) повторяющиеся элементы располагались внутри группы, заданной тем или иным группирующим фактором (сближенностью в A и общим участком в B). При межгрупповой попытке (B и E) элементы пары принадлежали разным группам. В нейтральном варианте группирующие факторы отсутствовали.

Существенным преимуществом количественных методов, в частности, описанных выше, является то, что они позволяют выполнить прецизионную оценку феноменологически неочевидных эффектов группировки. К примеру, Пальмер и Бек использовали метод детекции повтора для определения того, какие овалы — малые или большие — оказывают большее влияние на группировку при общем участке, если очевидно конфликтуют в рамках одного показа. Прежде (1992) Пальмер показал, что меньшие участки доминируют — рис. 6.1.11 A и 6.1.11 B. В дальнейшем, используя метод детекции повтора и путем некоторых манипуляций со стимулами, Пальмер и Бек смогли показать, что мелкие овалы оказывают большее группирующее влияние и что отличие в размере овалов важнее отличий в их ориентации. Удивительно, что доминанта малых овалов сохраняется даже тогда, когда к большим овалам добавлены «смайлики», превращающие эти овалы в «лица» (рис. 6.1.11 C). Сие указывает на то, что группировка в данной конкретной задаче не зависит от сближенности, а также от сходства с лицом человека (последнее, судя по всему, обретает значимость на более поздних этапах зрительной обработки).

#### 6.1.4 Этапность группировки

Вопрос о том, на каком этапе зрительной обработки происходит перцептивная группировка — наиважнейший в науке о зрении: Группировка — это ранний процесс, исполняемый на уровне структуры ретинального изображения, или он начинается



**Рис. 6.1.11** Влияние размера в общем участке. Результаты эксперимента по детекции повтора показали, что повторяющиеся пары внутри мелких участков (А) распознаются быстрее, чем те же пары внутри крупных (В). Сие справедливо даже тогда, когда крупные участки выполнены явно и однозначно в виде рожиц.

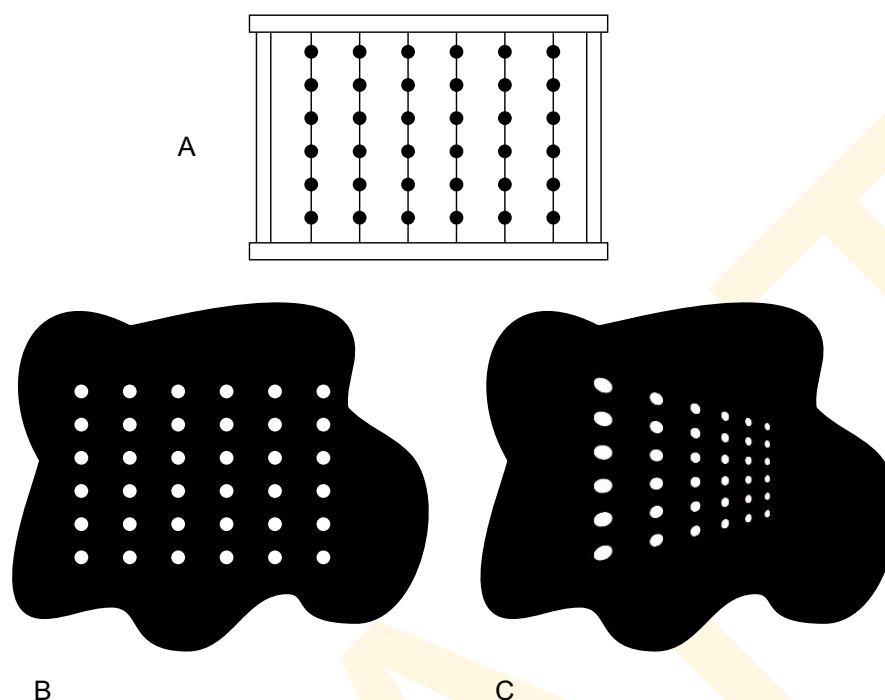
позже, когда информация о глубине пространства уже извлечена и перцептивная константность достигнута?<sup>1</sup>

Вертхаймер рассуждал о группировке, полагая, что она возникает на самом низком уровне, возможно соответствующем тому, который мы назвали «оптическим этапом» зрительной обработки. Эмпирического доказательства своей позиции ученый не дал, но общая точка зрения звучит так, что организация возникает именно на ранних этапах, дабы обеспечить процессы высокого уровня входным перцептивным материалом. Однако, такой подход крайне редко обсуждался всерьез, по крайней мере до недавнего времени. Впрочем, имеются кое-какие аргументы в пользу вертхаймеровской позиции, но обычно гештальтисты и не вспоминают о них, поскольку используют демонстрации, в которых глубина и константность иррелевантны: двухмерная стимуляция, предъявленная с фронтальной плоскости при равномерном освещении. В столь упрощенных условиях невозможно определить, на каком уровне работают главные факторы группировки — на уровне двухмерного ретинального изображения или на уровне трехмерной перцептивной структуры.

#### 6.1.4.1 Опыт Рока — Бросголя

Первый, добротный поставленный эксперимент по эксплицитному разделению прогнозов перцептивной организации на обоих этапах, опирался на фактор сближенности (Rock & Brosiole, 1964). Исследовался вопрос о том, где собственно зашита расстояния,

<sup>1</sup> Вспомним, что константность восприятия относится к способности зрения воспринимать неизменными свойства дистальной стимуляции, т.е. свойства объектов вне зависимости от вариаций стимуляции проксимальной (т.е. ретинального изображения), вызванных отличиями в условиях наблюдения. Детально этот вопрос мы рассмотрим в гл. 7.



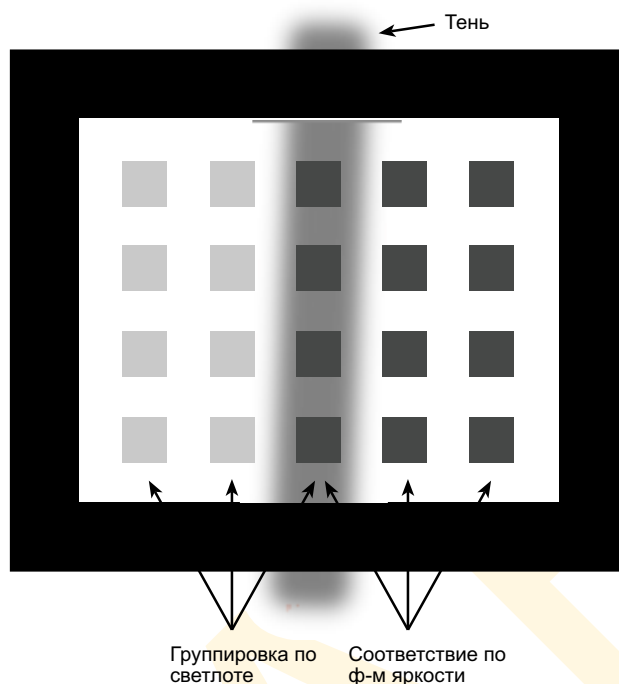
**Рис. 6.1.12** Ретинальное и перцептивное расстояния при группировке по сближенности. Когда шарики (А) рассматриваются во фронтальной плоскости, то, благодаря сближенности, они перцептивно организуются в колонки (В). Когда плоскость шариков повернута в глубину, то при бинокулярном рассматривании колончатая структура перцепта сохраняется, даже несмотря на то, что на сетчатке в правой части матрицы шарики оказываются существенно ближе друг к другу. Результат свидетельствует о том, что стереобработка влияет на группировку по сближенности.

управляющие группировкой по сближенности — в двухмерном ретинальном изображении или в трехмерном перцепте реальной сцены? Ирвин Рок и Леонард Бросголь использовали двухмерную прямоугольную матрицу светящихся шариков, предъявляемых наблюдателю (в затемненном помещении) либо фронтально (т.е. перпендикулярно зрительной оси) — либо наклоненной в глубину так, что *горизонтальный размер* уменьшался в зависимости от угла наклона (рис. 6.1.12). Когда матрица шариков рассматривалась во фронтальной плоскости (рис. 6.1.12 В), наблюдатели всегда докладывали о том, что шарики группируются в вертикальные колонки, но не в горизонтальные ряды.

Главный вопрос звучал так: Как пойдет группировка, когда ту же сетку шариков предъявят наблюдателю наклоненной в глубину, но так, что шарики сблизятся горизонтально (рис. 6.1.12 С)?<sup>1</sup> Неудивительно, что когда наблюдатели рассматривали наклоненную плоскость одним глазом, то есть, когда бинокулярный компонент исключался, они сообщали, что шарики организуются в ряды, т.е. согласно прогнозу по ретинальной сближенности. Сие, вероятно, возникало от того, что испытуемые ошибочно воспринимали плоскость шариков как расположенную фронтально (даже при ее наклоне более чем на 40° в глубину). Однако когда наблюдателям удавалось получить ве-

<sup>1</sup> Отметим, что в реальности они по-прежнему ближе друг к другу вертикально.





**Рис. 6.1.13** Группировка и константность светлоты. Когда центральная колонка квадратов воспринимается как затененная, то перцептивная группировка идет по признаку общей светлоты, то есть с квадратами слева, невзирая при этом на то, что фотометрическая яркость у квадратов центральной колонки та же, что и у образцов справа. Сей результат свидетельствует о том, что группировка по подобию ахроматической окраски определяема константностью светлоты.

ридикальный перцепт за счет бинокулярного просмотра, они докладывали о том, что видят наклоненную плоскость с кружками, организующимся в вертикальные *колонки*, т.е. так же, как и при фронтальном показе.

Итак, результат опыта Рока — Бросголя подтвердил гипотезу о том, что группировка возникает после стереоскопического восприятия глубины пространства.

#### 6.1.4.2 Опыт Рока — Ниджхэвена — Пальмера — Тудора

Рок, Ниджхэвен, Пальмер и Тудор (1992) в той же логике обратились к восприятию яркости и попытались ответить на вопрос: Возникнет ли группировка по светлоте ахроматических стимулов (основана на перцептивной оценке *яркости* элементов ретинального изображения) по достижении ее константности (см. раздел 3.3)? В первом эксперименте для отделения яркости от светлоты использовались наведенные тени: наблюдателям предъявлялись стимулы аналогичные показанным на рис. 6.1.13, после чего просили ответить, с кем группируются элементы центральной колонки: с элементами слева или справа от нее?

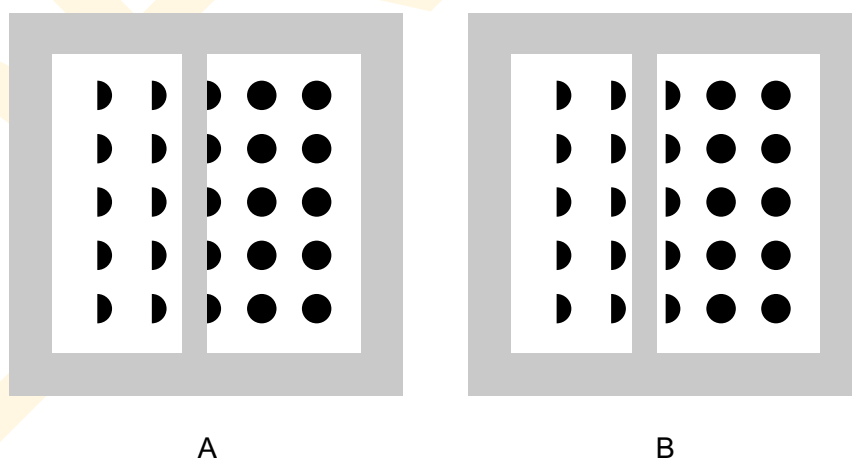
Стимул (рис. 6.1.13) предъявляли в строго заданных условиях: центральные квадратики были уравнены по отражательной способности с квадратами слева (попросту были выполнены из одного куска серой бумаги), но рассматривались при этом под тенью от картонной полоски. Затенение подобрали так, что их фотометрическая яр-

кость (т.е. количество света, отраженного от центральных квадратиков и летящего в глаз наблюдателю) оказалась идентична фотометрической яркости квадратиков справа. Далее рассудили так: если группировка базируется на относительно ранней обработке структуры ретинального изображения, центральные квадратик должны группироваться по яркости — то есть с квадратиками справа; ежели она основана на относительно поздних процессах, т.е. проходящих после восприятия затенения — группировка пойдет по светлоте, т.е. с квадратиками слева. Результат показал, что ситуация развивается по второму сценарию: группировка по подобию управляема *светлотой* квадратиков, но не яркостью их проекций на сетчатку. (Прочие эксперименты показали однако, что результат может зависеть от простого соотношения между яркостью квадратиков и их фонов).

#### 6.1.4.3 Опыт Пальмера — Неффа — Бек

Зрительная достройка также влияет на характер перцептивной группировки (Palmer, Neff, & Beck, 1996). Как было показано в гл. 1, достройка относится к тому факту, что когда наблюдатели смотрят на объект, частично закрытый другим, то имеет место устойчивая тенденция воспринимать его форму как завершенную позади кроющего объекта. Многие исследователи полагают, что сей процесс сравнительно поздний, проходящий после восприятия объектов и после того, как определены взаимоотношения по глубине. Если группировка по сходству формы определяется завершенной формой, это явится еще одним доказательством того, что сие относительно поздний процесс.

Стивен Пальмер, Джон Нефф и Диана Бек (1996) пытались разобраться в том, чем определяется группировка по форме — очертаниями неполных ретинальных элементов или перцептом формы полных? Используя тот же принцип организации показа, что и у Рока (1992), они построили стимул, в котором полукруги центральной колонки воспринимались как полные круги, частично закрытые вертикальной полоской (рис. 6.1.14 А). Если исходить из группировки на ранних этапах зрительного акта, то



**Рис. 6.1.14** Группировка и зрительная завершенность. Полукруги центральной колонки плотнее группируются с кругами справа, когда смотрятся полузакрытыми (А), нежели когда полностью открыты (В).

получается, что центральные элементы должны группироваться с элементами левой половины, поскольку их ретинальная форма — полукруг. Если же рассуждать в логике поздних этапов, то группировка должна пойти с правой половиной, поскольку кружки в этом случае будут восприниматься как цельные, но закрытые полоской.

Как видим, центральные фигуры группируются с правыми цельными кружочками, указывая тем самым на то, что группировка основана на подобию полной формы, нежели формы ретинальной. Вероятность того, что сие определено наличием закрывающей полоски, которая делит элементы на два участка согласно общему участку исключается контрольным сдвигом этой полоски (рис. 6.1.14 В): открывается реальный контур центральных элементов, позволяющий их полукруглой форме восприниматься однозначно. И хотя фактор общего участка оказывает определенное влияние на исход эксперимента, большинство испытуемых относят контрольные элементы к левой группе полукругов. Сие есть еще одно доказательство того, что группировка относительно сложный и достаточно поздний процесс в зрительном акте.

\* \* \*

Итак, результаты опытов убедительно свидетельствуют о том, что группировку нельзя относить к ранним, предконстантным зрительным процессам. Однако существует некая вероятность того, что группировка — это процесс, обладающий определенной временной протяженностью и включающий в себя компоненты как ранней, так и поздней обработки. Предварительная группировка может определиться на раннем, предконстантном этапе обработки ретинального изображения, но над ней может взять верх предметная информация (извлекаемая из глубины пространства, условий наблюдения, перекрытий и т.п.). Доказательством сего является ситуация, когда группировка на ранних этапах влияет на отработку зрительной константности, т.е. когда группировка предшествует последней.

Еще один аргумент в пользу того, что группировку можно разбить на раннюю и позднюю стадии, дают эксперименты, в которых ранние и поздние факторы объединяются, давая некие промежуточные результаты (к примеру Beck, 1975; Oslen & Attneave, 1970).

### 6.1.5 Прошлый опыт

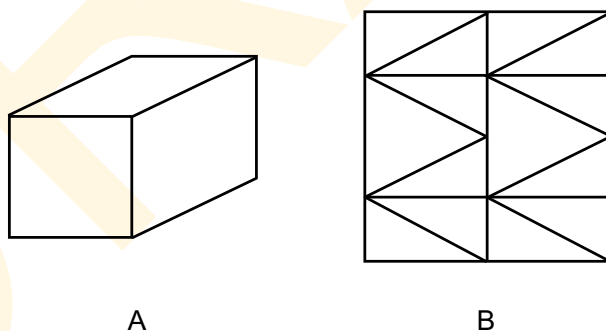
В завершение разговора о группировке следует отметить, что Вертхаймер (1923/1950) говорил еще об одном факторе перцептивной группировки, о котором сегодня почти не вспоминают — это т.н. *фактор прошлого опыта*. Суть в том, что элементы, ассоциированные зрением в прошлом, склонны к группировке в настоящем — рис. 6.1.15: в первый момент мы воспринимаем эту картинку как случайный набор черных пятен на белом фоне, но, спустя какое-то время, начинаем различать фигуру далматинца, обнюхивающего паранет улицы. Картинка радикально меняется — разрозненные точки группируются, поскольку теперь они суть составляющие собаки, а также улицы. Интересно и то, что стóит лишь один раз разглядеть собаку на этой картинке, как она так и останется навсегда собакой. Таким образом, прошлый опыт колоссально влияет на группировку и организацию, особенно если изображение неопределенно.

Фактор прошлого опыта тем отличен от прочих вертхаймеровых факторов, что не рассматривает собственно стимуляцию (в частности, ее геометрию), но *визуальную историю*, относящуюся к этой стимуляции. Однако прочие факторы группировки могут



**Рис. 6.1.15** Влияние прошлого опыта на перцептивную группировку. Если мы хотя бы раз разглядим фигуру далматинца на этом рисунке, то характер визуальной группировки *навсегда* изменится. Сие изменение можно отнести сугубо на счет прошлого опыта, оказывающего огромное влияние на перцептивную организацию неопределенных изображений.

Фото R. C. James.



**Рис. 6.1.16** Базовые факторы группировки могут взять верх над прошлым опытом. Перцепт параллелепипеда (А) может оказаться заблокированным прочими группирующими факторами (В).

блокировать детекцию знакомых объектов (к примеру Gottschaldt, 1929). На рис. 6.1.16 дан пример, в котором параллелепипед (рис. 6.1.16 А) полностью сокрыт комбинацией фигур (рис. 6.1.16 В). То есть, цельность, симметрия и прочие серьезные факторы практически полностью нейтрализуют распознавание.

Здесь следует учесть еще и то, что в отличие от примера с собакой, в котором точки исходно воспринимались неорганизованными, фигура на рис. 6.1.16 В сложена из плотно пригнанных блоков.

Влияние сближенности и распознанности объектов на перцептивную группировку весьма интересно, поскольку свидетельствует о том, что эффекты группировки возникают после объект-детекции. И сие закономерно, поскольку зафиксированное представление объекта несет информацию о том как различные его части сгруппированы и каковы отношения между ними. Если часть объекта (скажем, голова далматинца) идентифицирована первой, предыдущее знание о форме тела собак и их ног может оказаться использованным в реорганизации оставшейся части изображения. Сей пост-процесс указывает на то, что реорганизация эта возможно идет в течение всего акта восприятия, начинаясь на оптическом и заканчиваясь на категориальном этапе. При этом последующий результат замещает предыдущий.

## 6.2 РЕГИОНАЛЬНЫЙ АНАЛИЗ

Внимательный читатель легко отыщет пробел в гештальт-толковании вопросов перцептивной организации: гештальтисты тихо игнорируют объяснение того, откуда, собственно, берутся «элементы» их анализа. Такое впечатление, что Вертхаймер просто объявил о существовании таких элементов, полагая их феноменологически очевидными и не требующими доказательств. В таком случае, это пример одной из серьезных исследовательских ошибок, за которые, кстати говоря, гештальтисты ретиво критикуют других. На самом деле элементы вертхаймеровской демонстрации получены из структуры стимула опосредованно и требуют объяснения, включающего анализ факторов, управляющих существованием этих элементов как перцептивных объектов.

Очевидным базисом элементов перцептивного опыта, исходя из которого Вертхаймер рассуждает о принципах группировки — является анализ участков (регионов): отграниченных двумерных областей, составляющих пространственное подмножество ретинальных изображений. О концепции регионов как о базисе обработки ретинального изображения, мы еще не говорили в явном виде, сосредоточившись главным образом на одномерных конструкциях из линий и краев: в главе 4, к примеру, мы в основном обсуждали локализацию одномерных контуров, но почти ничего не сказали о том, почему контуры эти столь важны; в пятой главе мы обнаружили, что контуры задействуются в некоторых ключевых процессах восприятия глубины пространства, в частности краевой интерпретации и извлечения перспективной информации. Теперь же настало время рассмотреть другой важный аспект их перцептивной функции — *отграничивание двумерных областей*. Отграниченный участок — это ядро перцептивной организации, определяющее первый уровень зрительной обработки — уровень полноценных двумерных перцептивных единиц (на котором базируются дальнейшие процессы).

### 6.2.1 Равномерная связность

Пальмер и Рок (1994b) провели научный анализ того как вертхаймеровы гипотетические элементы формируются в логике некоего организационного принципа, который исследователь окрестил *равномерной связностью*: тенденцию нашего зрения целено воспринимать области однородных свойств — например, общих яркости, окраски, текстуры, движения и диспарантности. Т.е. о равномерно связанных участках Вертхай-



мер говорил, как об исходных единицах перцептивной организации,<sup>1</sup> и, как мы увидим далее, принцип равномерной связности соотносит наше представление о краевой детекции (см. гл. 4) с перцептивными организацией и группировкой.

Итак, давайте рассмотрим элементы оригинальной вертхаймеровой демонстрации как пример того, что деление ретинального изображения по участкам на базе равномерной связности может возникнуть на ранних этапах перцептивной организации. Точки, линии и углы на рис. 6.1.2 А — 6.1.2 F — это участки, объединенные общей яркостью и они целиком соответствуют элементам, к которым Вертхаймер апеллировал в своем анализе группировки. V-фигура на рис. 6.1.3 А, линии на рис. 6.1.2 G и 6.1.2 H, X-формы на рис. 6.1.2 I и песочноформный контур на рис. 6.1.2 J, согласно анализу Пальмера — Рока тоже являются равномерно связными (правда, их отношение к вертхаймеровым «элементам» чуть сложнее и будет рассмотрено нами несколько позже).

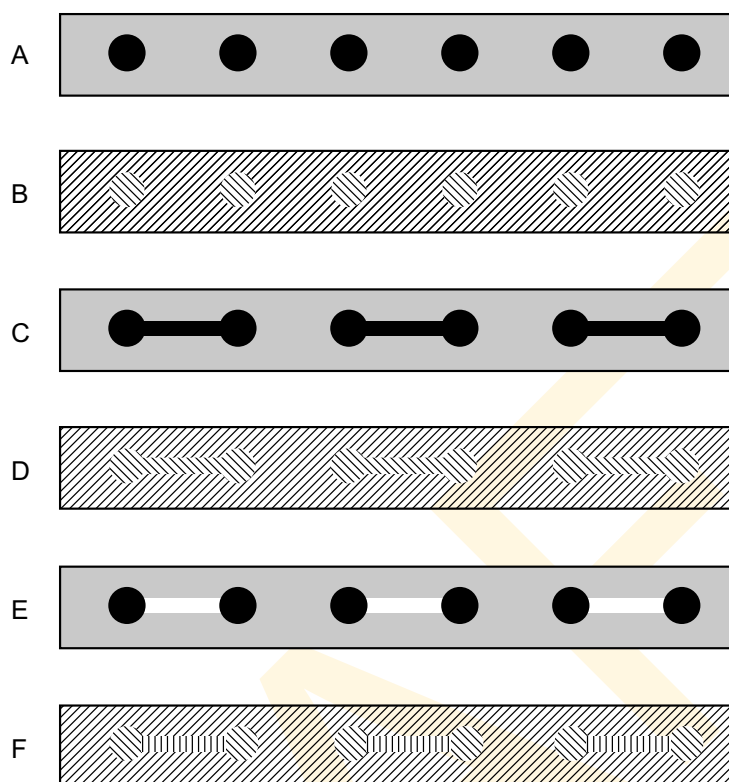
Мощь эффекта равномерной связности в перцептивной организации можно проиллюстрировать простым примером, в целом аналогичным вертхаймеровому: равномерно распределенные точки общей яркости на рис. 6.2.1 А смотрятся набором совершенно самостоятельных элементов; на рис. 6.2.1 В — аналогично, но по текстурам общей ориентации. Рис. 6.2.1 С и D демонстрирует то, что когда точки соединены участками общих с ними свойств, они сливаются в крупный, сложный и цельный элемент. Однако по рис. 6.2.1 Е и рис. 6.2.1 F хорошо видно, что когда эти же точки соединены участками отличных свойств, визуального объединения не происходит.

Может показаться однако, что равномерная связность — не более чем подобие по яркости и хроматике. К примеру, если тонкий лучик света, падающий на каждый ретинальный фоторецептор, взять как элемент, то разве нельзя равномерно связные участки объяснить группировкой этих элементов по яркости и хроматике? Возможно, Вертхаймер так и рассуждал об организации элементов. Но сходства по яркости и хроматике недостаточно для толкования перцептивного единства равномерно связанных элементов, поскольку оно не объясняет отличие между *связными* участками одного цвета и *несвязными*. То есть, без дополнительных ограничений в связности у нас отсутствует базис прогнозирования того, что два черных участка внутри одной и той же точки или штриха теснее связаны друг с другом, чем сравнимые с ними черные участки внутри разных точек или штрихов.<sup>2</sup> С феноменологических позиций, разумеется, одна отдельная точка плотнее организована как перцептивный объект, нежели любая пара раздельных точек. Сие наблюдение подталкивает нас к мысли, что равномерная связность — это один из наиглавнейших принципов перцептивной организации.

Стоит сказать, что Пальмер и Рок (1994а, 1994b) предупреждают от соблазна «разжаловать» равномерную связность с базового принципа перцептивной организации до принципа группировки. Равномерная связность не является принципом группиров-

<sup>1</sup> Кофка (1935) предвосхитил идею равномерной связности в своей статье о перцептивной организации, но не сделал выводов из своих научных наблюдений: его наброски никак не тянули на завершённую теорию до тех пор, пока не были пересмотрены в концепции Пальмера — Рока (1994а).

<sup>2</sup> Иной путь учета сего факта — это т.н. *ассоциативная группировка* (Geisler), идея которой в том, что если А сгруппирован с В и В сгруппирован с С, то А должен группироваться с С. Сия гипотеза может истолковать то, почему точки внутри равномерно связанного участка группируются плотнее, чем в участке обычном.



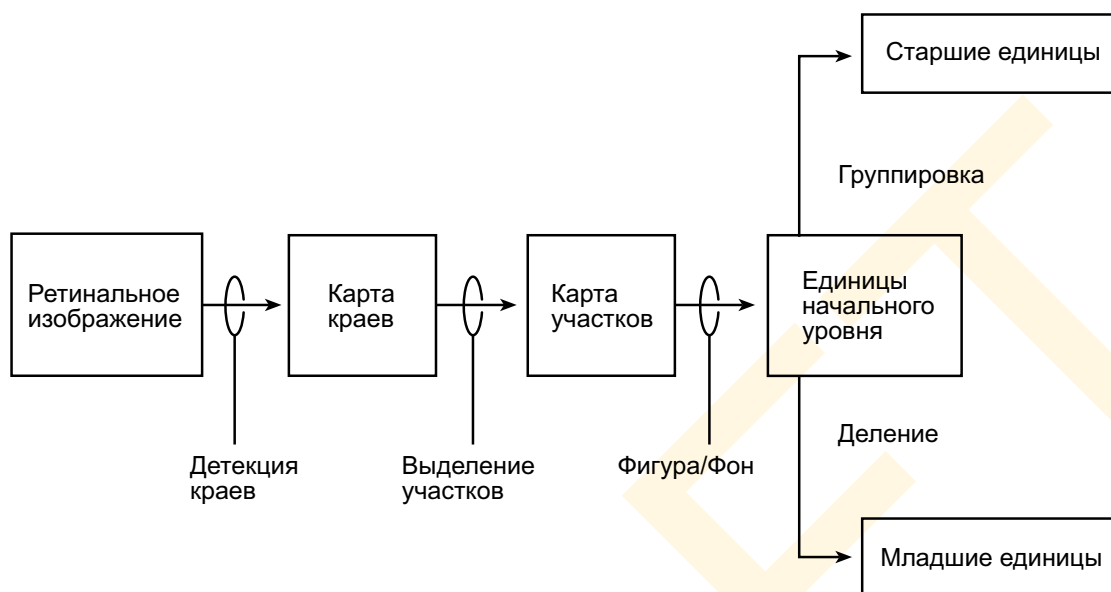
**Рис. 6.2.1** Равномерная связность. Связные участки равномерных визуальных свойств мы воспринимаем как самостоятельные, цельные элементы, если они определены общей яркостью (А и С), текстурой (В и D), либо какими-то другими простыми визуальными параметрами. Сходные элементы, но определенные разными свойствами (Е и F), лишены общеунитарного характера, как те, что определены равномерной связностью.

ки,<sup>1</sup> поскольку последний предполагает существование независимых элементов, группирующихся в определенных обстоятельствах, тогда как равномерная связность относится к цельному образу. По этой причине равномерная связность должна логически работать до того, как заработают принципы группировки. Проще говоря, равномерная связность — это принцип, отвечающий, в первую очередь, за *формирование элементов*, и который должен отрабатывать *до* процесса, работающего с этими элементами.

Итак, коль скоро равномерная связность столь важна и фундаментальна в перцептивной организации, то важно понимать, почему. Пальмер и Рок убедительно показали, что благодаря ее огромному информационному значению в диагностике связанных объектов (или частей объектов) в окружающем мире:

⇒ Если участок ретинального изображения представляет собой равномерно связанную область — сие однозначно плод стимуляции потоками, отраженными от единичного цельного объекта в сцене.

<sup>1</sup> Отметим, что элементную связность мы полагаем принципом группировки, но связность *равномерную* — нет. В концепции Пальмера — Рока сие принципиально разные факторы перцептивной организации, имеющие принципиально разную интерпретацию.



**Рис. 6.2.2** Общая схема концепции перцептивной организации, предложенная Пальмером и Роком (1994а). После того, как зрение детектирует края и выделит участки, начинают работу процессы, отделяющие фигуры от фонов и формирующие тем самым перцептивные единицы начального уровня. Затем в любой последовательности идут группировка и деление, образуя старшие и младшие единицы в иерархии «часть — целое».

Разумеется, из правила имеются исключения, например, природная маскировка животных — рис. 6.1.8. И это еще один пример того, как восприятие путается, когда эвристика не справляется со своей задачей. Однако такие ситуации редки и, несомненно, равномерная связность — это прочная эвристическая основа отыскания в ретинальном изображении участков, согласованных с цельными объектами во внешнем мире. Следовательно, первое, что следует сделать зрительной системе в ее пообъектном анализе сцены — это разделить ретинальное изображение на равномерно связанные участки.

Итак, в основу рассуждений Пальмер и Рок положили то, что равномерная связность — это наипервейший принцип двумерной перцептивной организации, на который опирается всё остальное. Целью является деление ретинального изображения на множество самостоятельных участков, именуемых *сегментами*, которые сродни сегментам витража. Таким манером идентифицированные участки можно затем организовывать иными процессами, такими, в частности, как отделение фигуры от фона, группировка двух и более сегментов вместе, разделение цельного сегмента на два и более субсегмента и пр. Схема, предложенная Пальмером и Роком (1994 а), и отражающая отношения между процессами перцептивной организации, дана на рис. 6.2.2.

Далее мы рассмотрим дополнительные организационные процессы детекции «фигура — фон», а в следующей главе подробно поговорим о сегрегации. Но прежде мы должны детально рассмотреть то, как именно ретинальное изображение можно разделить на равномерно связанные сегменты.

### 6.2.2 Регион-сегментация

⇒ Процесс дробления ретинального изображения на самостоятельные и независимые друг от друга области, выделяемые по признаку связности оптических свойств<sup>1</sup> — именуется *регион-сегментацией* (или просто — *сегментацией*).

Принято считать, что регион-сегментация — это процесс, в основе которого лежит принцип равномерной связности, реализуемый зрением либо за счет отбора по сходству (подобию) соседних участков ретинального изображения (и сие логически вполне возможно), но все же есть веские причины полагать, что зрительная система фактически детектирует отличия (или градиенты) в визуальных свойствах локалей и тем отделяет один участок ретинального изображения от другого, т.е. — сегментирует его.

#### 6.2.2.1 Граничные механизмы

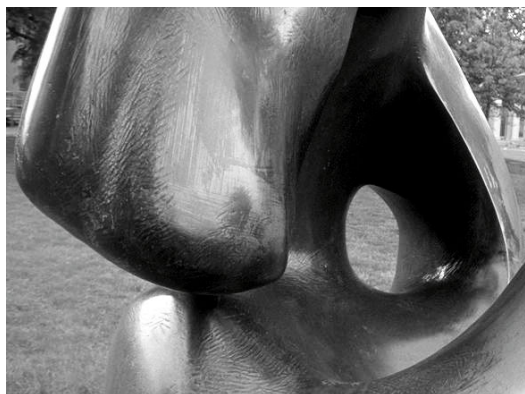
Косвенно мы уже коснулись граничного механизма регион-сегментации, когда говорили об алгоритмах краевой детекции (см. раздел 4.3.2). Всякий раз, когда края образуют замкнутый контур, требуется определить два участка: полностью отграниченный внутренний и частично отграниченный внешний. Следовательно, ретинальное изображение дробится на некое множество связанных друг с другом областей, используя для локализации замкнутых контуров алгоритм краевой детекции. Сия идея — это прочный теоретический мост между широко известными работами по физиологии и математике краевой детекции и гештальт-принципами перцептивной организации. Вспомним наш разговор о том, что одномерная краевая детекция может стать первым шагом к перцептивной организации, являясь при этом частью процесса двумерной регион-сегментации. Целью регион-сегментации, в свою очередь, является отыскание более или менее связных участков в ретинальном изображении.

Ранее мы уже описывали алгоритм краевой детекции, предложенный Марром и Хилдрет (см. раздел 4.3). Это простейший алгоритм и, разумеется, существуют другие, дающие сходные результаты, в частности, знаменитый алгоритм Канни (1986). Все эти схемы работают посредством эксплицитной локализации градиента фотометрической яркости: в алгоритме Марра — Хилдрет сие достигается сверткой изображения с множеством операторов второго порядка и нуль-детекцией на выходе. В данном аспекте одним из достоинств этого алгоритма является то, что края определены неизбежным проходом замкнутых контуров через ноль, чем они имплицитно определяют участок ретинального изображения (рис. 6.2.3).<sup>2</sup>

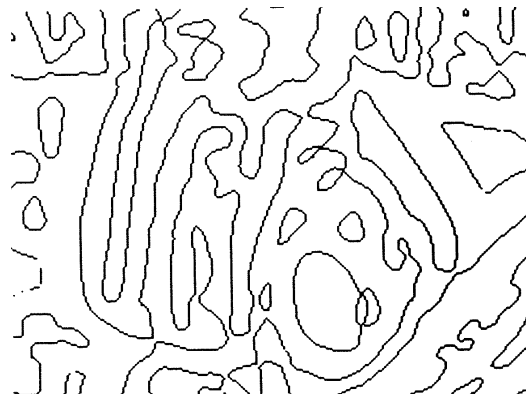
Таким образом, алгоритмы детекции яркостных краев можно использовать для деления ретинального изображения на участки более или менее связные по яркости. К примеру, при бесшумовой контурной стимуляции (в частности, такой, какую мы использовали для демонстрации классических принципов группировки — рис. 6.1.2) яркостных краевых алгоритмов будет достаточно для сегментации изображения на те же объекты, какие мы видим в реальной сцене. Однако в обработке большинства рети-

<sup>1</sup> В частности, таких как фотометрическая яркость, хроматика, текстура, движение или бинокулярная диспаратность.

<sup>2</sup> Не все краевые алгоритмы обладают столь удобным свойством. Тем, у которых его нет (например, алгоритм Канни) могут понадобиться дополнительные процессы группировки кусочковых краев в цельные контуры. Последние и определяют двумерные участки изображения.



А



В

**Рис. 6.2.3** Пример сегментации изображения, основанной на краевой детекции. В данном примере задействован алгоритм Марра — Хилдрет, позволивший разделить изображение скульптуры (А) на связные участки примерно одной фотометрической яркости (В). Одни участки соответствуют средним тонам объекта, другие — светам, третьи — теням, либо же представляют собой паразитные факторы.

По материалам Marr & Hildreth, 1980.

нальных изображений естественных сцен приходится учитывать множество других факторов.

Рассмотрим более выразительный пример: рис. 6.2.3 А — оригинал; рис. 6.2.3 В — результат сегментации оригинала по алгоритму нуль-детекции. Отметим, что край-оператор локализовал несколько перцептивно сомнительных участков. Некоторые из них — это результат относительно малых отличий в яркости и их можно исключить простым объединением соседних участков, краевой контраст между которыми достаточно низок (Yakimovsky, 1976).

Однако же намного важнее то, что в ретинальном изображении наличествуют отличные друг от друга участки, определенные относительно высоким яркостным градиентом, которые согласуются с различными фрагментами одного и того же объекта реальной сцены. В большинстве случаев падающий свет активно взаимодействует с поверхностями, образуя тени и блики. Последние перцептивно объединяются за счет нейронных процессов, реализующих константность светлоты. К примеру, на рис. 6.2.3 имеется множество участков высокой фотометрической яркости, группирующихся с соседними темными областями и в совокупности образующих темную поверхность скульптуры. Собственно перцепт скульптуры равномерно темный, а глянцеви́тость поверхности — плод интерпретации светлых участков бликами.

Прочие примеры выделения участков в ретинальном изображении одного и того же объекта происходят из окклюзии. Когда один объект находится перед другим, закрывая часть последнего, в ретинальном изображении формируются два (и более) участка, которые мы видим как фрагменты одного и того же предмета. Затем в силу вступают процессы группировки, собирающие эти участки воедино.

Такие феномены как тени, блики и окклюзия говорят нам о том, что для определения финальной организации реалистичного ретинального изображения (как состоя-



щего из согласованных с реальными объектами элементов), помимо регион-сегментации требуется множество иных процессов.

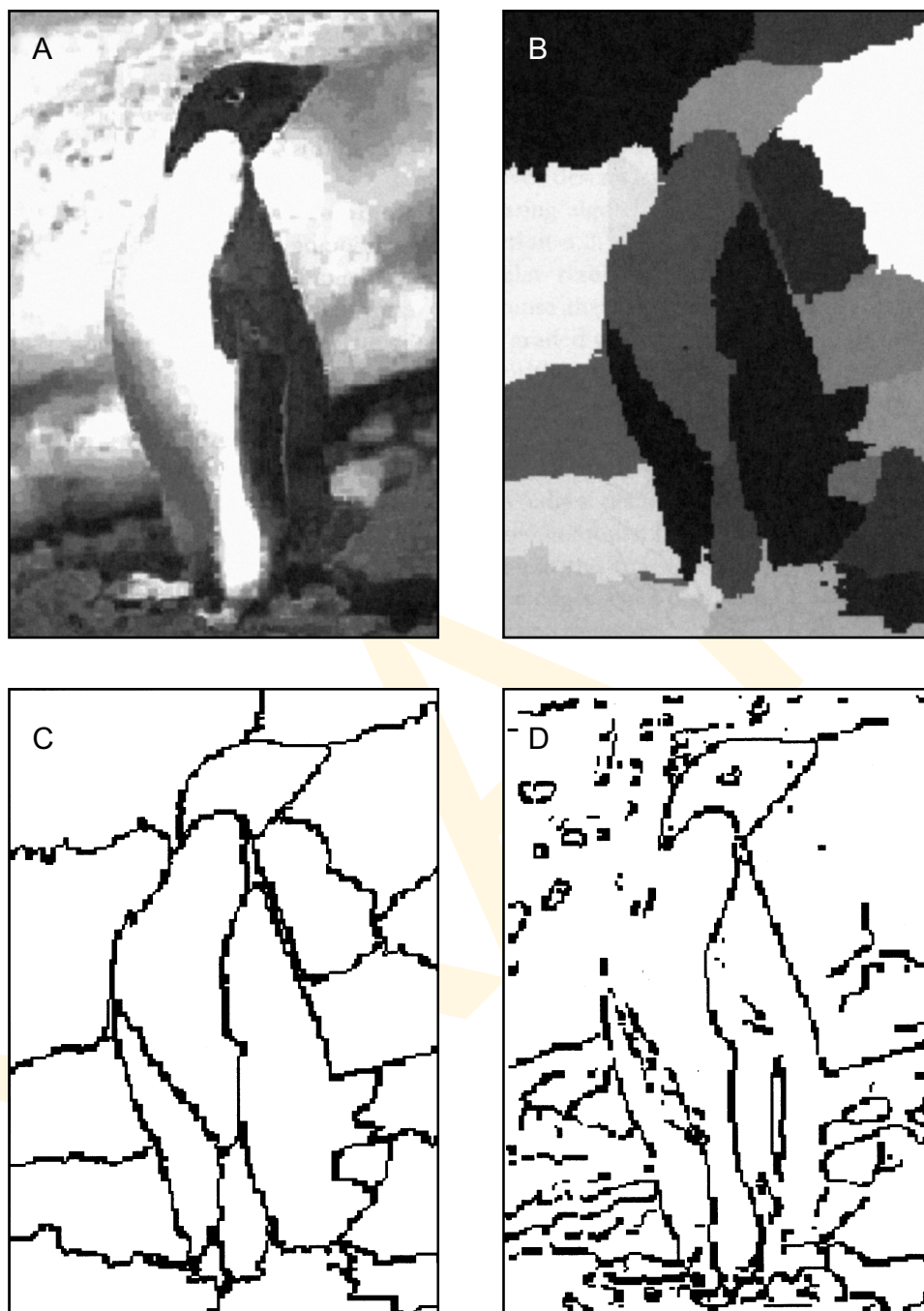
### 6.2.2.2 Региональный механизм

Заслуживает внимания тот факт, что граничные алгоритмы регион-сегментации радикально отличны от тех процессов, о которых говорят гештальт-психологи, когда рассуждают о феноменологии перцептивной группировки. Вертхаймеровы опыты и его гипотеза группировки через подобие свидетельствуют о неких глобальных процессах, «сочетающих» множество дискретных элементов, посредством, если угодно, «взаимопритяжения» за счет общих свойств. Отметим, что сие принципиально отличается от граничного принципа регион-сегментации, реализуемого через краевую детекцию: гештальт-подход свидетельствует о том, что вполне могут существовать иные механизмы регион-сегментации, не зависящие от предшествующих процессов локальной краевой детекции и обнаруживающие нужные участки почти что напрямую. В самом деле, возможно ведь даже то, что краевая детекция — это результат регион-сегментации, а не ее первопричина, то есть: если некий нейронный процесс каким-то образом отыскивает в ретинальном изображении равномерно связанные участки, края оказываются имплицитно (неявно) определены границами между этими участками. Как видим, сие диаметрально противоположно краевому механизму регион-сегментации.

Несмотря на то, что в целом все алгоритмы поиска границ, предложенные к сегодняшнему дню — исключительно локальны, их стоит рассмотреть глубже. Край-операторы, предложенные Марром и Хилдрет, Канни и др., отыскивают яркие градиенты, выполняя «вычисления» по очень мелким участкам ретинального изображения. Дабы определить области сегментации, найденные локальные края должно объединить во множество замкнутых контуров. Сей процесс математически довольно сложен и редко дает удовлетворительный результат (например рис. 6.2.4 D — работа краевого детектора Канни с изображением пингвина).

Более общую процедуру недавно предложил Дж. Малик (Leung & Malik, 1998; Shi & Malik, 1997): идея основана не на детекции краев или линий, а на прямом дроблении изображения на сегменты путем поиска пикселей, наиболее сходных внутри данной области и наиболее отличных между областями. Подобия в алгоритме Малика определены сходством по яркости, хроматике, текстуре и движению.

На рис. 6.2.4 дан пример изображения пингвина (часть A) и результат работы алгоритма (часть B). Границы участков (показаны в части C) весьма отличны от тех краев, что находят стандартные алгоритмы краевой детекции (D). Отметим, к примеру, что канни-алгоритм редко когда ловит связанные участки, тогда как региональный подход Малика отлично справляется с этой задачей. Конечно же, алгоритм Малика еще далек от совершенства, но добротен и, в первую очередь, ищет наиболее значимые области изображения, причем с минимумом паразитных разрозненных краев (что просто бич прочих методов). И хотя пока еще не совсем понятно, как такой сугубо технократический подход можно соотнести с фактической работой мозга, работа Малика — это серьезное напоминание о том, что локальный краевой принцип регион-сегментации — это далеко не единственный и далеко не лучший путь дробления ретинального изображения на равномерно связанные сегменты.



**Рис. 6.2.4** Пример регионального подхода к сегментации изображения. Маликов алгоритм идентифицирует равномерно связные участки (В) в изображении пингвина (А), используя при этом общий теоретико-графовый подход, опирающийся на максимальное сходство внутри участков и максимальное отличие между ними. Границы сегментов, выведенные оп алгоритму Малика — С, по алгоритму Канни — D. При этом отметим, что края в части D не дают добротного деления на связные участки.

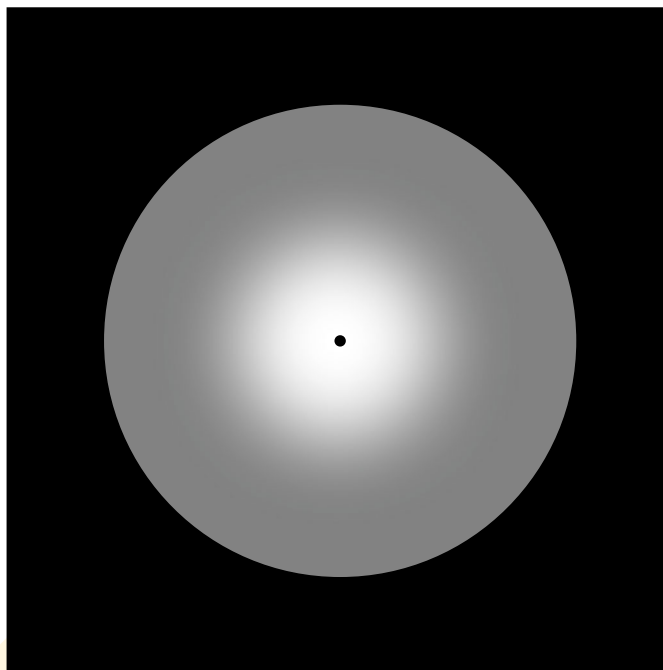
### 6.2.2.3 Стабилизированные ретинальные изображения

Итак, мы показали, что краевая детекция в целом может претендовать на роль механизма сегментации изображения, но пока еще не привели доказательств этого тезиса. Однако его уверенно поддерживают эксперименты по восприятию т.н. *стабилизированных ретинальных изображений* — изображений полностью стационарных на сетчатке (см. раздел 11.1). Удивительно то, что зрительная система фактически перестает откликаться (адаптируется) к оптической структуре такого стимула, и что после стимуляции в течение нескольких секунд эти изображения полностью теряют видность!<sup>1</sup>

Стабилизированные ретинальные изображения вполне адекватны идее краевой детекции как базиса региональной сегментации: они активно содействуют выводу о том, что в основе нашего восприятия формы и окраски того или иного участка объекта лежит в первую очередь краевая информация.

Возможно, наиболее убедительную демонстрацию дает нам простой, но весьма изящный эксперимент, поставленный американским нейрофизиологом Джоном Краускопфом (1963), предъявлявшим наблюдателям красный круг, окаймленный зеленым кольцом. В нормальных условиях просмотра стимул так и оставался красным кругом внутри зеленого кольца. Но Краускопф с помощью неких технических ухищрений стабилизировал внутренний контур между красным и зеленым участками, не трогая при этом всего остального. Результат оказался совершенно неожиданным: красный диск перестал восприниматься и полностью залился зеленым!

Сходный феномен мы можем наблюдать, примерно на минуту зафиксировав взгляд на точке в центре рис. 6.2.5. Поскольку мы не в силах полностью стабилизировать внутренний контур, он выполнен «размытым», дабы глазной тремор не оказывал на ход нашего опыта существенного влияния. В результате, по прошествии времени, нужного для зрительной адаптации, светлое пятно в центре перестанет восприниматься и заполнится серым цветом окружения. Как только мы расслабимся и глаза свобод-



**Рис. 6.2.5** Имитация стабилизированного контура. Сосредоточим взгляд на черной точке в центре картинке на 30-60 с. При этом обратим внимание на то, что светлое пятно в центре вначале начнет слегка «раскачиваться», а затем постепенно исчезнет, залившись окружающим серым (подробнее см. в тексте).

<sup>1</sup> Вспомним, что по этой самой причине мы не видим ни слепого пятна сетчатки, ни ее кровеносных сосудов.

но забегают по картинке, внутренний диск вновь появится, поскольку размытый край между областями вновь станет восприниматься.

Результаты краускопфова эксперимента таковы, как если бы и сам стимул, и границы областей задавались исключительно краевой информацией: коль скоро адаптация к внутреннему контуру полная (поскольку он стабилизирован на сетчатке), только внешний контур оказывает воздействие на зрение. То есть, ситуация идентична той, когда в стимуле представлен один лишь зеленый круг — наличие внешнего зеленого кольца без каких-либо внутренних границ. Сие в точности то, что мы воспринимаем, когда внутренний контур стабилизирован, причем вне зависимости от того, что физически внутренний красный круг продолжает упорно экспонировать сетчатку.

Сей примечательный феномен смотрится сугубо лабораторным эффектом, но на самом деле подобное встречается и в естественных условиях, например, в слепом пятне и в тенях от кровеносных сосудов, падающих на фоторецепторы — сие естественным путем стабилизированные изображения, поскольку они непрерывно двигаются синхронно с сетчаткой. То есть, мы буквально слепы к ним в обычных условиях зрения.

#### 6.2.2.4 Фрагменты и фрагментация

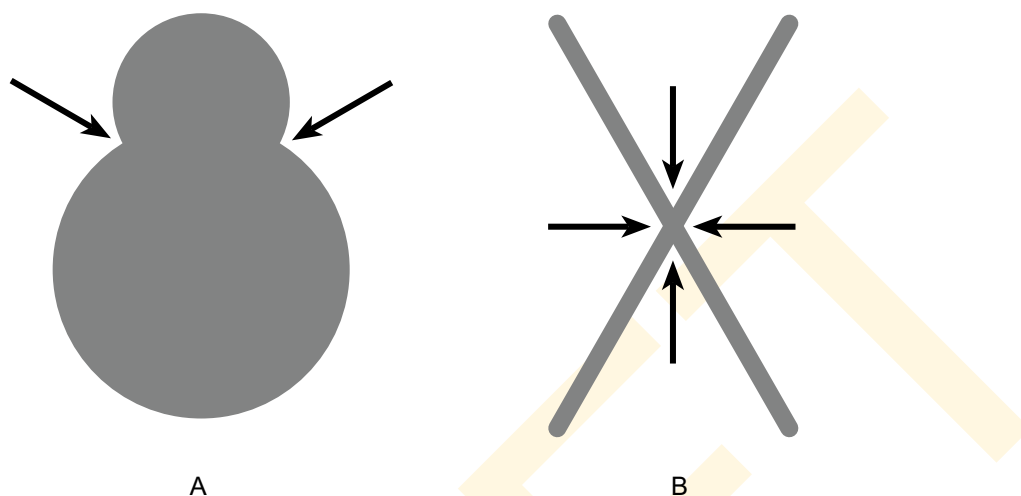
Итак, мы обсудили два базовых процесса перцептивной организации — регион-сегментацию и группировку. При этом резонно предположить, что регион-сегментация должна проходить раньше, являя собой весьма примитивный процесс, поскольку всего лишь определяет элементы, которыми оперируют процессы группировки. Но в восприятие объектов вовлечен еще и третий процесс — деление единичного элемента на составляющие. Сей процесс именуется *фрагментацией*.

Фрагментация — значимый аспект перцептивной организации, поскольку определяет то, что субрегионы перцептивной единицы восприняты как «сочетанные» максимально адекватно. Для иллюстрации вновь рассмотрим равномерно связанные сегменты рис. 6.2.1 С, памятуя что принцип равномерной связности определяет каждый сегмент как цельный и самостоятельный. При этом мы воспринимаем эти «гантельки» и как состоящие из четко читаемых фрагментов, а именно, — двух кружков, соединенных палочкой.

Рассмотрим также Х-образную фигуру на рис. 6.1.2I. Вопреки очевидной цельности анализ организации этой фигуры как состоящей из двух пересекающихся линий согласно вертхаймерову принципу цельности требует того, чтобы она вначале воспринималась состоящей из четырех сегментов, группирующихся затем попарно. По Пальмеру и Року (1994а), однако, перцепт цельной фигуры возникает первым, за которым идет дробление (фрагментация) на четыре сегмента. Однако сегменты эти могут затем сгруппироваться через вертхаймерову цельность.

Каким же образом фрагментация делит связанные сегменты ретинального изображения на составляющие? Во-первых, отметим, что налицо зависимость от формы объекта: залитый круг или прямая линия, к примеру, не имеют естественной фрагментации, тогда как два накладываются друг на друга круга или две пересекающихся линии — имеют (см. рис. 6.2.6). Если мы взглянем в места, в которых происходит естественное деление этих фигур, то обнаружим, что они разделены в местах т.н. *глубокой вогнутости* — т.е. в тех точках, где у контура имеется крутой изгиб внутрь — рис. 6.2.6 (Hoffman & Richards, 1984). Мы говорим, что фрагментация возникает там, где имеется *пара* таких вогнутостей.

Подробнее на фрагментах и фрагментации мы остановимся в гл. 7 и 8, когда будем детально разбирать восприятие формы объектов. Нынче же главным вопросом для нас



**Рис. 6.2.6** Фрагментация по глубоким вогнутостям. Стрелками показаны вогнутости в контурах равномерно связанных участков. Фрагментация обычно идет по парам таких вогнутостей.

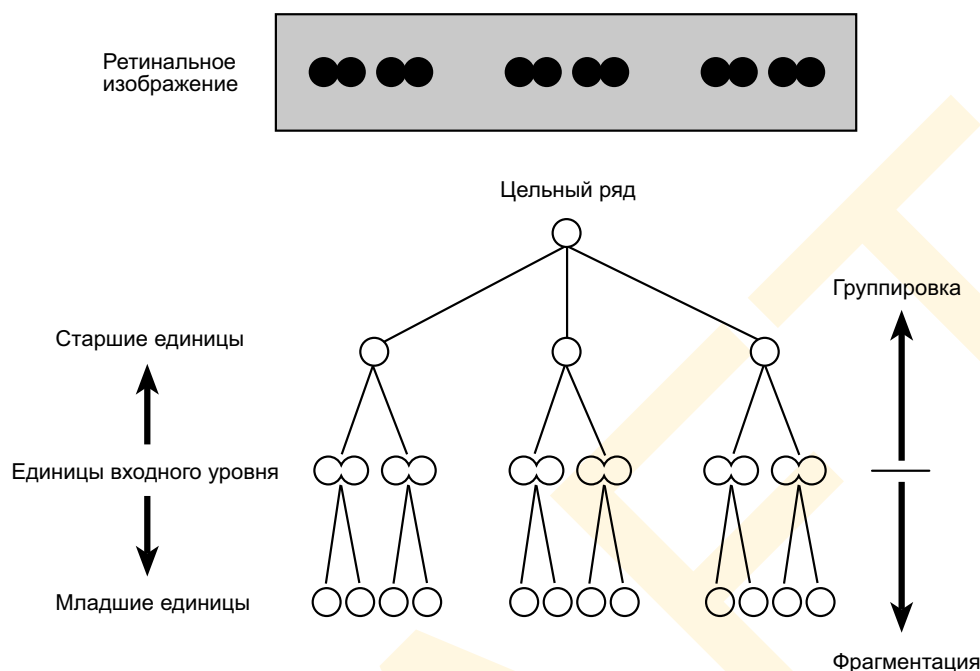
является то, как фрагментация соотносится с процессами региональной сегрегации и группировки.

Подобно группировке, фрагментация предполагает наличие исходного элемента, и нужно, чтобы региональная сегрегация сформировала такие элементы. Отметим при этом, что фрагментация диаметрально противоположна группировке, поскольку берет единичную, цельную перцептивную единицу (элемент) и делит ее на две и более единиц. Но при этом не существует ограничений на последовательность, в которой работают фрагментация и группировка — они вполне могут действовать одновременно. Почему так происходит иллюстрирует схема Пальмера — Рока (рис. 6.2.2), на которой показаны фрагментация и группировка, работающие параллельно, но при этом после того, как определены сегменты.

На рис. 6.2.7 Пальмер и Рок (1994а) дают пример того как три организационных процесса — региональная сегрегация, группировка и фрагментация — соотносятся друг с другом, достигая при этом сложной перцептивной организации. Исследуемый стимул дан сверху. Этот стимул мы можем описать как ряд, состоящий из трех парных фигур, однако заметим сперва, что сие описание предполагает некую иерархию: на высшем ее уровне находится верховная цельная сущность — ряд; на предыдущем ему уровне — составляющие его три парные фигуры. Еще ниже — компоненты пар. И, наконец, в самом низу располагаются «фрагменты», на которые эти фигуры зрительно разделены. Как достигнута сия иерархия?

Согласно концепции Пальмера — Рока первой должна отрабатывать региональная сегрегация, поскольку это единственный процесс, работающий с неорганизованными ретинальными изображениями. Процессы региональной сегрегации превращают различные участки нейронного ретинального изображения во множество непересекающихся, равномерно связанных участков. Затем в силу вступают процессы, именуемые «организацией «фигура — фон»». Организация «фигура — фон» определяет то, какие участки ретинального изображения следует рассматривать как объекты (фигуры), а какие как фон, оценивая каждый контур как принадлежащий к группе тех или дру-





**Рис. 6.2.7** Построение иерархии «часть — целое» согласно концепции Пальмера — Рока. Изображение вверху — массив стимуляции. Графы внизу — уровни организации, а также то, как их можно получить из единиц входного уровня, состоящих, в свою очередь, из равномерно связанных элементов.

гих. Участки, признанные фигурами, образуют первые единицы перцептивной организации и формируют входной массив иерархии класса «часть — целое» (см. рис. 6.2.7). Пальмер и Рок назвали единицы этого уровня *единицами входного уровня*.

Теперь, когда входной уровень определен, оставшуюся часть иерархии «часть — целое» можно достроить, применив процессы группировки и фрагментации, то есть: единицы входного уровня либо собираются (за счет процессов группировки) в единицы вышестоящие — старшие, либо (за счет процессов фрагментации) дробятся на единицы подчиненные — младшие. При этом ни кто не диктует нам последовательность, в которой оба эти процесса должны идти — они могут работать одновременно, расширяя систему «часть — целое» в обоих направлениях. В текущем примере группировка объединяет фигуры входного уровня в двухединичные группы (старшие фигуры на рис. 6.2.7), а затем из трех пар формирует цельный ряд. Уровень, подчиненный входному уровню, построен фрагментацией каждой равномерно связанной фигуры на два кружка в точках глубоких вогнутостей по центру каждой единицы входного уровня.

Итак, теоретический анализ Пальмера — Рока (1994а) навел мост через зияющий провал между классическими гештальт-идеями перцептивной организации и современными количественными концептами науки о зрении. Многое на этом пути уже пройдено, но, разумеется, далеко не всё (к примеру, Geissler; Rensink & Enns, 1995). Тот факт, что до сих пор не существует доказательства тому, что единицы входного уровня обрабатываются первыми в иерархии «часть — целое», говорит о том, что рассуждения Пальмера — Рока должны приниматься читателем с некоторой осторожностью.

### 6.2.3 Текстурная сегрегация

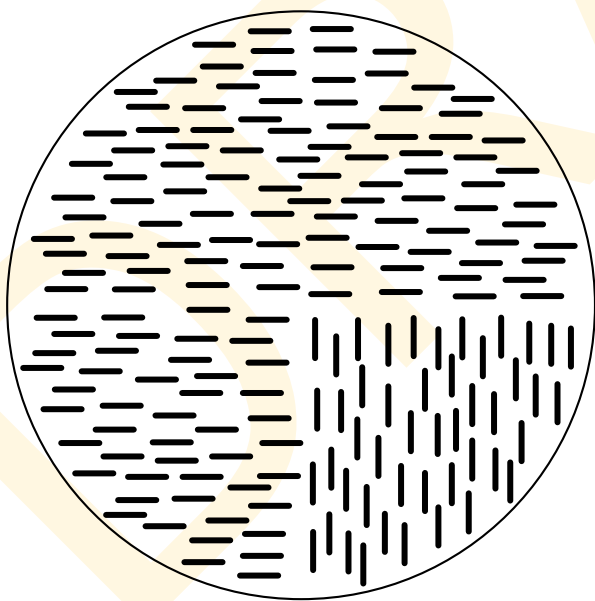
Пальмер и Рок (1994а) предположили, что равномерно связанные участки ретинального изображения вероятнее всего определяются параметрами более высокого уровня, нежели фотометрическая яркость и хроматика. Их рассуждения основаны на углубленном изучении вопроса т.н. *текстурной сегрегации*. Текстурная сегрегация — это принцип, согласно которому ретинальное изображение может сегментироваться исключительно на базе текстурной информации. И здесь было бы уместно дать простое и однозначное определение «текстуры», но, как мы увидим далее, сие далеко не просто.

Текстурная сегрегация обычно изучалась по изображениям, составленным из множества дискретных элементов, например, таких, что показаны на рис. 6.2.8: мы видим, что нижний правый квадрант легко отделяется от всей остальной части круга, даже вне осознанного восприятия формы отдельных его элементов. Очевидно, что текстурная сегрегация плотно связана и с классической группировкой, и с регион-сегментацией, но при этом по-разному.

Связь с классической группировкой заключена в том, что текстурную сегрегацию можно рассматривать как результат группировки текстурных элементов по классическому гештальт-принципу подобия. Однако, как уже отмечалось ранее, не все подобия одинаково эффективны (см. рис. 6.1.3), а Вертхаймер при этом ничего не говорил об особенностях разных видов подобия.

Текстурную сегрегацию можно понимать и как частный случай регион-сегментации, когда край-операторы отлавливают пространственные градиенты в т.н. «все еще неопределенных размерностях» текстуры, аналогично детекции резких перепадов яркости (гл. 4).

Прорыв в нашем понимании механизмов текстурной сегрегации произошел благодаря второму подходу, продвинувшему методы регион-сегментации до анализа текстурных свойств.



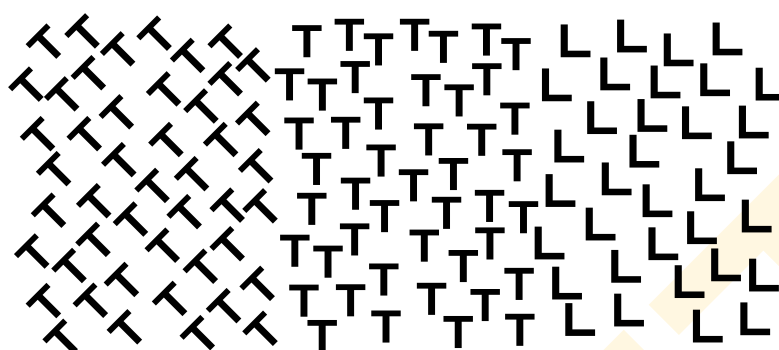
**Рис. 6.2.8** Текстурная сегрегация. Наше зрение с легкостью делит данное изображение на два участка по ориентации текстурных элементов.

По материалам Beck, 1966.

#### 6.2.3.1 Признаки текстуры

Первые серьезные исследования особенностей зрительного восприятия текстур выполнил американский физиолог Жакоб Бек (Университет штата Орегон). Перед его испытуемыми стояла задача определить, какой участок в стимуле, состоящем из множества дискретных текстурных элементов, наиболее отличен от других (рис. 6.2.9 А). К примеру, очевидно, что мы легко и быстро отделим наклоненные литеры «Т» слева на рис. 6.2.9 А от строго вертикальных в центре.

Бек начал с гипотезы о том, что стимул, состоящий из большого числа элементов группируется по сходству формы, то есть, согласно гештальт-принципу подобия. Но вскоре ученый обна-



А. Текстурная сегрегация



В. Подобие по форме

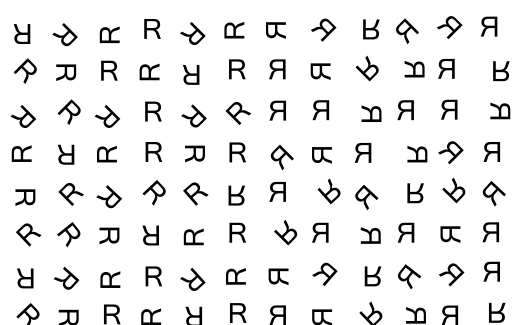
**Рис. 6.2.9** Текстурная сегрегация и подобие по форме. Текстуры увереннее делятся по признаку ориентации элементов, нежели по паттерн-идентичности (А), тогда как в восприятии подобия цельных фигур над ориентацией линий доминирует форма (В).

ружил, что факторы, управляющие текстурной сегрегацией могут быть иными. Например, явно выраженная текстурная сегрегация на рис. 6.2.9 А диаметрально противоположна прогнозу по подобию формы: когда испытуемые оценивают по форме те же самые элементы как самостоятельные фигуры (см. рис. 6.2.9 В), повернутая «Т» гораздо более тяготеет к «Т» вертикальной, чем к «L». По результатам множества аналогичных экспериментов Бек (1972, 1982) пришел к выводу, что текстурная сегрегация — это результат детекции отличий в т.н. *элементной плотности* (среднем количестве элементов на единицу площади) по простейшим атрибутам, таким как ориентация линий, общая яркость, хроматика, размер и движение.

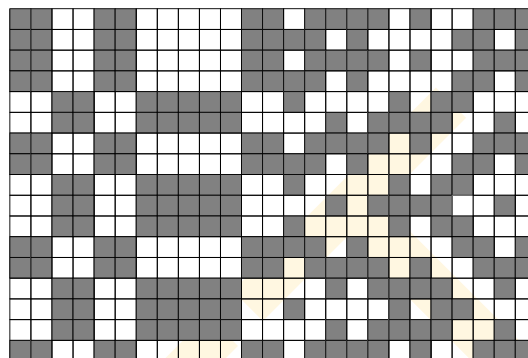
### 6.2.3.2 Текстурная сегрегация как параллельный процесс

Еще одну весьма дельную мысль по текстурной сегрегации подал Бела Юлеш — изобретатель стохастических стереограмм. Он показал, что текстуры различаются двумя возможными способами: либо через «обычную» текстурную сегрегацию (которая проходит легко и одновременно по всему полю зрения), либо посредством «осознанного рассматривания». Согласно Юлешу осознанное рассматривание задействует последовательную фокусировку внимания на различных частях предмета в попытке найти отличия по форме среди отдельных элементов. В итоге, исследователь предположил, что нормальная текстурная сегрегация является *преаттентивным* или, как еще говорят психологи — *предвнимательным* процессом, отрабатывающим перед тем как начнется фокусировка внимания.<sup>1</sup>

В своих исследованиях механизма текстурной сегрегации Юлеш изучил множество хитро построенных текстур, параллельно неделимых, несмотря на четко определенные физические отличия между составляющими их элементами. На рис. 6.2.10 А дан при-



А



В

**Рис. 6.2.10** Примеры юлешевых стимулов. А — физически различные текстуры (прямые и зеркальные литеры «Р»), неотличимые друг от друга без серьезного усилия. В — контрпример ранней юлешевой концепции различения текстур, основанной на дипольной статистике (подробности в тексте).

По материалам Julesz, 1981.

мер: нормально расположенные литеры «Р» слева не отличимы от зеркальных им справа без серьезного зрительного усилия. Ученый предположил, что физически отличные текстуры различимы зрением тогда и только тогда, когда они обладают разными *статистическими* свойствами (т.н. *дипольная статистика* изображения). Многие эксперименты подтвердили эту точку зрения (см. Julesz, 1975). Однако среди исследованных Юлешем стимулов были и такие, что показаны на рис. 6.2.10 В: картинка обладает статистикой второго и даже третьего порядков, но в ней не выделяются текстуры. Последнее вынудило Юлеша отказаться от рассуждений, основанных на статистических свойствах стимуляции и принять точку зрения Бека — исследователь склонился к тому, что текстуры делятся путем детекции изменений в плотности простых локальных текстурных признаков, которые он окрестил *текстонами* (Julesz, 1981).

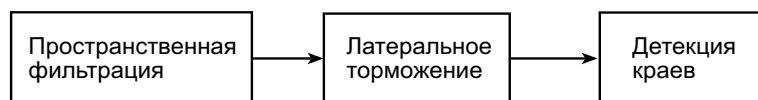
По Юлешу (1981) существует три типа текстонов:

- вытянутые пятна (характеризуемые окраской, длиной, шириной, ориентацией, бинокулярной диспаратностью и частотой мерцания);
- оконцовки линий;
- перекрестья линий.

Затем ученый показал, что в зрительной системе наличествуют детекторы, чувствительные к этим текстонам и что сегрегация идет через различную меру активации этих детекторов. Отметим, что юлешевы текстоны весьма схожи с приоритетными формами стимуляции клеток V1 по Хьюбелу и Визелю (1968), а также напоминают некоторые примитивные элементы маррова опорного скетча (1982).

Наконец отметим, что работы Юлеша и Бека по текстурной сегрегации исключительно ценны для нас, поскольку дают посыл к установу научных связей между дан-

<sup>1</sup> В разделе 11.2.5 мы покажем, что хотя такая обработка и происходит до фокусировки *внимания* на отдельных объектах, она не может возникнуть до начала работы общих механизмов внимания.



**Рис. 6.2.11** Общая схема трехстадийной текстурной сегрегации по Малику — Перона (подробнее см. в тексте).

ными о текстурной сегрегации, физиологией зрительной коры и количественными теориями зрения. И есть надежда, что сие приведет к созданию некоей общей концепции, описывающей эти связи.

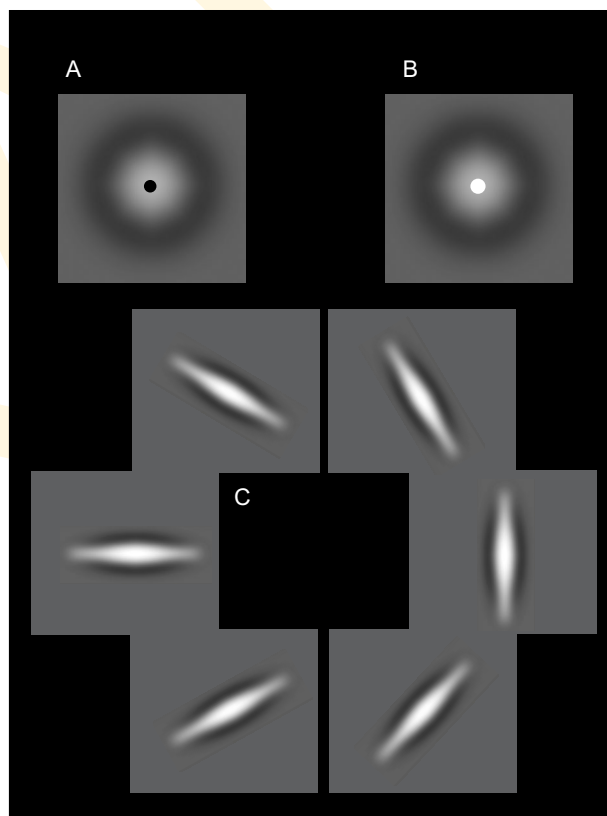
### 6.2.3.3 Модель Малика — Пероны

Сравнительно недавно Джитендра Малик и Петро Перона (1990) предложили биологически адекватную количественную концепцию текстурной сегрегации, основанную на детекции краев клетками V1. Суть этой концепции (которую мы коротко рассмотрели в разделе 4.3.3) сводится к трем этапам (рис. 6.2.11):

- этап фильтровки;
- этап латерального торможения;
- этап обсчета градиентов (краевая детекция).

Мы рассмотрим эту модель по двум причинам: во-первых, потому что сегодня она одна из наиболее удачных; во-вторых потому, что сие отличный пример междисциплинарного подхода к исследованию зрения — сочетание психологии восприятия, биологии и математики дает возможность продвинуться много дальше, чем позволяет каждая из этих дисциплин в отдельности.

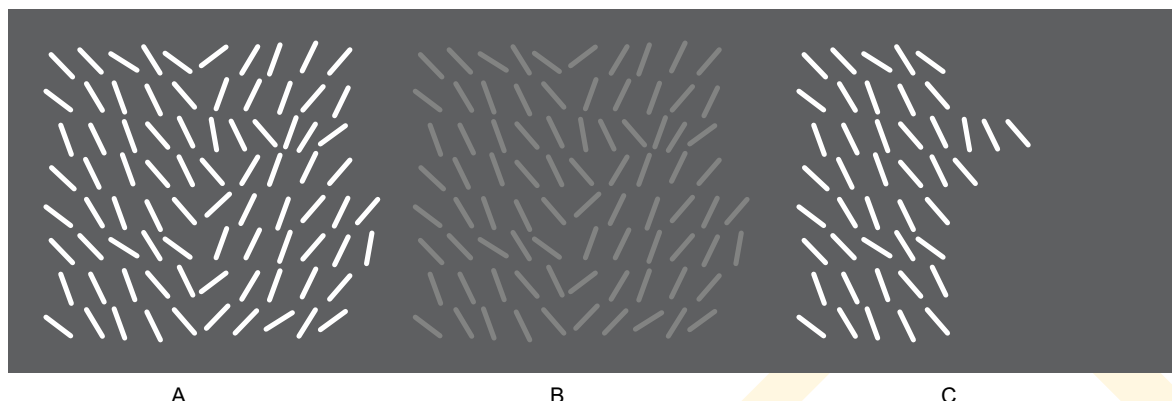
**Стадия фильтровки.** Итак, начальная стадия — стадия фильтровки — моделирует выходной сигнал клеток V1 (см. раздел 4.1) Фильтры, использованные Маликом и Перона, это рецептивные поля класса центр-окружение (рис. 6.2.12 А и 6.2.12 В), а также разноориентированные поля — *бардетекторы* (*детекторы штрихов*) — рис. 6.2.12 С. Поле зрения плотно заселяется фильтрами каждого из типов, в результате чего общий выходной сигнал по каждой пространственной позиции аппрокси-



**Рис. 6.2.12** Примеры фильтров Малика — Пероны, некоторые из которых — суть центрально симметричные рецептивные поля класса центр/окружение (А и В); другие — ориентированные бардетекторы (С). Аналогичные рецептивные поля обнаружены в V1.

По материалам Malik & Perona, 1990.





**Рис. 6.2.13** Примеры выходного сигнала стадии фильтровки (концепция Малика — Пероны). А — ретинальное изображение. В — выходной сигнал фильтров светлых штрихов, позволяющий отделить левую половину стимула от правой.

По материалам Malik & Perona, 1990.

мирует свертку ретинального изображения с фильтром того или иного типа, то есть, с соответствующим рецептивным полем.<sup>1</sup>

На рис. 6.2.13 А дана текстура из вытянутых пятен различных ориентаций. На рис. 6.2.13 В — выходной сигнал по каждой позиции этого изображения, выдаваемый множеством фильтров темных штрихов. На рис. 6.2.13 С дан выходной сигнал по каждой позиции изображения, выданный фильтром светлых штрихов той же ориентации. Выходные нейронные изображения являют собой результат свертки входного оптического изображения (часть А) с фильтром соответствующего типа. Светлота указывает на величину нейронной активности клетки V1 с данным рецептивным полем в соответствующей ретинальной локале. Мы говорим при этом, что выходное нейронное изображение примерно соответствует ретинотопической карте активности корковых клеток, с тем или иным специфическим рецептивным полем.

Отметим, что фильтр светлых штрихов дает много бóльший выходной сигнал на левой стороне текстуры (нежели на правой). Так происходит потому, что фильтр этот настроен на ориентацию пятен слева и совершенно интактен к той, что справа. Таким образом, имеет место существенное отличие (градиент) в работе данного фильтра по левой и правой половинам стимула. В выходном нейронном изображении это отличие представлено бóльшим отличием в общей яркости левой и правой сторон.<sup>2</sup> К сравнению: фильтры темных штрихов в той же ориентации (рис. 6.2.13 В) дают весьма слабый отклик на левой стороне и, следовательно, только очень легкий градиент по тому же самому входному ретинальному изображению.

**Латеральное торможение.** Второй этап работы модели Малика — Пероны позиционирует латеральное торможение пространственно соседних клеток с разными рецептивными полями. Смысл такого торможения — подавить (или хотя бы редуцировать)

<sup>1</sup> Понятие «свертка» толкуется в разделе 4.3.

<sup>2</sup> Отметим, что хотя сие отличие существует *имплицитно* в выходном сигнале первой стадии (стадии фильтровки), оно не детектируется *эксплицитно* до тех пор, пока не пройдет финальный этап работы модели.



**Рис. 6.2.14** Границы текстур в картине Густава Климта, найденные алгоритмом Малика — Пероны.

По материалам Malik & Perona, 1990.

ложные отклики (которые обычно весьма слабы). Посредством такого механизма, к примеру, слабые отклики фильтров темных штрихов (рис. 6.2.13 В) подавляются мощным сигналом фильтров штрихов светлых (рис. 6.2.13 С). Несмотря на то, что до сих пор нет полной уверенности в существовании ингибиторных связей в зрительной коре, модель Малика — Пероны не противоречит известным биологическим фактам и вполне согласуется с тем, что мы сегодня знаем об устройстве зрительной коры и ее функции.

**Краевая детекция.** Третья и заключительная стадия работы модели в явном виде просчитывает силу текстурных градиентов: она берет выходной сигнал стадии латерального торможения по фильтру каждого типа и анализирует этот сигнал через множество очень грубых (низкого разрешения) край-операторов в различных ориентациях и позициях, и по тем же схемам, что мы обсуждали в разделе 4.3.2, когда говорили о количественных концепциях краевой детекции. Операторы Малика — Пероны плотно соотносятся с текстурными элементами, то есть усредняют выходной сигнал множества отдельных фильтров с одинаковыми рецептивными полями. Раздельный анализ данного типа выполняется по каждому множеству рецептивных полей. Границы текстур отслеживаются по максимумам выходного сигнала.

Отметим, что у модели Малика — Пероны есть масса полезных свойств:

- во-первых, в отличие от концепций Бека и Юлеша, модель Малика — Пероны эксплицитно и полностью описывается в количественных показателях. Сие значит, что возможно (и действительно имеет место) внедрение этой модели в компьютерные программы, фактически исполняющие текстурную сегрегацию изображений;

- во-вторых, также вопреки Беку и Юлешу, модель применима к *любым* изображениям, а не только собранным из дискретных элементов. Иллюстрируя сие, Малик и Перона использовали компьютерную программу, отыскивающую текстурные границы в картине Густава Климта (рис. 6.2.14). И хотя мы помним о некоторых девиациях

в восприятии одного и того же стимула разными людьми, работа модели вполне уверенно регистрирует текстурные границы в сложных изображениях;

— в-третьих, мера того, насколько четко модель разделяет текстуры плотно согласуется с работой зрительной системы человека по тем же текстурам (Krose, 1986, 1987), а ее прогнозы (после шкалирования) четко садятся в данные, полученные на людях. На рис. 6.2.15 дано сравнение работы человеческого зрения и модели Малика — Пероны по 10 сильно разнящимся текстурам. И хотя модель еще далека от совершенства, результаты уже впечатляют;

— в-четвертых, модель Малика — Пероны интегрирует отдельные факторы в общий прогноз по любому стимулирующему изображению. Правда, задействуется лишь малая толика классических вертхаймеровых принципов группировки, а именно подобие по форме и по окраске. Однако, несомненно, сие хоть и первый, но очень важный шаг.

— в-пятых, текстурные размерности модели плотнее согласуются со структурой рецептивных полей клеток V1, нежели с естественными оптическими параметрами объектов сцены.

Таким образом, модель Малика — Пероны убедительно свидетельствует о том, что система обработки текстур может явиться одним из множества модулей, на которых основан выходной сигнал простых клеток зрительной коры (см. рис. 4.3.20).

Однако открытым сегодня остается вопрос о том, могут ли текстурные алгоритмы Малика — Пероны работать с классической вертхаймеровой группировкой по подобию? Здесь мы заняли такую позицию, что текстурную сегрегацию удобнее рассматривать как этап регион-сегментации, основанный на текстурных свойствах, нежели как этап классической группировки по подобию, но такой подход может оказаться и неверным. В таком случае алгоритмы, подобные алгоритму Малика — Пероны должно расширить до работы с обоими группами феноменов (1992).

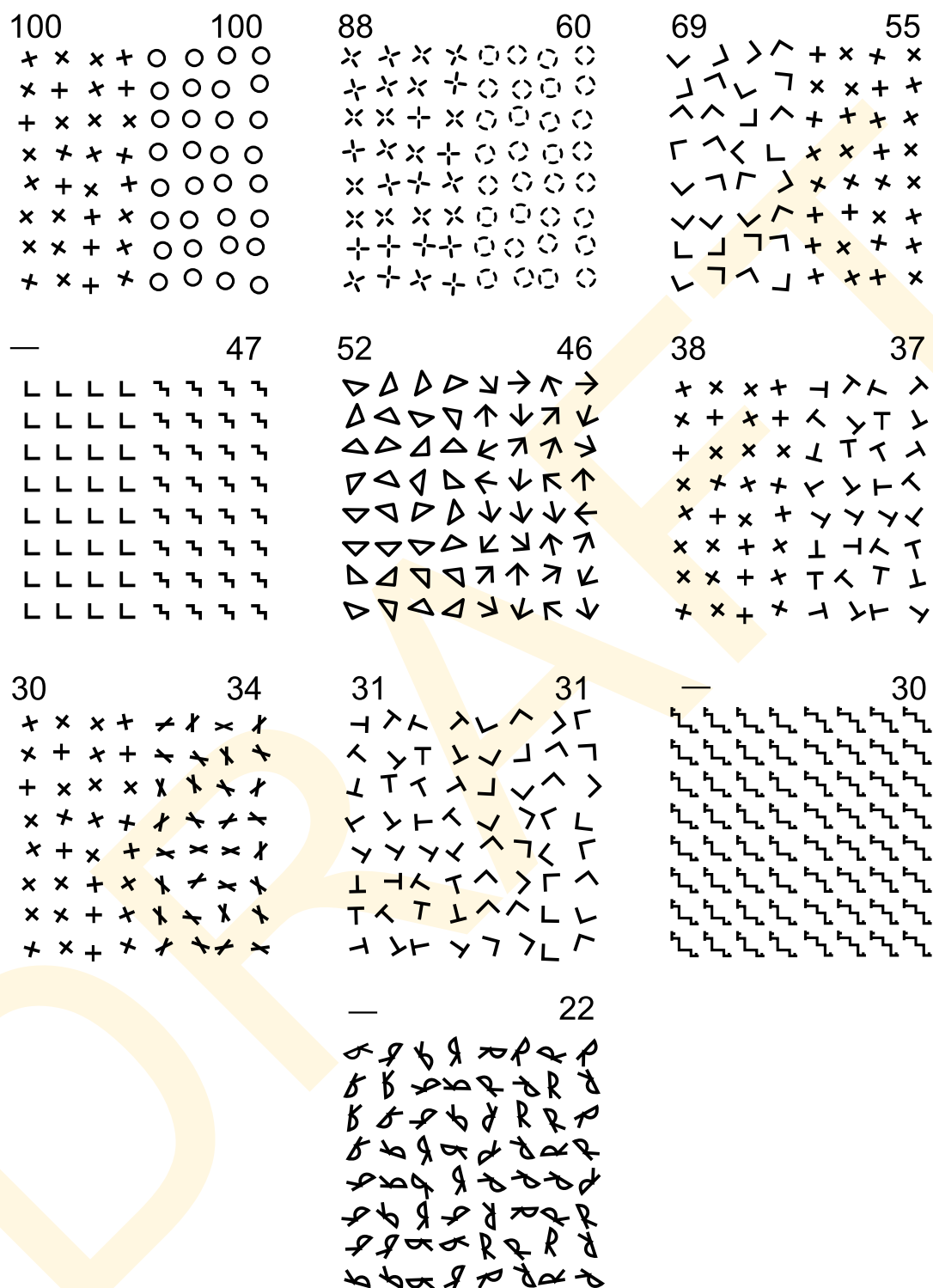
Итак, понятно, что прочие феномены группировки, в частности те, что базируются на сближенности, цельности и некоторых вариантах общего тока — это процессы особого рода, требующие особой теории. Позднее мы еще коснемся этого вопроса.

### 6.3 ОРГАНИЗАЦИЯ «ФИГУРА — ФОН»

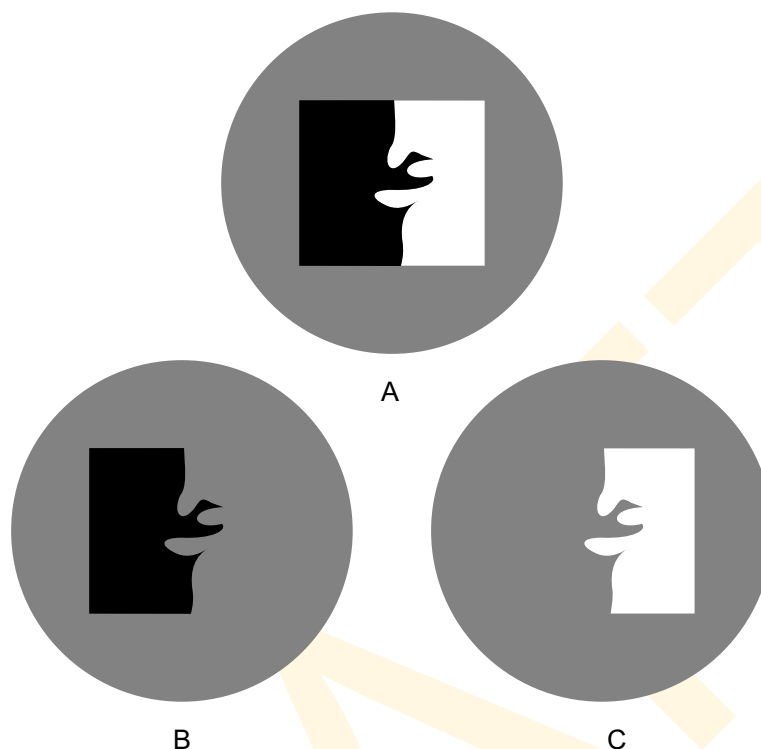
Ежели цель перцептивной организации — создание некоей иерархии объектов и групп объектов, то регион-сегментация — это не более чем первый шаг в этой работе, поскольку оптические участки ретинальных изображений редко напрямую согласуются с объектами естественных сцен. Даже при простейшей искусственной стимуляции, в частности такой, как на рис. 6.3.1 А, такого согласования не возникает. Регион-сегментация разделит это изображение на два равномерно связанных участка (один белый, другой черный), но при этом мы увидим либо цельный белый объект на черном фоне, либо черный на белом. То есть, мы говорим, что данный стимул совершенно неопределенный: то мы видим некую черную фигуру с «когтями», то наоборот — белую с «язычками» (как показано на отдельных и однозначно толкуемых рисунках внизу (6.3.1 В и 6.3.1 С).

Итак, сие еще один аспект перцептивной организации, который датский психолог Эдгар Рубин (1921) окрестил как *организация «фигура — фон»*. «Фигуроподобный» участок — это *фигура*; «фоноподобный» — *фон*.

Рубин выстроил свой концепт на феноменологическом анализе отличий между субъективными перцептами фигур и фонов (см. табл. 6.3.1): фигура всегда воспринимается расположенной ближе к наблюдателю и обладает формой, заданной контуром



**Рис. 6.2.15** Соответствие теорических данных экспериментальным в решении задачи текстурной сегрегации. Десять текстур помечены парой условных величин: слева величина различимости текстуры зрением человека; справа — прогноз, сделанный моделью Малика — Пероны. По материалам Malik & Perona, 1990.



**Рис. 6.3.1** Неопределенная организация фигура-фон. Черно-белый контур в центре части А относятся либо к левому, либо к правому участкам. Когтеподобный черный образ (В) — результат присвоения контура левому участку; пальцеподобный (С) — правому.

По материалам Rock, 1975.

(то есть контур «принадлежит» или «присвоен» фигурному участку, но не фону); фон при этом воспринимается как расположенный дальше и распространяется позади контура.<sup>1</sup>

**ТАБЛИЦА 6.3.1**

Фигура	Фон
Фигуроподобная	Не фигуроподобная
Ближе к наблюдателю	Дальше от наблюдателя
Отграниченная контуром	Распространяется позади контура
Форма задана контуром	Форма не зависит от контура

Экспериментальные результаты подтвердили рубинский тезис о том, что как собственно фигуру мы склонны воспринимать стимуляцию лишь по одну сторону от

<sup>1</sup> Рубин предположил также, что фигура выглядит плотнее, чем фон той же окраски, но это менее важное и достаточно спорное утверждение.

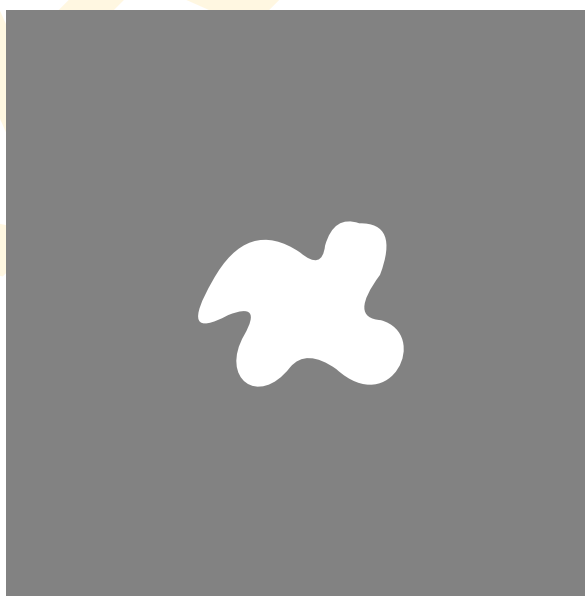


контура. Спустя много лет оригинальные данные Рубина (Rubin, 1921) были методологически усовершенствованы Роком (1983) на базе парадигмы т.н. *последовательно распознающей памяти*. Наблюдателям предъявлялась серия неопределенных стимулов класса «фигура — фон», в частности таких, что показаны на рис. 6.3.1 А. Половину испытуемых просили разглядеть белую фигуру половину — черную. После показа, наблюдателям предъявляли серии однозначных серых тестовых картинок, но лишь с одним участком: некоторые имели форму фигуры, некоторые — фона, а некоторых испытуемые ранее не видели. Увидев знакомую фигуру, наблюдатели говорили — «старая». Результаты опыта показали, что испытуемые хорошо запомнили форму тестовой фигуры, но форму фона — очень слабо, и отвечали скорее на удачу. Очевидный приоритет памяти именно на фигуры, а не на фон, говорит либо о том, что наблюдатели никогда ранее не воспринимали фон, как обладающий формой, либо о том, что они селективно забывают эту форму. Поскольку последнее представляется крайне маловероятным, Рок остановился на толковании, согласном с исходными посылами Рубина.

Итак, результаты опыта показали, что зрительная система отдает явное предпочтение только одной стороне отграниченной контуром области, а вторую сторону толкует однозначно как поверхность позади нее. Тот факт, что фигура всегда воспринимается, как лежащая впереди от фона, указывает на то, что организация класса «фигура — фон» теснейшим образом соотносится с восприятием глубины пространства, в частности с пикториальной информацией, извлекаемой из окклюзии. В поддержку этой идеи отметим, что все отличия между фигурой и фоном (сведены в таблицу 6.3.1), полностью согласуются с нашим представлением о восприятии глубины. Действительно, организацию класса «фигура — фон» можно интерпретировать как часть процессов краевой интерпретации (см. разделе 5.5.7), в частности тех, что отвечают за классификацию краев.

Теперь от неопределенного по критерию «фигура — фон» рис. 6.3.1 обратимся к рис. 6.3.2: мы моментально и однозначно увидим белую кляксу на сером фоне. Отметим при этом, что эту же картинку можно воспринять как неправильной формы дырку в сером квадрате, позади которой видна белая фоновая поверхность. Однако такой перцепт никогда не возникнет первым. Тот факт, что мы в первую очередь склонны к интерпретации белой фигуры как «кляксы» и что такое восприятие мощно доминирует над перцептом дырки в сером фоне, демонстрирует то, что зрительная система избирательна в интерпретации разных видов областей стимуляции как фигур.

Теперь нам предстоит разобраться в



**Рис. 6.3.2** Однозначная организация «фигура — фон». Белое пятно в центре исходно всегда воспринимается фигурой, поскольку окружено серым фоном, меньше его и более выпукло. И почти никогда оно не воспримется белой поверхностью, видимой сквозь дырку в квадратной серой плашке (по крайней мере до тех пор, пока явно не укажут на такую возможность).

том, как именно факторы стимуляции управляют делением на фигуры и фоны, и как сие соотносится с информацией, поступающей из внешнего оптического массива.

### 6.3.1 Факторы организации «фигура — фон»

Аналогично гештальт-принципам перцептивной группировки факторы организации «фигура — фон» подчиняются условию *ceteris paribus*, то есть, интересующий нас фактор дает определенный эффект, если все прочие исключены или нейтрализованы. Наиболее значимыми являются следующие:

#### 6.3.1.1 Окруженность

⇒ Когда при прочих равных одна область полностью окружена другой, то окруженный участок воспринимается как фигура, а окружающая его область — как фон.

Рис. 6.3.2 проиллюстрирует сей тезис: картинка легче читается как белое пятно на сером фоне, нежели как дырка в серой поверхности, позади которой виден белый фон. То есть, мы полагаем окруженность мощнейшим фактором деления стимуляции на фигуру и фон.

#### 6.3.1.2 Размер

⇒ При прочих равных меньший равномерно связный участок в стимуляции воспринимается как фигура.

Сей принцип четко работает и на рис. 6.3.2, поскольку белый участок очевидно меньше по размеру окружающего его серого поля. Однако «прочие» здесь не равны, поскольку на восприятие влияет окруженность. В этом отношении более чистым примером является вертушка на рис. 6.3.3 В, в которой черные лепестки, благодаря их меньшему размеру, воспринимаются как фигура на белом фоне. Сравним этот вариант с рис. 6.3.3 А, в котором лепестки одинаковы по размеру и, следовательно, максимально неопределенны в том смысле, какие из них полагать фигурой, какие фоном.

#### 6.3.1.3 Ориентация

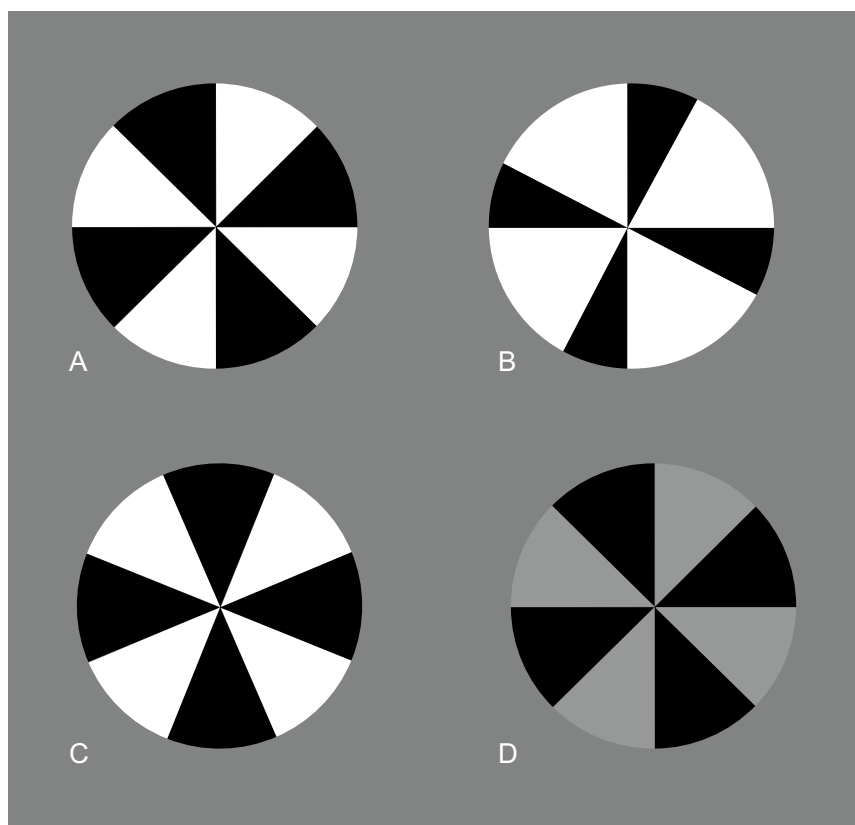
Если равновеликие лепестки сориентировать так, как показано рис. 6.3.3 С, то вертикальный черный крест скорее воспримется фигурой, чем белый косой. То есть:

⇒ При прочих равных ориентация — это фактор организации «фигура — фон».

#### 6.3.1.4 Контраст

⇒ При прочих равных участки с большим яркостным и/или хроматическим контрастом по отношению к фону воспринимаются как фигура.

Принцип проиллюстрирован рис. 6.3.3 D, в котором белые лепестки вертушки затемнены и контраст с фоном понижен, что повышает вероятность толкования черных лепестков как фигуры.



**Рис. 6.3.3** Принципы организации фигура-фон. А — высоко неопределенный стимул: и черные и белые лопасти вертушки могут равновелико толковаться как фигура. В, С и D — черные лопасти толкуются фигурой: в В — из-за относительно малого размера, в С — из-за строгой горизонтально-вертикальной ориентации, в D — из-за большего контраста с окружением.

### 6.3.1.5 Симметрия

По Рубину еще одним фактором организации «фигура — фон» является симметрия:

⇒ При прочих равных симметричные участки скорее воспримутся фигурой.

Эффект проиллюстрирован рис. 6.3.4 А, на котором мы явно видим симметричные черные фигуры на белом фоне.

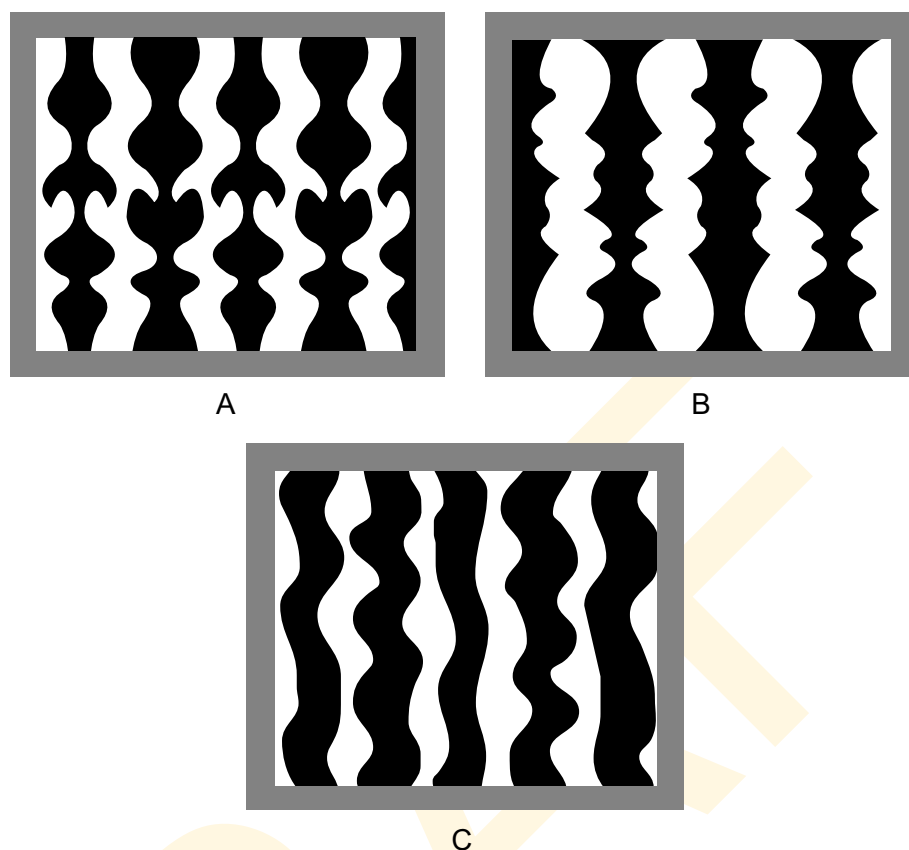
### 6.3.1.6 Выпуклость

Следующий фактор, определяющий деление «фигура — фон», — выпуклость:

⇒ При прочих равных, пятно с выпуклостями скорее воспримется как фигура, а с вогнутостями — как фон.

Фактор выпуклости иногда называют «законом внутренней стороны Метцгера»<sup>1</sup> (Metzger, 1953) и который позднее был проработан Каниша и Джербино (1976), пока-

<sup>1</sup> Вольфганг Метцгер — известный немецкий гештальт-психолог (1899 — 1979). — *Прим. пер.*



**Рис. 6.3.4** Симметрия, выпуклость и параллелизм в организации фигура-фон. А — симметричные участки склонны восприниматься фигурами. В — выпуклость может взять верх над симметрией. С — параллельность контуров как важный фактор восприятия стимула как фигуры.

По материалам Kanizsha & Gerbino, 1976 (А и В), а также Metzger, 1953 (С).

завшими, что выпуклость может доминировать над симметрией. На рис. 6.3.4 В, к примеру, даны ассиметричные выпуклые участки, которые мы скорее воспримем как фигуры, даже несмотря на то, что черные строго симметричны.

### 6.3.1.7 Параллельность

И еще один фактор — параллельность контуров. На рис. 6.3.4 С показан мецгеровский пример (1953), иллюстрирующий то, что:

⇒ При прочих равных участков со строго параллельными сторонами скорее воспримется фигурой, нежели фоном.

\* \* \*

Стоит сказать, что подчиненность описанных принципов критерию *ceteris paribus* — это их общее слабое место, в первую очередь, из-за неспособности к прогнозированию результата, когда различные конфликтующие факторы работают одновременно. К сожалению, со времен Рубина прогресса в этом секторе науки о зрении почти нет. И, хотя Кинкер, Сейновски, Хинтон и Шумахер (1986) предложили некую про-

стейшую архитектуру нейронной сети, дающей некоторые интересные варианты организации «фигура — фон», она так и не прошла систематической оценки на соответствие человеческому восприятию.

По ходу нашего сравнения организации «фигура — фон» с группировкой давайте задумаемся, почему нам раньше было необходимо, чтобы организация эта отработывала *до* того, как вступят в силу процессы группировки и фрагментации? В отношении группировки причина в том, что элементы входного уровня, с которыми работает группировка, должны уже быть отделены от фона поскольку, ежели сие не так, каждый участок будет группироваться со своим ближайшим соседом по принципу связности элементов (или же сближенности). Однако, ежели элементы входного уровня идентифицированы после того, как отработала организация «фигура — фон», они начинают отделяться друг от друга фоном и могут активно группироваться по принципам, описанным ранее.

Фрагментация объектов (см. раздел 7.6.2) по вполне понятным причинам требует предварительного деления стимуляции на фигуру и фон. Если фрагментация базируется на делении фигур по глубоким впадинам, то мы должны знать, какая сторона контура составляет фигуру, поскольку вогнутость (противу выпуклости) можно определить только относительно внутренности фигуры. Таким же манером искривленный контур, к примеру, является выпуклым в отношении одной внутренней стороны и вогнутым в отношении другой. Еще один тонкий момент: фон в своих вогнутостях не делится на фрагменты — только фигуры. Таким образом, «фигура — фон» организация должна всегда срабатывать *до* того, как начнутся фрагментация и группировка.

### 6.3.2 Экологические рассуждения

Некоторые полагают, что «фигура — фон» организацию можно реализовать, попросту отдав зрительное предпочтение какому-то одному участку в поле зрения. Однако на рис. 6.3.1, любой участок можно положить фигурой. Поэтому, несмотря на то, что зрительно мы с легкостью отделяем фигуры от их фонов, разговор о вопросах «фигура — фон» организации далеко не так прост, как может показаться на первый взгляд.

Наша способность отделять фон говорит о том, что «фигура — фон»-организация, вероятнее всего, происходит *до* того, как мы обратим внимание на что-либо. Таким образом, вероятнее всего, наше внимание *привлекаемо* фигурой, нежели ею определяемо. Иными словами, внимание управляемо «фигура — фон» организацией, а не наоборот. Существуют веские экологические предпосылки к тому, что мы в первую очередь выделяем из стимула фигуру, а уже после ее фон, поскольку фигура всегда ближе фона и, следовательно, представляет для особи первейший интерес. Однако внимание не может определяться одной лишь организацией «фигура — фон»: будь оно так, мы не смогли бы выделять фоны, даже если бы старались. Таким образом становится понятным, что внимание гибко переключается с фигуры на фон в зависимости от целей наблюдателя, но при этом всегда имеется жесткий приоритет именно фигур.

Экологические рассуждения подводят нас к функциональной гипотезе о том, почему зрительная система позиционирует «фигура — фон»-организацию и почему на нее оказывают влияние специфические факторы, описанные Рубиным и др. Функциональная роль организации «фигура — фон», вероятно, в том, чтобы одни участки двухмерного ретинального изображения оперативно соотнести с объектами сцены, а другие — с промежутками между этими объектами. Связь же с вниманием состоит в



том, чтобы переключать это внимание на анализ, в первую очередь, самих объектов, а не промежутков между ними (свойства последних второстепенны). Сие предпочтение объясняется рядом причин, но не только тем, что объекты как таковые экологически приоритетны.

Объекты сцены — это сущности, обладающие стабильными свойствами даже тогда, когда точка наблюдения меняется или когда движутся сами объекты: в конце концов, форма фонового участка в ретинальном изображении исключительно зависима от точки наблюдения и от объектов, закрывающих фон (Rock, 1983). Таким образом, скорее выделяя фигуры, нежели фоны, особь селективно обрабатывает т.н. *неакцидентальные признаки поля зрения*.<sup>1</sup> Контур всегда скорее задает фигуру, нежели фон, поскольку (если фигура ближе) так и происходит в реальности. Таким образом, принципы организации «фигура — фон» работают на веридикальность восприятия окружающего мира.

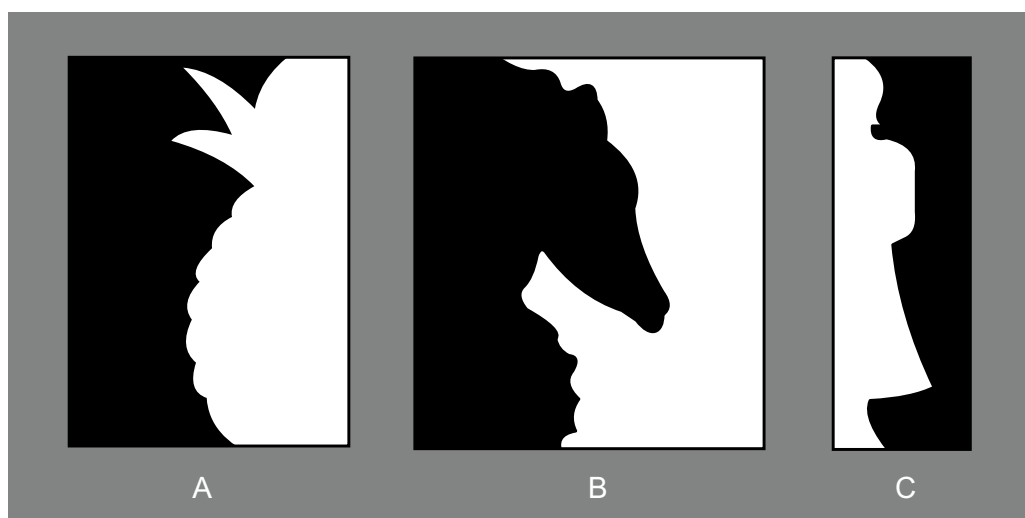
С этой точки зрения факторы, влияющие на организацию «фигура — фон», можно интерпретировать как относящиеся к визуальному отделению объектов от промежутков между этими объектами. Меньший размер и окруженность тоже могут иметь значение, поскольку всяко вероятнее то, что мы увидим мелкий объект на фоне большого, нежели дырку в большом объекте. Определенное значение может иметь и ориентация, поскольку большинство объектов реального мира склонны к большей гравитационной устойчивости, либо в горизонтальном, либо в вертикальном направлениях, нежели наклоненном. Реальные объекты более склонны к выпуклости, а их окружения, соответственно, — к вогнутости. Наконец, всегда вероятнее, что симметричные или параллельные стороны принадлежат объекту, нежели промежутку между объектами, где подобное может возникнуть лишь случайно. Таким образом, факторы организации «фигура — фон» систематически соотносятся со свойствами, отличающими двухмерные участки, согласованные с объектами сцены, от участков, согласованных с промежутками между этими объектами. Следовательно, эти свойства — весьма эффективная эвристика отделения тех участков ретинального изображения, что согласуются с объектами внешнего мира, от тех, что согласуются с фонами этих объектов.

### 6.3.3 Роль смыслов

Давеча мы показали, что процессы, лежащие в основе организации «фигура — фон» можно полагать способом, которым зрительная система человека определяет то, каким образом одни участки ретинальной проекции согласуются с объектами сцены (или главными сегментами этих объектов), а какие с частично закрытыми поверхностями позади первых. Ежели сие так, то вполне вероятно, что на организацию «фигура — фон» может оказать влияние форма участка ретинального изображения, который в большей степени согласуется с уже знакомым объектом. Недавние эксперименты американского психолога Мэри Питерсон показали, что организация «фигура — фон» действительно подвержена влиянию прошлого визуального опыта (Peterson & Gibson, 1991, 1993, 1994; Peterson, Harvey & Weidenbacher, 1991).

В экспериментах Питерсон испытуемым предъявлялись изображения, неопределенные по критерию «фигура — фон», в частности, что показаны на рис. 6.3.5. При этом половина каждого рисунка несла в себе определенный визуальный смысл (ананас, морской конек, девушка), а другая нет. Оба участка были более или менее равны в

<sup>1</sup> Мы вернемся к концепции неакцидентальности в конце текущей главы как к возможности унификации принципов перцептивной организации.



**Рис. 6.3.5** Роль смыслов в организации «фигура — фон». Неопределенная «фигура — фон»-стимуляция построена так, что один участок наполнен визуальным смыслом, а другой — нет. Смысловой участок наблюдатель, в первую очередь, воспринимает как фигуру, а ее перцепт сохранится весьма продолжительное время.

По материалам Peterson, Harvey & Weidenbacher, 1991.

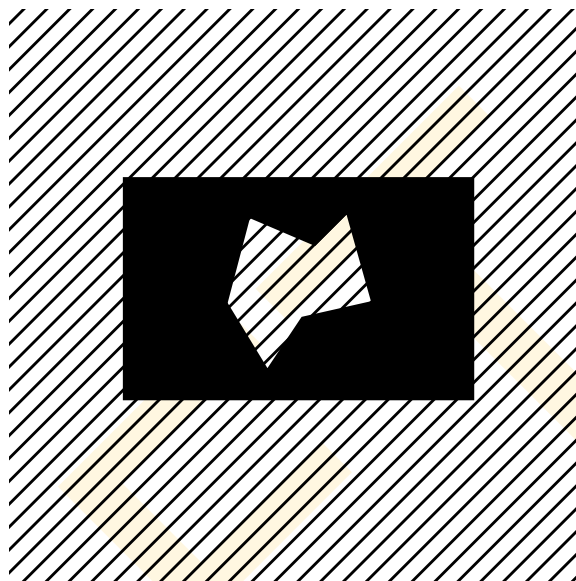
логике стандартных параметров «фигура — фон», таких как занимаемая площадь, окруженность и выпуклость. Испытуемых просили, чтобы, увидев белую фигуру на черном фоне, они нажимали белую кнопку, а черную фигуру на белом — черную. Результаты показали, что участок, согласующийся со смысловым объектом, воспринимался как фигура существенно дольше, чем участок бессмысленный. Испытуемые также были склонны воспринимать смысловой участок, в первую очередь, как фигуру.

Несмотря на то, что распределение по смысловым и несмысловым участкам выполнялось со всей тщательностью и по классическим факторам «фигура — фон», нельзя быть совершенно уверенным в том, что отличия в стимуляции помимо отличий в смыслах неспособны вызвать те же эффекты. Поэтому Питерсон сравнила продолжительность фигурной перцепции по смысловым и несмысловым участкам, когда одни и те же фигуры предъявлялись в нормальном и в перевернутом виде. Исследователь предположила, что смысловые объекты будут легко узнаваемы только в нормальном положении, и что какие бы то ни было отличия, найденные при нормальном положении, исчезнут при переворачивании вверх ногами. Так в точности и оказалось.

Результаты экспериментов Питерсон свидетельствует о том, что обработка «фигура — фон» подвержена влиянию нашего знания о специфической форме объекта. Отметим, что сей вывод полностью согласуется с тем рассуждением, что цель обработки «фигура — фон» — определить, какие участки ретинального изображения согласуются с объектами, а какие с промежутками между ними. Если участок ретинального изображения идентифицируется как сходный по форме с известным объектом, этот участок однозначно толкуется как объект, но не как промежуток.

Стоит сказать, что результаты Питерсон подняли интереснейшую проблему. Классическая точка зрения всегда была такова, что организация «фигура — фон» должна предшествовать распознаванию объекта, поскольку последнее требует наличия объекта-кандидата (скажем, фигуры) с которой идет работа. Питерсон (Peterson &

Gibson, 1994) интерпретировала свои результаты диаметрально противоположным образом, предположив вместо этого о наличии некоего «префигурального» процесса распознавания, работающего с обеими сторонами каждого контура *до* того, как начнется обработка «фигура — фон». Пальмер и Рок (1994b) показали, что логическую приоритетность «фигура — фон» можно восстановить, поскольку результаты в след идущих процессов распознавания объектов дают обратную связь на процессы «фигура — фон». Сие позволяет неопределенные стимулы «фигура — фон» (в частности, питерсеновские) определить дальнейшими процессами, в частности, процессами распознавания (причем скорее чем того требует предшествующая распознаванию объектов обработка «фигура — фон»). Недавно созданная модель, имитирующая некую нейронную сеть, продемонстрировала то, что данный вид архитектуры обратной связи фактически дает множество отличий, о которых и сообщала в свое время Питерсон (Vecera & O'Reilly, 1998). Какая из версий верна — пока не ясно, но можно не сомневаться в том факте, что процессы «фигура — фон» в какой-то мере подвержены влиянию процессов распознавания.



**Рис. 6.3.6** Отверстия в организации фигура-фон. Внутренность отверстия воспринимается фоном, но при этом обладает формой. Сие проявление конфликта с главной особенностью организации фигура-фон, то есть с тем, что границы задают форму только фигурам.

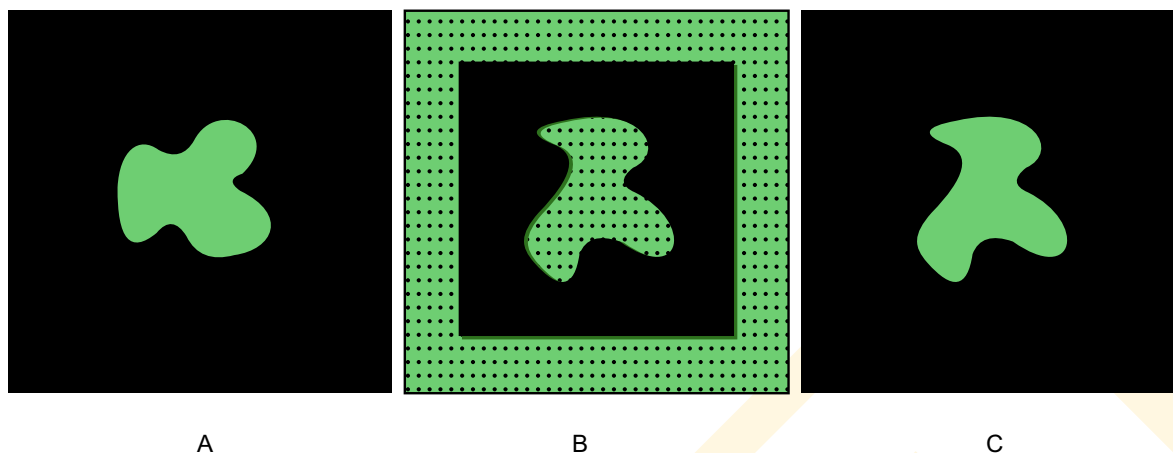
### 6.3.4 Отверстия

Выше мы говорили о том, что окруженная область может восприниматься либо как выпуклая фигура на некоем фоне, либо как отверстие в некоей поверхности, позади которой находится другая поверхность. Сия неопределенность вносит сумятицу в стандартный анализ отношений «фигура — фон». Однако начнем с определения:

⇒ Отверстия — это внутренние участки объектов или поверхностей, не содержащие вещества, материи (не считая воздуха).

То есть, мы говорим, что отверстие проникает сквозь весь объект, в результате чего становится видимой поверхность позади него (рис. 6.3.6). Главная сложность в этой ситуации — это т.н. псевдофигуральный статус отверстий. Это значит, что отверстия воспринимаются самостоятельными феноменологическими сущностями, имеющими свою собственную форму, даже несмотря на то, что фактически представляют собой лишь пустое место (Casati & Varzi, 1994).

Недавние эксперименты Рока, Пальмера и Хьюма показали, что отверстия действительно воспринимаются, как нечто обладающее формой. Исследователи предъявляли испытуемым комплекты двухмерных объектов, половина из которых представляла собой небольшие зеленые картонки неправильной формы, расположенные на черном



**Рис. 6.3.7** Стимулы в эксперименте Рока — Пальмера — Хьюма. Половина из них представляла собой бесформенное пятно на зеленом фоне (А). Половина — отверстие в зеленой плашке, в которой виднелся некий фон (В). Память на форму фигур и отверстий проверялась предъявлением тестовых предметов, чья форма была идентична форме ранее показанных фигур или отверстий. Предъявлялись также отвлекающие фигуры, ранее не показанные.

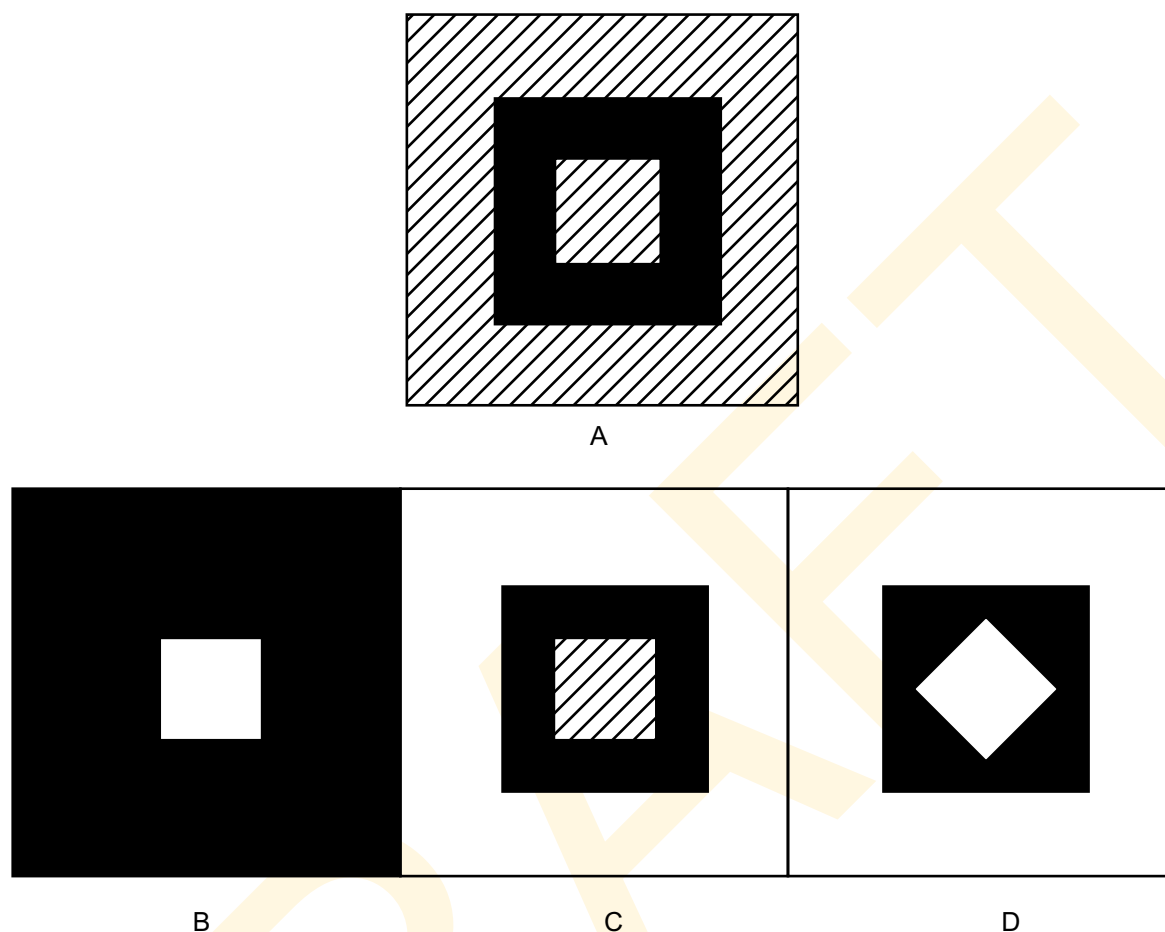
фоне (рис. 6.3.7 А), а половина — крупные черные картонные квадраты с отверстиями, расположенные на расстоянии нескольких сантиметров от зеленого фона, покрытого точками (рис. 6.3.7 В).<sup>1</sup> После просмотра основных образцов (которые воспринимались испытуемыми устойчиво веридикально, то есть — либо как сплошные, либо с отверстиями) наблюдателям предъявили сплошные тест-объекты (без отверстий), некоторые из которых были теми же пятнами (часть А), некоторые имели форму отверстий (то есть, дырок в квадратных картонках — часть С), а некоторых наблюдатели ранее не видели. Задача состояла в том, чтобы, судя по форме, выбрать виденные ранее.

Рок, Пальмер и Хьюм обнаружили, что, несмотря на исходные теоретические ожидания согласные классическим понятиям о фигурах и фонах, наша опознающая память работает одинаково и в отношении отверстий, и в отношении фигур. То есть, удалось показать, что зрение скорее припишет форму отверстию, нежели воспримет стимуляцию как бесформенный фон позади фигуры. Ряд прочих экспериментов подтвердили эти выводы.

Стоит отметить, что в рамках стандартной концепции «фигура — фон» тот факт, что отверстия воспринимаются имеющими форму, ведет к парадоксу: ежели контур отверстия принадлежит окружающему объекту (как, собственно, и должно быть, когда отверстие воспринимается как отверстие), то в результате должна вырисовываться форма *окружающего* объекта. Но каким образом наблюдатель видит отверстие, как обладающее собственной формой? Получается, что внутренний контур выполняет двойную работу, задавая форму участков по обеим своим сторонам.

Один из возможных ответов звучит так, что восприятие фигуры (т.е. участка, имеющего детерминированную форму) может сильно отличаться от процессов, фактически представляющих форму данного участка ретинального изображения. Мы мо-

<sup>1</sup> Точки были нанесены специально, дабы предоставить зрению испытуемого необходимую стереоскопическую информацию.



**Рис. 6.3.8** Факторы, влияющие на восприятие отверстий. Перцепт отверстия возникнет скорее, если окружающий участок смотрится фигурой, если очевидна цельность фона и если отношения с внешним контуром неакцидентальны (А). Когда одно из условий не выполняется, перцепт отверстия возникает с большим трудом.

жем воспринимать форму продырявленных объектов в логике объекта цельного с отверстием определенной формы в определенной позиции. Согласно сей точке зрения в определении фигуры всё однозначно — это участок, окружающий отверстие. Фигура эта тоже воспринимается как обладающая формой, заданной внутренней стороной своего внешнего контура и внешней стороной контура внутреннего. Но при этом форма пока еще имплицитна — эксплицитной она станет благодаря работе дальнейших процессов описания формы (см. гл. 8), один из которых работает с отверстиями в логике «цельный с дыркой». И это должен быть высокоуровневый описательный процесс, меняющий статус отверстия так, что его форму можно толковать в логике его внешнего контура, который может стать внутренним контуром объекта. Таким образом, отверстие — это, с одной стороны, основа определения отношений по глубине (а также того, какой материал находится позади открытого пространства), но с другой, отверстие — это фигура для описания формы.

Прочие интересные выводы касаются условий, в которых отверстие выглядит так, как показано на рис. 6.3.8 А: тот факт, что отверстие воспринимается нами именно



как отверстие, но не как некий маленький объект на передней поверхности отдельного большого объекта — требует пояснения. Действительно, когда факторов интерпретации стимула как отверстия недостаточно, внутренний участок воспримется так, как показано на рис. 6.3.8 В, 6.3.8 С и 6.3.8 D. Некоторые из факторов, склоняющих наше восприятие к перцепту отверстия следующие:

#### 1. Окружающая фигура

⇒ Перцепт отверстия требует того, чтобы окружающая область имела внешнюю границу.

То есть воспринималась фигурой, поскольку отверстие может существовать только внутри фигуры. На рис. 6.3.8 В дан пример, в котором окружение обычно воспринимается как фон, а белый квадрат — как фигура, но не как отверстие.

#### 2. Цельность фона.

⇒ Перцепт отверстия требует подтверждения тому, что видимая внутри отверстия поверхность — это та же самая поверхность, что и позади внешних границ объекта.

Многие факторы потенциально удовлетворяют этому требованию, и наиболее важные из них — это соответствие по окраске, текстуре либо цельность (явное продолжение) элементов фона. На рис. 6.3.8 С дан пример, в котором перцепт отверстия маловероятен, поскольку внутренний участок не согласуется с внешним.

#### 3. Отношения с внешней границей.

⇒ Перцепт отверстия требует того, чтобы внутренняя граница объекта неакцидентально соотносилась с внешней.

По крайней мере сие, усиливает восприятие внутреннего участка как отверстия. Перцепту отверстия также содействует параллельность контуров, что проиллюстрировано рис. 6.3.8 D, где небольшое изменение в ориентации внутреннего квадрата привело к тому, что квадрат этот воспринимается скорее как фигура на черном фоне, нежели как отверстие в нем (Bozzi, 1975).

\* \* \*

Итак, в целом, чем сильнее выражен каждый из перечисленных признаков, тем выше вероятность того, что внутренний участок объекта воспримется отверстием. Отметим, что сказанное об отверстиях, подчиняясь критерию *ceteris paribus*, не работает, если «прочие» окажутся неравными.

### 6.4. ЗРИТЕЛЬНАЯ ИНТЕРПОЛЯЦИЯ

Даже после того как участки сходных визуальных свойств сегментированы и фигуры отделены от своих фонов, зрительная система еще далека от завершения процесса организации поля зрения по объектам. Одним из основных осложняющих дело обстоятельств является то, что большинство поверхностей непрозрачны и постоянно закрывают от нас фрагменты предметов. По мере движения наблюдателя в пространстве частично закрытые поверхности то далее укрываются, то наоборот — открываются из под края ближних поверхностей (см. раздел 5.4.4). Очевидно, что зрению необходимо

что-то еще дополнительное, позволяющее справиться с кусочковостью видимых поверхностей и изменчивостью их компоновки, и сделать выводы о скрытых фрагментах. Мы полагаем, что в решении этого вопроса зрительная система задействует механизм, получивший название *зрительной интерполяции*.

Конечно же у этого механизма есть ряд ограничений, в первую очередь связанных с тем, что строится он на догадках о том, чего не видно. К примеру, полностью закрытые объекты редко когда интерполируемы, поскольку попросту нет оснований для такой интерполяции,<sup>1</sup> и даже частично видимые объекты зачастую реконструируются некорректно. Но даже в этом случае мы оказываемся неплохо информированными о характере полузакрытых предметов. Давайте прямо сейчас поглядим вокруг и попытаемся оценить число полузакрытых предметов в поле зрения: их на удивление много, но мы обычно не обращаем на сей факт никакого внимания.

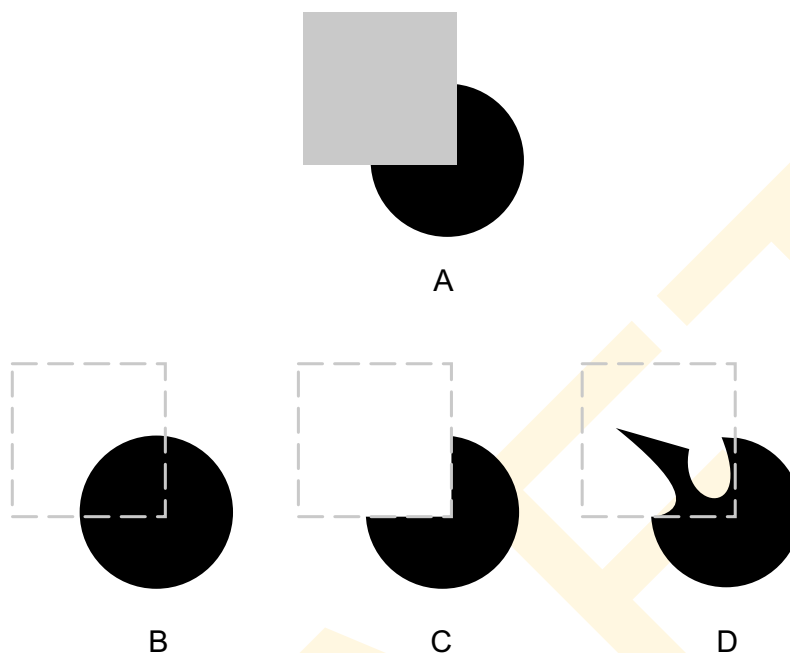
Итак, прежде чем приступить к описанию феноменов зрительной интерполяции, следует показать отличие между перцептивной завершенностью (возникающей, когда поверхность частично закрыта) и заполнением (например как у слепого пятна): заполнение обладает всей полнотой зрительного ощущения соответствующего сенсорного качества по отсутствующему участку ретинального изображения; зрительная же интерполяция не дает ощущения цельной поверхности, но лишь *перцептивное знание* — или, точнее, *уверенность* в ее свойствах.

### 6.4.1 Зрительная завершенность

Феномен *зрительной завершенности* (или как его еще называют *амодальной завершенности*) состоит в том, что зрительная система автоматически воспринимает частично закрытые поверхности и предметы как единые и цельные по форме, текстуре и окраске. На рис. 6.4.1 А дан простейший пример: мы сходу видим его как цельный круг позади квадрата (рис. 6.4.1 В), даже несмотря на то, что четверть поверхности круга закрыта. И хотя прилагательные «амодальный» и «зрительный», казалось бы, противоречат друг другу, оба относятся к разным аспектам одного и того же явления — феномена перцептивной завершенности. Завершенность вполне можно назвать «амодальной», поскольку во всех сенсорных модальностях она не зависит ни от локальной стимуляции, ни от ощущения. Однако, с тем же успехом ее можно назвать и «зрительной», поскольку полнота восприятия предмета косвенно опирается на зрительную информацию в каком-то из участков ретинального изображения.

В отношении амодальной завершенности мы говорим о некоей феноменологической неопределенности: с одной стороны, мы воспринимаем круг как завершенный (вернее, мы *догадываемся*, что это так), даже несмотря на то, что у нас нет зрительного ощущения от всей его поверхности. По-видимому, достаточно оказаться где-то на пути между полным сенсорным ощущением от видимого объекта и чисто когнитивной осведомленностью о существовании этого предмета, скажем, в соседней комнате. Если закрытая часть контура относительно мала, то ощущение от его полной формы ока-

<sup>1</sup> Любопытно, что полностью закрытые поверхности видимых объектов, например, такие как задняя и нижние грани куба — зрительно завершенны, поскольку остальные части предмета полностью видимы. Еще одна ситуация, когда полностью закрытые объекты оказываются перцептивно завершенными — это ситуация, в которой интересующий нас предмет представляет собой сегмент большего тесно связанного с ним предмета.



**Рис. 6.4.1** Зрительная завершенность. Частично закрытые объекты (например, круг в части А), воспринимаются чаще как завершенные, но при этом расположенные позади кроющего объекта (В). Логически же возможны какие угодно варианты, например, мозаичный (С), либо любой другой (например, D).

жется достаточно близким к реальности (т.е. к тому, как если бы мы его видели полностью), даже невзирая на частичную укрываемость.

Как уже отмечалось в первой главе нашей книги, зрительная завершенность логически недоопределена. Реальность в отношении рис. 6.4.1 А может оказаться такой как показано на рис. 6.4.1 В — квадрат, частично закрывший круг; такой как на рис. 6.4.1 С — квадрат, упершийся в вырез круга; квадрат перед кругом с неправильной формы выпячиванием; бесконечное число прочих вариантов. Создается впечатление, что у зрительной системы имеются четкие установки на обеспечение полноты восприятия частично закрытых объектов, и установки эти очевидно нацелены на достижение веридикальности этого восприятия. Каким же образом?

Феномен амодальной завершенности толкуют разные теории. Мы рассмотрим три из них: теорию *фигуральной типичности*, *фигуральной простоты* и *теорию соотносимости*.

#### 6.4.1.1 Концепция фигуральной типичности

Одна из причин того, что мы воспринимаем темную фигуру на рис. 6.4.1 А цельным кругом, кроется в том, что в прошлом мы в великом множестве видели такие круги. У нас, разумеется, есть опыт наблюдения фигур и в три четверти круга, но обычных кругов в нашей жизни было намного больше. Сие рассуждение подводит нас к идее фигуральной типичности: мы воспринимаем частично закрытые фигуры согласно типичным, повседневно встречающимся формам, согласуемым с визуальной информацией, заключенной в стимуле. Одно из слабых мест в этой концепции таково, что ведь целю



**Рис. 6.4.2** Типичность и зрительная завершенность. Чтобы достичь завершенности, частично закрытые фигуры не обязаны обладать типичной формой (А), но при этом типичность и контекст могут оказать радикальное влияние на результирующий перцепт (В).

мы воспринимаем и неизвестные ранее объекты, причем с тем же успехом и с той же легкостью, что и знакомые. Пример амодально завершенной, но ранее не виденной фигуры находим на рис. 6.4.2 А: визуальных данных в этом стимуле ничуть не больше, чем в примере с кругом с рис. 6.4.1 А.

Сей аргумент вовсе не означает, что типичность не играет роли в возникновении амодальной завершенности восприятия — он лишь указывает на то, что в процесс вовлечено нечто большее, чем простое сходство с уже виденным. Тем не менее, у нас имеется ряд весьма эффектных примеров влияния фигуральной типичности на результирующий перцепт. На рис. 6.4.2 В показана частично закрытая буква среди прочих букв. Любой из нас воспримет ее как литеру «В», хотя логически ее можно истолковать и как «Р», и как латинскую «R», и как любую другую геометрическую фигуру, вообще не похожую на литеру. Тот факт, что «СЛОВО» — это хорошо знакомое слово, а «СЛОРО» или, тем более, «СЛОРО» — отнюдь, представляется весьма важным фактором в данном примере. Однако же схожесть на уровне цельной фигуры вопрос не исчерпывается и мы должны рассмотреть еще несколько вариантов толкования феномена амодальной завершенности.

#### 6.4.1.2 Концепция фигуральной простоты

Вторая причина перцептивной завершенности закрытых фигур заключена в том, что сие результат «наипростейшего» восприятия. К примеру, квадрат, закрывший круг на рис. 6.4.1 А, интуитивно проще, чем любая другая фигура (и сие справедливо в отношении любой неизвестной формы). Толкование феноменов перцептивной организации в логике максимальной простоты (или, эквивалентно — минимальной сложности) — это ведущий теоретический посыл гештальт-психологии (см. раздел 2.1.2). И мы видим, что он блестяще объясняет многие перцептивные феномены.

Из уст гештальтистов гипотеза звучит так, что зрительная система руководствуется *принципом вероятности простейшего* или *принципом прягнанца*,<sup>1</sup> (Hochberg & McAlister, 1953). Принцип гласит, что:

<sup>1</sup> Позднее его стали формулировать *принципом минимализма*.

⇒ Перцепт будет настолько добротным, насколько позволяют превалирующие условия.

Термин «добротность» относится к степени фигуральной простоты (или регулярности), а превалирующие условия — к структуре текущего стимулирующего ретинального изображения. Проблема же заключена в том, что для прогнозирования перцептов нам требуется четкое определение «добротности». И сие сложность, с которой гештальтисты никогда не сталкивались в полной мере. Мы детально обсудим эту проблему в 8-й главе, когда вплотную займемся разбором вопросов фигуральной добротности. Сейчас лишь проиллюстрируем основную идею.

Итак, предположим, что добротность фигуры поверяется лишь числом осей билатеральной симметрии: чем больше осей, тем выше фигуральная добротность.<sup>1</sup> В данном случае перцепт квадрата, закрывающего круг — это простейшая версия, поскольку круг билатерально симметричен по всем осям, проходящим через его центр. Мозаичная интерпретация имеет только одну ось симметрии (примерно по левой диагонали, проходящей через центр), то есть мы видим, что нерегулярно завершенный круг в целом несимметричен. То есть, перцепт квадрата перед кругом — это простейшая интерпретация. Как будет показано в гл. 8 — сказанное является сильным упрощением, но оно иллюстрирует варианты аналитического подхода к амодальной завершенности в рамках гештальт-идеологии и, в частности, критерия прягнанц.

Одна из проблем, связанных со всеми толкованиями амодальной завершенности состоит в том, что, хотя толкования эти и учитывают множество фактов, существуют прочие факты, по которым не получается дать корректный прогноз. Один контрпример дан на рис. 6.4.3 А: обычно квадрат воспринимается расположенным перед другим квадратом, у которого отсечен правый нижний угол (рис. 6.3.4 В). Однако согласно определению, построенному на симметрии, простейший вариант завершенности — это шестигранник, показанный на рис. 6.4.3 С, поскольку имеет две оси симметрии (обе диагональные), тогда как квадрат без угла — только одну (левая диагональ).

Однако сей пример отнюдь не приговор идее симметрии, он лишь указывает на недочет в определении фигуральной добротности. К примеру, критично важным фактором фигуральной добротности может стать количество сторон. Действительно, определяя добротность исключительно в показателях числа сторон получаем вполне корректный прогноз: предпочитаемый квадрат без угла имеет лишь пять сторон, тогда как шестигранник — шесть. Разумеется, существует множество иных факторов, совокупно определяющих фигуральную добротность.

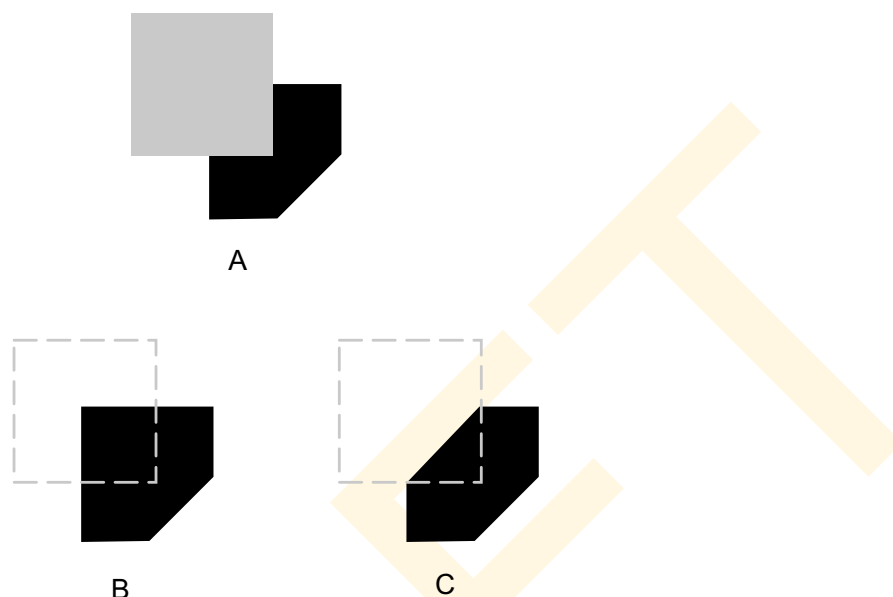
Итак, приведенный выше пример, иллюстрирует общий подход к теориям амодальной завершенности, основанным на прягнанце: они тем добротнее, чем проще.

#### 6.4.1.3 Теория соотносимости

Теории третьей группы толкуют феномен зрительной завершенности, напрямую апеллируя к экологической структуре закрытых контуров. Например, если контур одного объекта частично закрыт другим объектом, образуется вертекс Т-типа. Непрерывный контур (верхняя планка «Т») интерпретируется как кроющий край, чья поверхность закрывает другой край (ножка «Т»). Затем приходится предположить, что

<sup>1</sup> На самом деле всё гораздо сложнее, но для иллюстрации мы упрощаем ситуацию.





**Рис. 6.4.3** Простота в концепции завершенности. Закрытая черная фигура (А) скорее воспримется как имеющая форму, показанную в В, нежели в С.

закрытый край (и принадлежащая ему поверхность) как-то связана с другим закрытым краем в сцене.

Одну из экологических концепций завершенности предложили Филипп Кельман и Томас Шипли (Пенсильванский университет, 1991), назвав ее *теорией соотносимости*. Сразу отметим, что теорию соотносимости можно понимать как некое расширение к классическому принципу группировки по цельности.

Идея соотносимости зиждется на четырех факторах, первый из которых т.н. *прерывность*, и мы говорим, что прерывность — это необходимое, но недостаточное условие зрительной интерполяции амодальных контуров. Итак:

⇒ Прерывность контура — это резкая смена направления этого контура.<sup>1</sup>

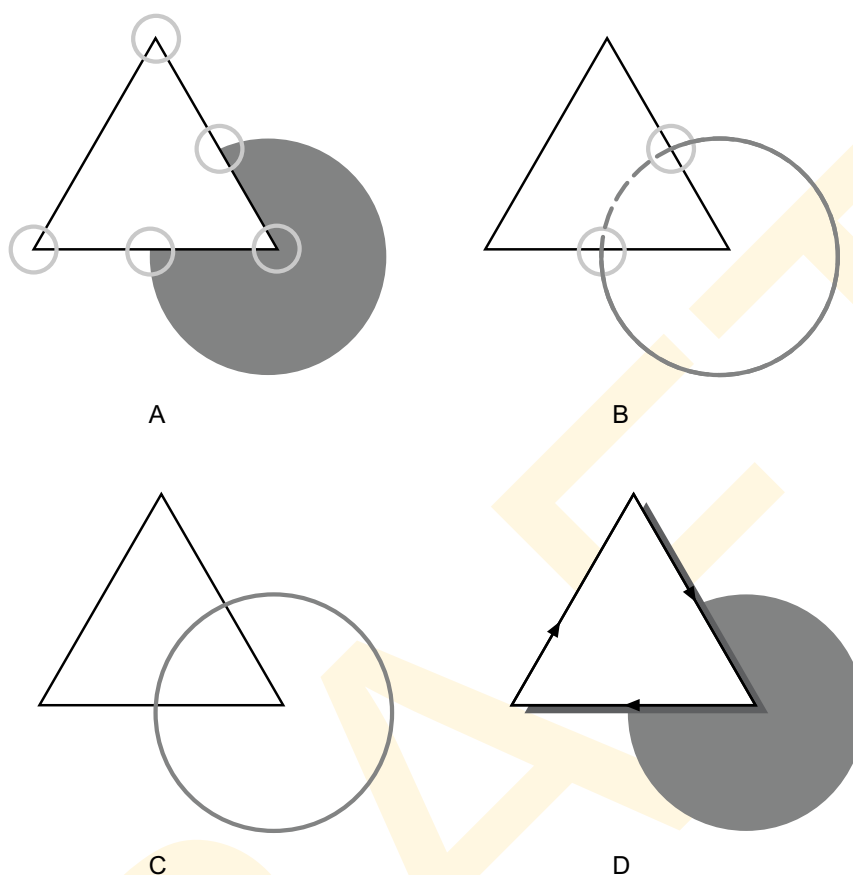
Так прерывность краев появляется на всех пересечениях контуров, включая незавершенные (например, такие как углы квадрата). Чтобы глубже понять смысл прерывности, рассмотрим рис. 6.4.4 А. Прерывности сами по себе недостаточны для генерации амодальной завершенности, поскольку она зависима от того, можно или нельзя одни прерывности соотнести с другими прерывностями в ретинальном изображении (что и следует из названия теории).

Второй фактор:

⇒ Амодально завершенными контуры воспринимаются тогда, когда прерывные края соотносимы (см. рис. 6.4.4 В).

Согласно Кельману — Шипли два края соотносимы тогда и только тогда, когда: их

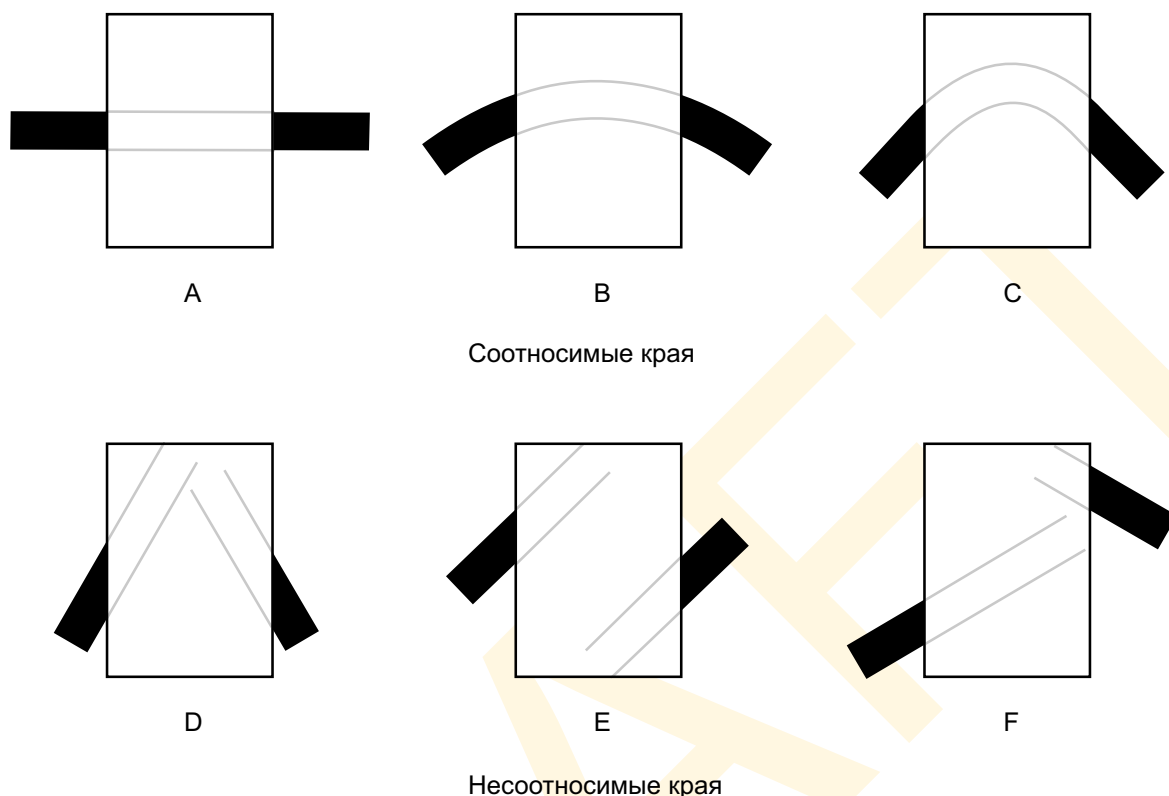
<sup>1</sup> Логически, прерывность края определена как прерывность первой производной математической функции, описывающей край в пространстве.



**Рис. 6.4.4** Этапы зрительного завершения согласно концепции соотносимости Кельмана — Шипли. На первом этапе идентифицируются прерывности (А), затем завершаются соотносимые пары (В). После чего закрытые фигуры, определяемые завершенными краями, формируют новые перцептивные единицы (С) и, наконец, определяются соотношения по глубине (D).

*продолжения* пересекаются под углом не менее  $90^\circ$  и когда их можно плавно соединить друг с другом (рис. 6.4.4 В). Чтобы уяснить смысл сей концепции, посмотрим на рис. 6.4.5, в котором даны три варианта соотносимых краев и три варианта несоотносимых. Когда края коллинеарны (часть А) — то есть, когда их продолжения сходятся под углом  $180^\circ$  — соотносимость абсолютная. Она несколько снижается, когда угол между продолжениями начинает уменьшаться (В и С) и отвесно падает до нуля (или почти до нуля) при угле ниже  $90^\circ$  (D). Края также несоотносимы, когда их *продолжения* вообще не пересекаются, либо, когда края параллельны друг другу и разнесены (Е), либо же, когда продолжение одного пересекает другой край по собственно второму краю, а не по его продолжению (F).

Отметим, что сие определение позиционирует соотносимость тех или иных пар краев, но не описывает при этом того, как зрительная система определяет, какой именно край с каким соотносится. А сие крайне важно, потому что край может соотноситься с многими другими, но только один из них воспримется как его продолжение. Предположительно зрительная система пытается найти оптимальные соответствия путем поиска наистрожайших соотносимостей.



**Рис. 6.4.5** Условия соотносимости краев. Поскольку края в А, В и С соотносимы, фигуры, частями которых они являются, завершены позади кроющего четырехугольника. Края в D, E и F несоотносимы и поэтому материнские фигуры не завершены (подробнее см. в тексте).

По материалам Kellman & Shipley, 1991.

Следующий критерий:

⇒ Новая перцептивная единица формируется тогда, когда амодально завершенные края образуют замкнутую область.

Сей тезис попросту означает, что прерывные края, соотносясь друг с другом, образуют множество, формирующее закрытый контур (см. рис. 6.4.4 С). Обычно новая перцептивная единица согласуется с объектом сцены, но им может стать и отверстие (чего мы уже коротко коснулись, когда рассматривали отношения между мнимыми контурами и амодальной завершенностью).

Наконец:

⇒ Перцептивным единицам присваивается определенное удаление от наблюдателя на основе имеющихся факторов глубины пространства (рис. 6.4.4 D).

К примеру, информация, извлекаемая из окклюзии (интерпозиции) описывает то, что амодально завершенные края расположены позади кроющего объекта. Сие проиллюстрировано рис. 6.4.4D, где стрелки и правило правой руки (см. раздел 5.5.7) указывают на то, какая поверхность перед какой расположена.

Может показаться странным то, что такой сложный процесс требуется для воспри-

ятия такой простой вещи, как треугольник перед кругом (рис. 6.4.4). Однако эксперименты показали, что процесс завершения простых геометрических форм, например квадратов и кругов, требует порядка 400 мс (Sekuler & Palmer, 1992). Однако же соответствуют процессы, проходящие в этот период, положениям теории соотносимости или нет — по сей день остается загадкой.

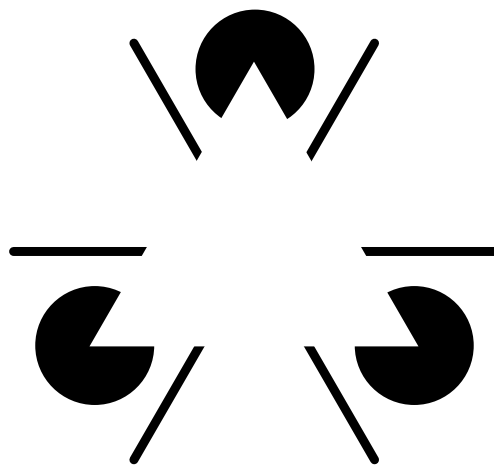
### 6.4.2 Мнимые контуры

Следующий вид зрительной интерполяции дает забавную иллюзию, при которой видимые контуры реально отсутствуют в стимуляции. Феномен *мнимых контуров* (или, как их еще называют — *субъективных контуров*) известен довольно давно (Schumann, 1904), но всплеск интереса к нему случился благодаря эффектным демонстрациям итальянского психолога Джетано Каниша (1955, 1979). Один из наиболее популярных примеров — *треугольник Каниша* (рис. 6.4.6): белый иллюзорный треугольник легко воспринимается как заданная мнимым контуром, т.к. собственно стимул содержит лишь три черных круга с вырезами и три отрезка линий. Большинство наблюдателей сообщают, что видят вполне определенные яркостные края как раз в тех местах, где и должны быть контуры треугольника. Любопытно, что внутренняя часть фигуры воспринимается при этом чуть светлее окружения.

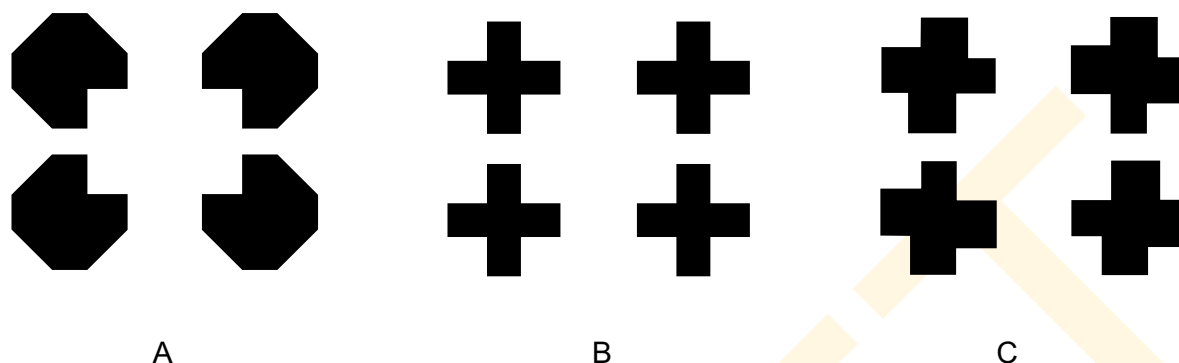
Появлению мнимых контуров содействует амодальная завершенность наводящих элементов изображения. К примеру, когда мы зрим иллюзию треугольника на рис. 6.4.6, круги и линии мы воспринимаем расположенными как бы позади этого треугольника, если наводящие элементы воспринимаются незавершенными, то мнимых контуров и фигур не появляется: на рис. 6.4.7 А иллюзорный четырехугольник виден благодаря тому, что наводящие элементы воспринимаются завершенными восьмигранниками; на рис. 6.4.7 В, однако, иллюзии четырехугольника не возникает, даже несмотря на то, что угловые вырезы точно такие же, как и на рис. 6.4.7 А.

Особенность рис. 6.4.7 В в том, что вырезанные углы воспринимаются неотъемлемыми фрагментами перекрестий. Следовательно, внутренние выемки не требуют перцептивного «объяснения» через кроющие объекты. Один только фактор может частично относиться к вопросу — симметрия: чтобы демонстрация воспринялась как частично закрытая белым прямоугольником поверхность, симметрия перекрестий должна толковаться зрительной системой как случайно возникшая из специфического расположения кроющего белого прямоугольника и наводящих элементов. В этом случае симметрия не станет блокировать видность мнимых контуров.

На рис. 6.4.7 С показано, что ассиметричные крестоподобные фигуры тоже не в силах дать иллюзию контуров до тех пор, пока их стороны остаются параллельными. И вновь объяснение прячется в крайне малой вероят-



**Рис. 6.4.6** Мнимые контуры. Любой из нас легко увидит белый треугольник на фоне частично «закрытых» им кругов и линий. Любопытно, что внутренность треугольника воспринимается чуть светлее, чем окружение (даже несмотря на то, что фотометрическая яркость треугольника та же, что и у окружения).



**Рис. 6.4.7** Условия восприятия мнимых фигур. Белый прямоугольник устойчиво и легко читается на фоне четырех черных восьмигранников (А), но почти не виден на фоне других фигур (В и С), несмотря на присутствие тех же локальных контуров: судя по всему, факторы симметрии и параллельности краев блокируют восприятие.

ности того, чтобы стороны кроющего четырехугольника оказались непременно параллельны сторонам наводящих фигур (Albert, 1993).

Итак, мы убедились в том, что мнимые фигуры и контуры — это относительно сложный феномен, зависящий от многих локальных условий стимуляции. И сие отличный пример того гештальт-тезиса, что перцепт во многом определяется цельной структурой и что его порой невозможно спрогнозировать по структуре отдельных частей.

Рис. 6.4.8 А продемонстрирует нам то, что мнимые фигуры, подобно амодально завершенным, хоть и не обладают типичностью формы как треугольники или круги, но при этом воспринимаются с той же легкостью (Gregory, 1972). Прочие весьма эффектные примеры мнимых контуров дают сегменты линий, которые оканчиваются так, как будто покрыты белым непрозрачным кругом — рис. 6.4.8 В. Мнимые контуры могут восприниматься даже тогда, когда сегменты линий смещены — либо внутрь, либо наружу от иллюзорной границы, образованной их концами (рис. 6.4.8 С).

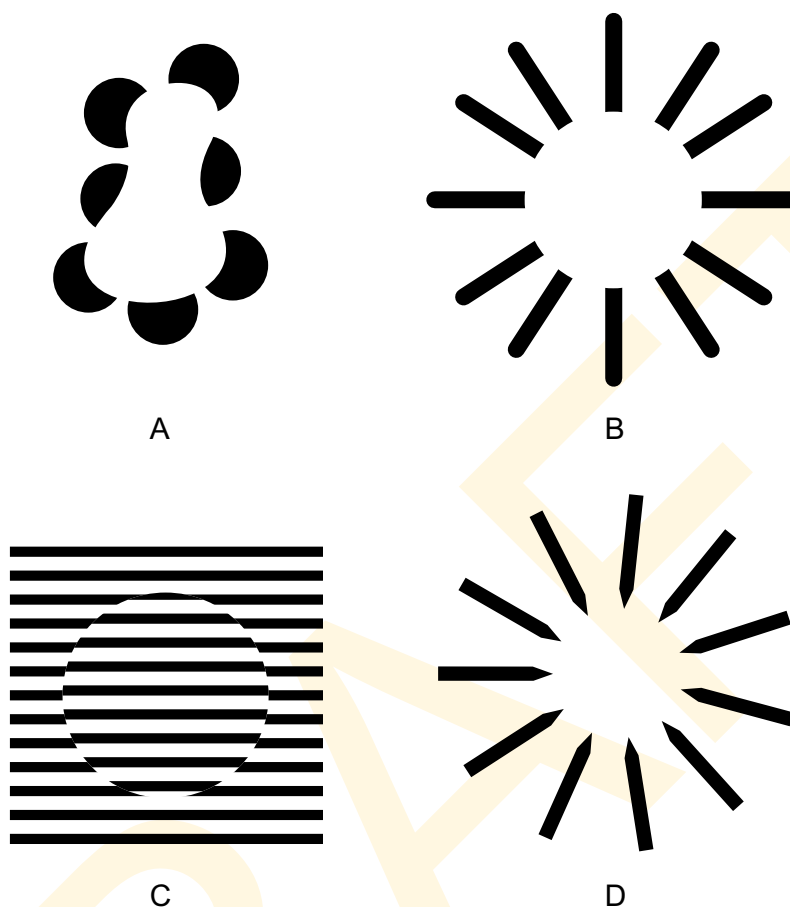
Канадский психолог Джон Кеннеди (1988) показал, что мнимые контуры, индуцированные линиями, целиком зависимы от характера оконцовок этих линий: мы видим, что на рис. 6.4.8 D мнимый контур ослаблен или даже вовсе отсутствует, поскольку оконцовки линий заострены. Аналогичное ослабление иллюзии наблюдаем при закруглении концов. Сии факты тесно связаны с неким общим положением, что чем резче обрыв линий вдоль «кроющего» контура, тем сильнее перцепт окклюзии (сравним рис. 6.4.8 В с рис. 6.4.8 D).

#### 6.4.2.1 Связь с амодальной завершенностью

Мы уже говорили об одном из возможных путей, которым субъективные контуры и мнимые фигуры соотносятся с амодальной завершенностью: наводящие элементы амодально завершаются позади иллюзорной фигуры. Однако Кельман убедительно показал, что связь между двумя феноменами намного глубже, чем кажется на первый взгляд (Kellman & Loukides, 1987; Kellman & Shipley, 1991): ученый продемонстрировал, что оба результата происходят из одного и того же множества подлежащих процессов и оба описываются в логике теории соотносимости.

Кельман заметил, что вместо мнимых контуров многие иллюзорные фигуры ино-





**Рис. 6.4.8** Еще один пример мнимых фигур. Мнимые фигуры могут быть ранее неизвестными (А), а также наведенными исключительно оконцовками линий (В и С). При этом концы линий должны формировать гладкий контур: заостренные концы не дают эффекта (D).

гда дают альтернативные перцепты: к примеру, треугольник Каниша на рис. 6.4.9 А, можно увидеть (хоть и с определенным усилием) как белую фигуру на равномерно черном фоне, видимую через три круглых отверстия в кроющей белой поверхности. Рис. 6.4.9 В даст тот же перцепт, но много легче, поскольку в нем отрисованы дуги отверстий. Ключевым здесь является то, что мнимых контуров треугольника нет — скорее стороны треугольника амодально завершены позади укрытия.

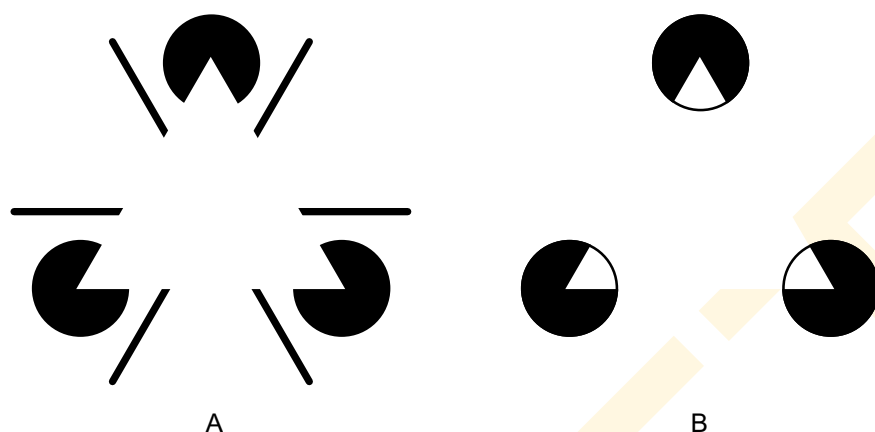
Кельман и Шипли (1991) выявили еще одну важную взаимосвязь между амодальной завершенностью и мнимыми контурами: многие конфигурации иллюзорных контуров можно конвертировать в амодально завершенные перцепты. Простейшая трехступенчатая процедура такой конверсии проиллюстрирована рис. 6.4.10:

1. Начнем с мнимой фигуры, показанной на рис. 6.4.10 А.

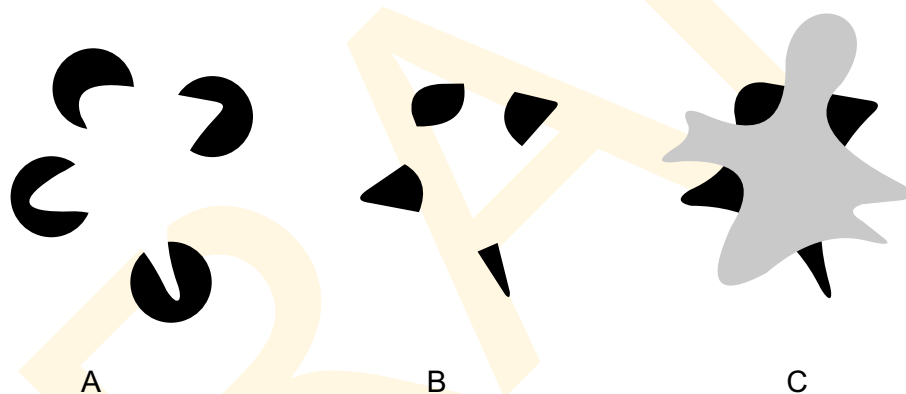
2. Выделим наводящие элементы — рис. 6.4.10 В.

3. Добавим кроющую фигуру, чьи контуры мягко повторяют внутренние контуры нашей добавки (серая фигура на рис. 6.4.10С).

Результатом явится пример амодально завершенной фигуры, в которой контуры — ранее мнимые — теперь воспринимаются как закрытые новой фигурой.

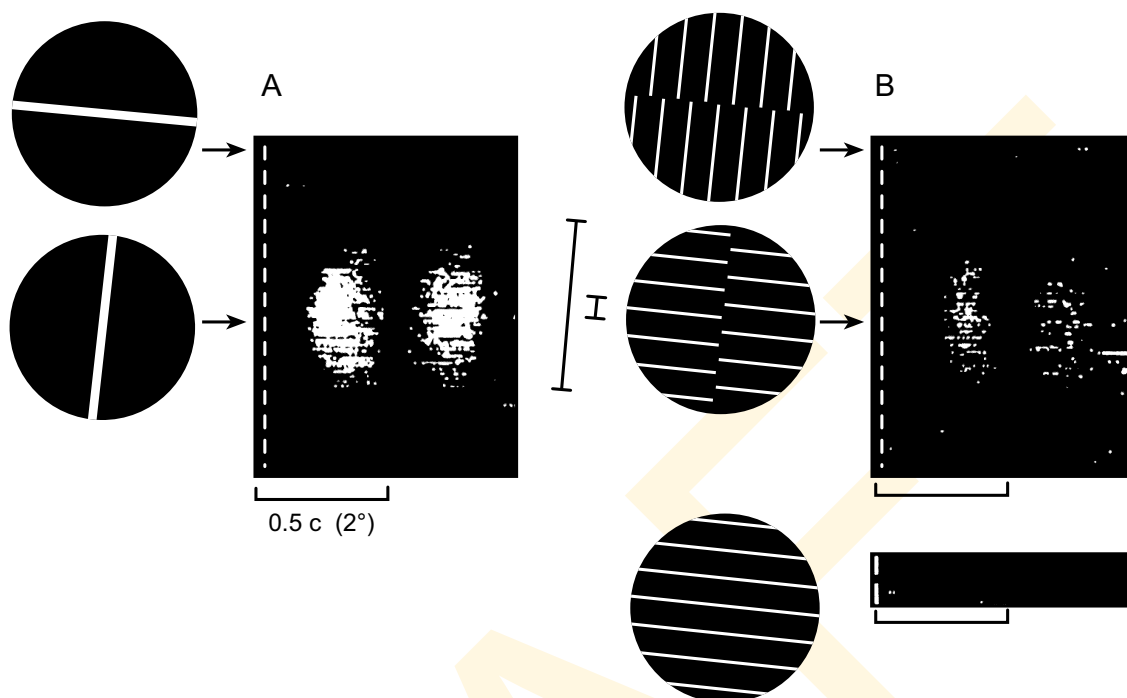


**Рис. 6.4.9** Мнимые контуры, амодальная завершенность и взаимоотношения по глубине. Мнимые контуры (А) пропадают, когда треугольник оказывается виден как бы через отверстия (В) — в данном случае стороны треугольника воспринимаются амодально завершенными. По материалам Kellman & Shipley, 1991.



**Рис. 6.4.10** Конверсия мнимой фигуры в амодально завершенную. Начав с иллюзорной фигуры (А), выделим наводящие элементы (В) и добавим кроющую фигуру, контуры которой плавно повторяют внутренние контуры наводящих элементов (С).

Столь тесная связь между двумя феноменами натолкнула Кельмана на мысль, что, фактически, мнимые контуры и амодальная завершенность — это разные проявления одного и того же процесса, процесса, лежащего в основе формирования перцептивной единицы (что и предполагает теория соотносимости). Но если в обоих случаях действует один и тот же процесс, то почему оба феномена выглядят столь разными? Ответ находим в четвертом тезисе, касающемся связи с глубиной пространства, а именно: результат восприятия оказывается зависим от перцептивных отношений по глубине между фигурами — если отсутствующие контуры, это часть ближней кроющей фигуры — перцепт мнимых контуров возникнет; ежели отсутствующие контуры, это фрагмент дальней — закрытой фигуры, то они окажутся амодально завершенными позади фигуры кроющей.



**Рис. 6.4.11** Отклики нейронов V2 на стимуляцию реальными и мнимыми контурами. Нейроны V2 откликаются на фактические движущиеся линии околоректальной ориентации, на что указывает высокая плотность белых точек спайковой записи (А). Эти же клетки селективно чувствительны к мнимым контурам той же ориентации, но не к линиям, наводящим эти контуры (В).

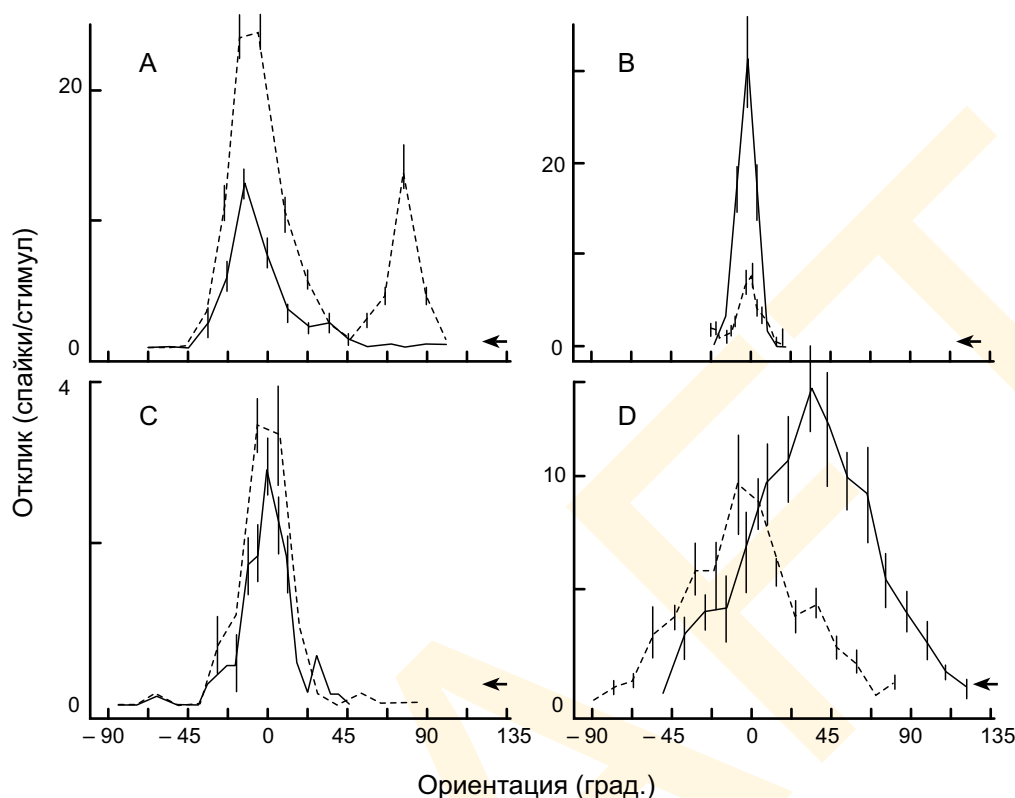
По материалам von der Heydt & Peterhans, 1989 а.

#### 6.4.2.2 Физиология мнимых контуров

Сравнительно недавно (благодаря успехам физиологии коры) удалось идентифицировать клетки V2, дающие отклик на наличие мнимых контуров в стимуляции.<sup>1</sup> Развернутый тест, предложенный американским нейрофизиологом Рудигером фон дер Хейдтом, показал, что 40% клеток V2, селективных по ориентации, возбуждаются, когда глазу предъявлен стимул, индуцирующий появление мнимых контуров (Peterhans & von der Heydt, 1989, 1991; von der Heydt, Peterhans, & Baumgartner, 1984; von der Heydt & Peterhans, 1989).

На рис. 6.4.11 даны примеры стимулов, использованных в эксперименте Хейдта. В начале регистрировалась частота пульсации каждой клетки в ответ на стимуляцию сетчатки белой полоской в шестнадцати ориентациях (рис. 6.4.11 А). Коль скоро клетки селективны по ориентации (что отмечено острыми пиками в сплошных кривых на рис. 6.4.12), то ожидался и, в итоге, был зафиксирован их отклик и на стимуляцию, потенциально несущую мнимый контур — рис. 6.4.11 В. Было обнаружено, что очень многие нейроны V2 в той или иной мере реагируют и на мнимые контуры в стимуляции: рис. 6.4.12 С демонстрирует нам кривые отклика клеток, реакция которых на реальные и мнимые контуры почти что одна и та же; на рис. 6.4.12 А и 6.4.12 В даны слу-

<sup>1</sup> Рецептивные поля клеток V2 изначально похожи на таковые у клеток V1 (см. раздел 4.1.2).

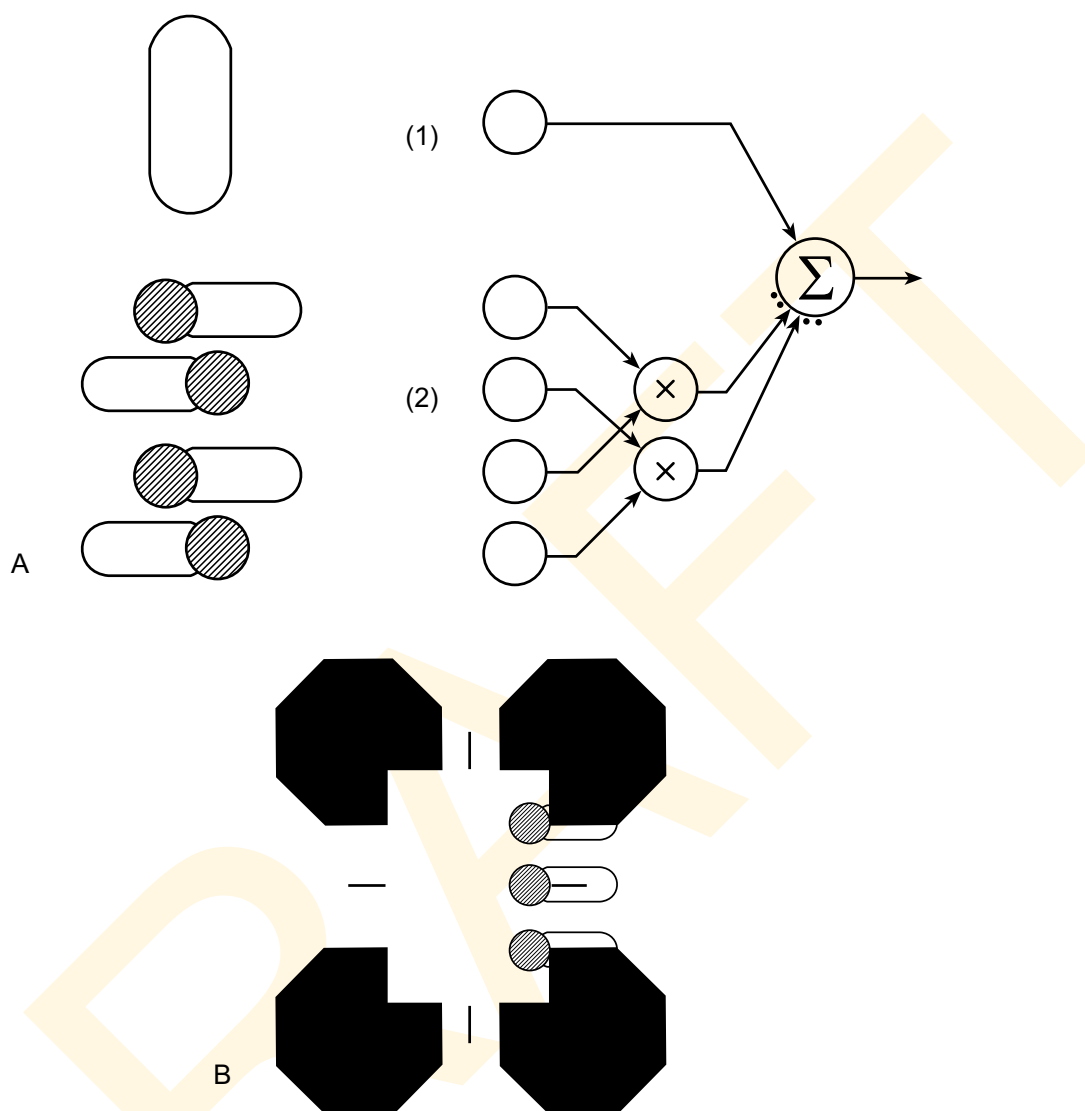


**Рис. 6.4.12** Функции отклика нейронов V2 к реальным и мнимым контурам различных ориентаций. Отклики четырех контурочувствительных клеток V2 на светлые штрихи (сплошные линии) и мнимые контуры (пунктиры) весьма сходны у одних клеток и сильно разнятся у других. По материалам von der Heydt & Peterhans, 1989 а.

чаи, в которых клетки, настроенные на стимул одной ориентации, обладают разной чувствительностью к реальным и к мнимым контурам; на рис. 6.4.12 D дан пример, в котором оптимальная ориентация по реальным контурам отлична от таковой по контурам мнимым на  $45^\circ$ . Отметим, что клетки, показанные на рис. 6.4.12 A откликаются на ориентацию реальных линий в иллюзорогенном стимуле, но другие — нет.

Каким же образом возникает отклик на мнимые контуры? Существует множество объяснений, одно из которых называют *гипотезой контурных клеток*. Согласно ей выходной сигнал множества клеток V2, чувствительных к ориентации стимула — интегрируется (рис. 6.4.13). Входной сигнал эти клетки получают в том числе от простых и сложных клеток V1, приоритетная ориентация стимуляции для которых та же, что и у клеток V2. Вторым вариантом входного сигнала — это сигнал от множества ортогонально ориентированных стопорных клеток V1, чьи *ингибирующие концы* расположены вдоль приоритетной ориентации контура. Такая клетка будет демонстрировать функцию настройки аналогичную той, что показана на рис. 6.4.12 A с откликами как на реальные так и на мнимые контуры, и в ориентациях, отличных от  $90^\circ$ . На рис. 6.4.12 B показано то, как приблизительно должны располагаться стопорные рецептивные поля, чтобы реагировать на несуществующие контуры мнимого прямоугольника.

Отметим, что гипотеза контурных клеток не дает исчерпывающего объяснения происхождению мнимых контуров, но лишь предлагает удобное толкование тому, как от-



**Рис. 6.4.13** Схема работы контурных клеток V2. Контурные клетки ( $\Sigma$ ) суммируют сигнал по двум параллельным проводящим путям: (1) по пути, идущему от клеток — краевых детекторов в V1 и (2) группирующему пути, идущему от перпендикулярных стопорных клеток.

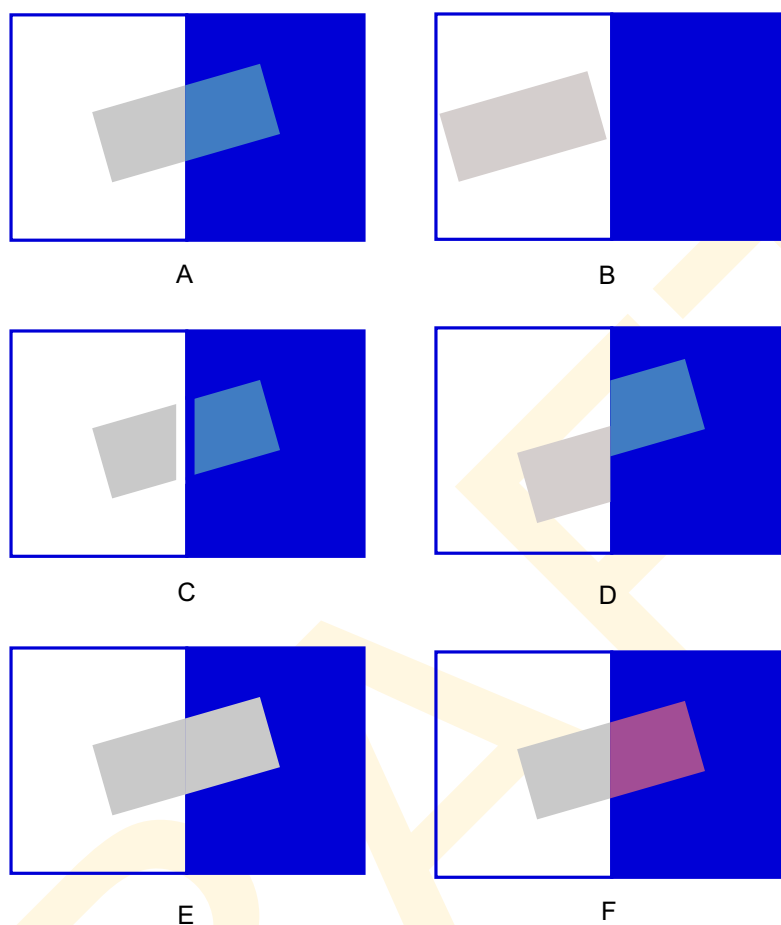
По материалам von der Heydt & Peterhans, 1989 а.

клик этих клеток получается из выходного сигнала нейронов V1. В частности, гипотеза эта не дает объяснения того, почему иллюзорные фигуры воспринимаются расположенными перед наводящими элементами и не в силах объяснить, почему рис. 6.4.7 А дает мнимые контуры, а рис. 6.4.7 В — нет. Для толкования столь тонких, но необычайно важных феноменов, нужна намного более развитая научная концепция.

### 6.4.3 Восприятие прозрачности

Еще одной чертой перцептивной организации, плотно соотносимой с феноменом амодальной завершенности и феноменом мнимых контуров, является восприятие





**Рис. 6.4.14** Условия восприятия прозрачности. Наблюдатель воспримет прозрачность, когда во взаимоотношениях как минимум четырех участков задействована одна полупрозрачная поверхность, расположенная перед двумя (и более) соседствующими поверхностями (А). Если каким-то образом это условие нарушено (В, С, D, Е и F) — перцепта прозрачности не возникает.

*прозрачности* — перцепция объекта, видимого через ближележащий полупрозрачный объект.

⇒ Полупрозрачный объект — это объект, в большей мере пропускающий свет, отраженный от позади лежащих объектов, нежели поглощающий или отражающий его назад.

В условиях полупрозрачности свет, воздействующий на сетчатку (в данной ретинальной локали), переносит информацию *как минимум* о двух внешних точках одного направления (от точки наблюдения): о точке на собственно полупрозрачной поверхности и о точке на непрозрачной поверхности, видимой через первую (рис. 6.4.14 А). Мы сказали «как минимум две точки», поскольку на пути светового потока может оказаться несколько полупрозрачных поверхностей. Однако для простоты мы будем рассматривать лишь один полупрозрачный слой.

Восприятие прозрачности зависит от определенного сочетания пространственной и яркостно-хроматической стимуляции. Изменение хотя бы в одном из этих условий ве-

дет к блокировке эффекта. К примеру, перцепт прозрачности возникнет, если полупрозрачная поверхность позиционирована так, что вещественные края на непрозрачной поверхности, расположенной позади первой, видны как сквозь нее, так и простираются за нее — рис. 6.4.14 А. Когда сие имеет место, то возникает феномен, именуемый *цветоделением* — цвет того участка ретинального изображения, что является проекцией от полупрозрачной поверхности, воспринимается «комбинацией» двух цветовых ощущений — одного, «принадлежащего» фону, другого — светопропускающей поверхности.

Эффект цветоделения не возникнет, если нарушены пространственные, яркостные, либо хроматические соотношения в стимуляции. Ежели светопропускающая поверхность заключена полностью внутри одного отражающего участка (рис. 6.4.14 В), стимул воспримется как некая единая непрозрачная поверхность с участком иной отражательной способности, но без признаков прозрачности. Перцепт прозрачности также может исчезнуть (либо сильно ослабнуть), если разрушить цельность единицы пропускающего участка (рис. 6.4.14 С и 6.4.14 D). Таким образом, цветоделение — это эффектный пример значения пространственной организации в восприятии прозрачности — эффект не возникает без должной перцептивной организации участков (Metelli, 1974). При этом исполнения одних лишь пространственных условий недостаточно, поскольку отклонение по другим параметрам стимуляции заблокирует восприятие прозрачности строго в той же пространственной конфигурации, что дает перцептивную прозрачность при подходящих яркостно-хроматических условиях (см. рис. 6.4.14 Е и 6.14.4А).

Теперь мы коротко рассмотрим яркостно-хроматические условия восприятия прозрачности. В простейшей двухслойной демонстрации (рис. 6.4.15) даны четыре пространственных участка: у двух из них (А и В) разные коэффициенты отражения; у двух других (С и D) — один общий, но свет от последних — это свет, прошедший сквозь *пропускающую* поверхность. Суть цветоделения, возникающего в зрительной системе, в том, что в каждом из двух участков ретинальной проекции этой фигуры (С и D) заключен лишь один вариант яркостно-хроматической стимуляции, но при этом по каждому из участков зрение извлекает информацию о двух поверхностях: о фоне и о пропускающей поверхности поверх него.

Сие еще один пример недоопределенности задачи инверсии, поскольку в наличии больше неизвестных (светлота пропускающего слоя и обоих непрозрачных участков), чем известных (фотометрическая яркость в каждом из участков наложения). Понятно, что дополнительная информация, необходимая для решения этой задачи, поступает из эвристических допущений, сделанных в отношении хроматических взаимоотношений с окружающими пространственными контекстами.

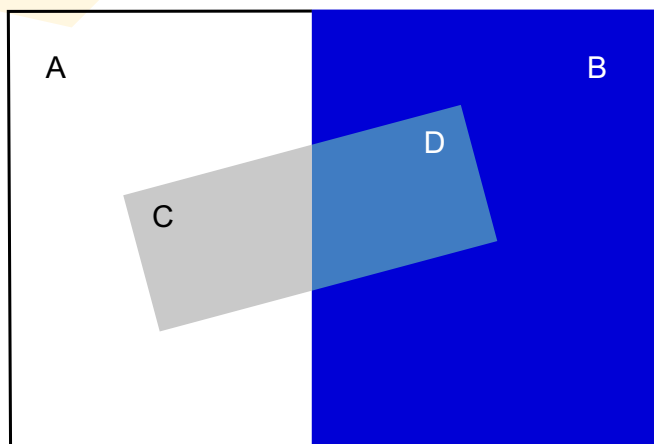


Рис. 6.4.15 Яркостно-хроматические отношения в восприятии прозрачности. Фотометрическая яркость и цветность участков наложения (С и D) должны лежать где-то между таковыми у открытых участков (А и В) и у самостоятельной полупрозрачной поверхности.

Критичность контекста можно продемонстрировать, откадрировав участок наложения (например, с помощью картонки с вырезом) — тогда сразу исчезнут и цветоделение, и прозрачность. Следовательно, обрезка удаляет из стимула критично важную информацию, необходимую для возникновения эффекта цветоделения.

В качественном смысле мы можем сказать, что цветовой перцепт по каждому из участков наложения «делится» на компонент прозрачного и компонент непрозрачного слоев. Метелли (1974) и Джербино (1994) дали количественный анализ условий такого деления, но сейчас мы скажем лишь о том, что прозрачность требует того, чтобы светлота участка наложения была промежуточной между светлотой непрозрачного слоя и светлотой пропускающего. К примеру, если непрозрачный слой черный, а пропускающий — средне-серый, совокупный эффект по участку наложения будет темно-серым. Если непрозрачный слой белый, а пропускающий — средне-серый, то участок наложения будет светло-серым.

#### 6.4.4 Фигуроделение

Наконец, имеется еще один феномен, рассматриваемый нами как пример зрительной интерполяции — *фигуроделение* (которое мы называли так по аналогии с цветоделением). При фигуроделении единый гомогенный участок перцептивно делится на две фигуры одного и того же цвета — одна фигура располагается перед другой, закрывая ее — рис. 6.4.16.

У феномена фигуроделения имеется ряд весьма любопытных свойств:

**1. Предопределенность.** Для деления равномерно связанного участка не требуется никакой локальной сенсорной информации. То есть, внутри участка нет ни яркостных, ни хроматических краев, которые задали бы две фигуры.

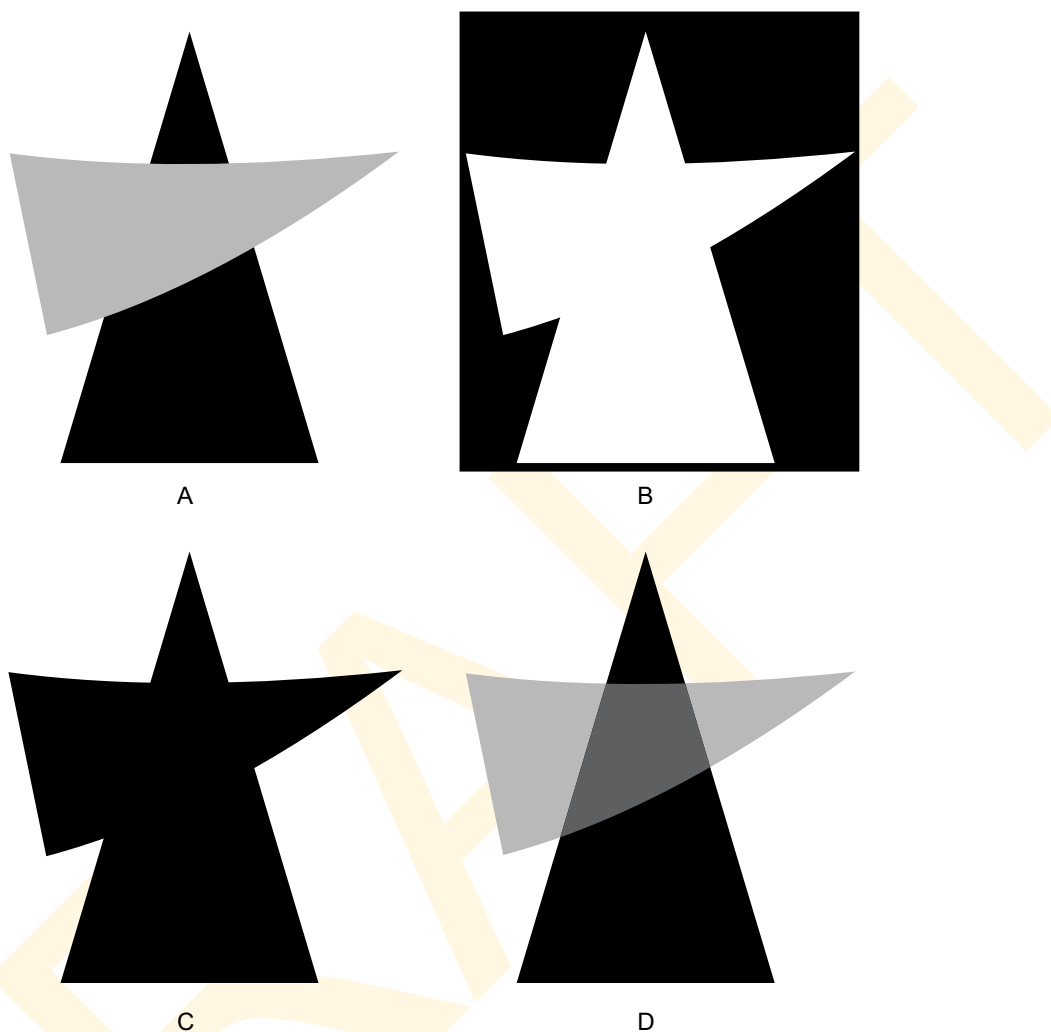
**2. Мнимые контуры.** Зрительная система достраивает отсутствующие контуры там, где ожидает их найти. Таким образом, когда начинается фигуроделение, мнимые контуры появляются во внутренней части участка, там где ближняя фигура закрывает дальнюю.

**3. Завершенность.** Коль скоро мнимые контуры фигуры сформировались, зрительная система завершает, достраивает участки фигуры, закрытые ближней фигурой.

**4. Неопределенность.** Условия стимуляции не определяют, какая фигура окажется перед, а какая позади — равновероятны оба варианта.

**5. Мультистабильность.** Если мы на некоторое время сосредоточим взгляд на такой демонстрации, то взаимоотношения по глубине двух частей начнут спонтанно обращаться. Когда сие происходит, мнимые контуры начинают динамически перемежаться — то мы видим их у одной фигуры, то у другой.

**Рис. 6.4.16** Фигуроделение. Цельный гомогенный участок перцептивно делится на два «накладывающихся» объекта, один из которых частично «закрит» другим. Отношения по глубине при этом остаются неопределенными.



**Рис. 6.4.17** Пример отношений между различными видами организационных феноменов. Две одинаковых фигуры использованы для демонстрации зрительной завершенности (А), мнимых контуров (В), фигурodelения (С) и перцептивной прозрачности (D).

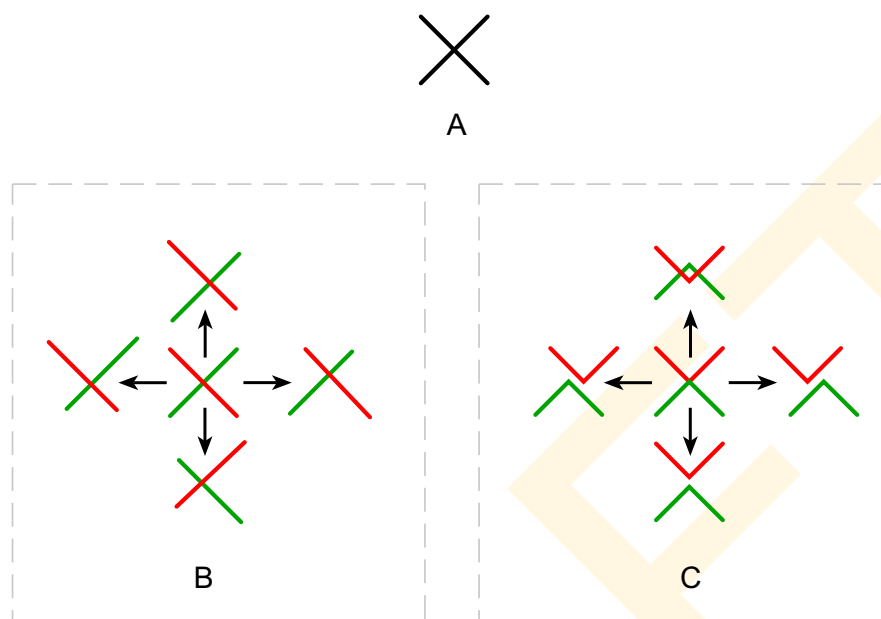
По материалам Kellman & Shipley, 1991.

Тесная взаимосвязь четырех феноменов перцептивной организации — амодальная завершенность, мнимые контуры, перцептивная прозрачность и фигурodelение — проиллюстрированы рис. 6.4.17.

Наконец отметим, что на самом деле феноменов, которые подпадают под кельманову гипотезу зрительной интерполяции несколько больше, но поскольку они имеют отношение к движению, мы обсудим их в 10-й главе, когда будем говорить о динамических событиях в сцене.

### 6.4.5 Принцип неонакцидентальности

Многие феномены перцептивной организации толкуемы в рамках концепта т.н. *нонакцидентальности (предрешенности)*: гипотезы о том, что зрительная система избе-



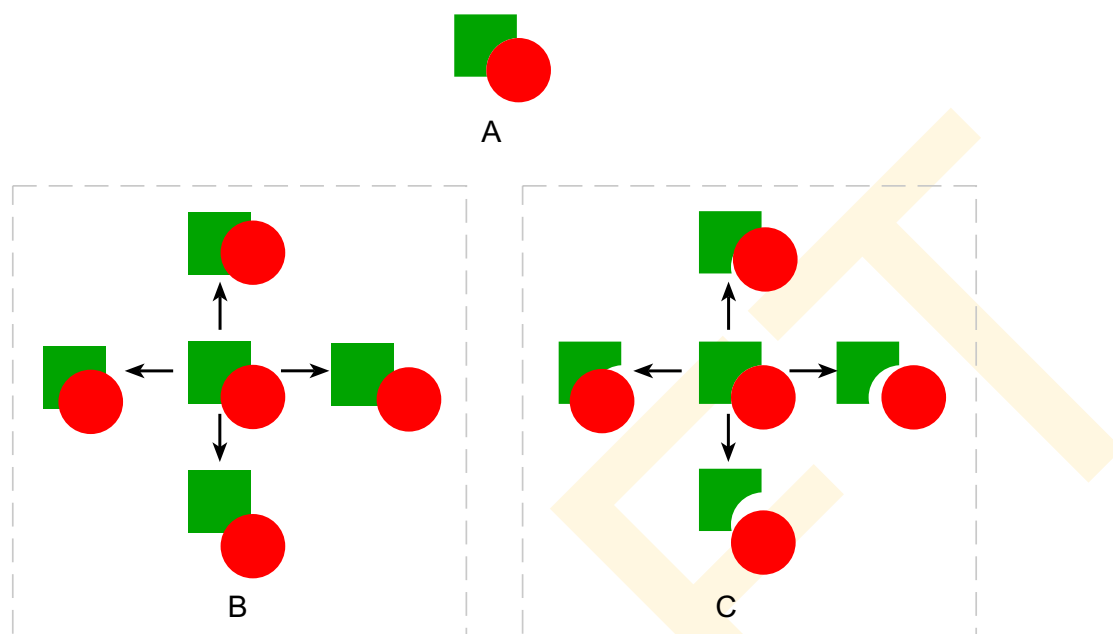
**Рис. 6.4.18** Нонакцидентальность и цельность. Легкие изменения в положении элементов (А) приводят к меньшим изменениям в геометрии изображения, если оно интерпретируется как состоящее из двух неразрывных линий на разных удалениях друг от друга (В). Те же изменения в положении приводят к гораздо большему изменению, если изображение толкуется как состоящее из двух углов, случайно сошедшихся своими вершинами (С).

гает интерпретации структурных закономерностей как возникших из неправдоподобных зрительных случайностей. В психологии первым о нонакцидентальности заговорил Ирвин Рок (1983) как о принципе *ухода от случайностей*. Он показал, что многим феноменом перцептивной организации можно дать объяснение, предположив, что зрительная система избегает интерпретаций, при которых те или иные свойства ретинального изображения — это плод какой-либо случайности какого-то вида. Независимо от Рока с аналогичными идеями, но в количественном аспекте, выступил Дэвид Лоу (Бостонский университет, 1985), а его исходные посылы были подхвачены затем прочими исследователями (Albert & Hoffman, 1995).

Феномены группировки дают множество примеров, эффектно иллюстрирующих принцип нонакцидентальности. Скажем, если два участка движутся в одном направлении и с одной скоростью, то принадлежать двум разным объектам они могут лишь случайно, поскольку те оказались случайно движущимися в одном направлении. Однако, вероятнее всего, участки эти принадлежат *одному* объекту. Итак, мы говорим, что, ежели участки эти относятся к разным объектам, сходство их движения — акцидентально, случайно; если оба участка принадлежат одному и тому же предмету — закономерно, неакцидентально. Зрительная система всегда выбирает неакцидентальную интерпретацию — то есть, второй вариант.

Сие справедливо и в отношении цельности. Мы воспринимаем, что Х-фигура — на рис. 6.4.18 А — это плод пересечения двух прямых линий (рис. 6.4.18 В), но не двух углов, которые случайно встретились своими вершинами (рис. 6.4.18 С), т. к. последний вариант явно маловероятен.





**Рис. 6.4.19** Нонакцидентальность и зрительная завершенность. Легкое смещение элементов (А) ведет к малым изменениям в геометрии изображения, если оно интерпретируется как круг, расположенный перед квадратом (В). Такое же смещение ведет к гораздо большим изменениям, если изображение трактуется как круг, подогнанный к квадрату с выемкой (С).

Чтобы понять, какая из ситуаций вероятнее и закономернее, можно понаблюдать за последствиями легких изменений в позиции объектов или в точке наблюдения. Отметим, что в данном случае, слегка сдвинув оба сегмента в любом из направлений, в устройстве самой фигуры изменится совсем немного (рис. 6.4.18 В), но легкое смещение углов в любом из направлений приведет к радикальным изменениям в структуре объекта (рис. 6.4.18 С). Логически конечно возможно, что цельность сегментов — плод случайной приводки двух углов, но намного вероятнее, что они попросту цельны.

Данный нами выше анализ факторов «фигура — фон» организации также стыкуется с критерием нонакцидентальности. Мы полагаем, к примеру, что симметричные ретинальные участки склонны восприниматься фигурами, поскольку реальные объекты гораздо более склонны к симметрии, нежели просветы между ними. Просветы между предметами оказываются симметричными только в экстраординарных и сугубо случайных обстоятельствах, и небольшие изменения в их позициях друг относительно друга или относительно наблюдателя с легкостью разрушают любую возможную акцидентальную симметрию. Тогда как, смещая симметричный предмет, мы не нарушаем его симметрии. Сие справедливо и в отношении параллельности, поскольку множество объектов имеют параллельные контуры, но просветы между объектами оказываются отграничены параллельными контурами чаще всего вследствие случайности.

Аналогично мы рассуждаем в отношении зрительной завершенности, визуальной прозрачности и мнимых фигур. Логически возможно то, что видимые участки двух областей на рис. 6.4.19 А представляют собой мозаику точно подогнанных форм, но сие потребует много большего числа совпадений, чем когда один объект просто закроет собой другой. И вновь, ежели сместить оба мозаичных объекта в каком-то одном направлении, возникнут качественные изменения в конфигурации (рис. 6.4.19 С). Эти измене-

ния — суть указание на то, что расположение — плод маловероятного совпадения. Сдвигая две наложенных друг на друга фигуры, мы получим минимум количественных изменений (рис. 6.4.19 В). То есть, скажем, акцидентальная конфигурация требует много более точного расположения, чем нонакцидентальная. То же относится и к расположению участков при мнимых контурах и прозрачности. В обоих этих случаях перцептивно предпочтительная интерпретация требует гораздо меньшего числа и гораздо меньшей вероятности совпадений во взаиморасположении, чем альтернативная.

Принцип нонакцидентальности дает очень четкое объяснение многих ключевых феноменов перцептивной организации. Мы наблюдаем ее как частный случай в реализации гельмгольцева вероятностного принципа, когда зрительная система воспринимает ситуацию как наиболее вероятно организовавшую данную стимулирующую картину на сетчатке.

Главное альтернативное толкование — это гештальт-принцип прыгннца: восприятие окажется настолько «правильным», насколько превалирующие условия обеспечат простоту и регулярность, нежели вероятность. Мы коротко уже касались этой идеи в контексте завершенности, но мы вернемся к ней в восьмой главе, когда поведем разговор о представлении формы.

## 6.5 МУЛЬТИСТАБИЛЬНОСТЬ

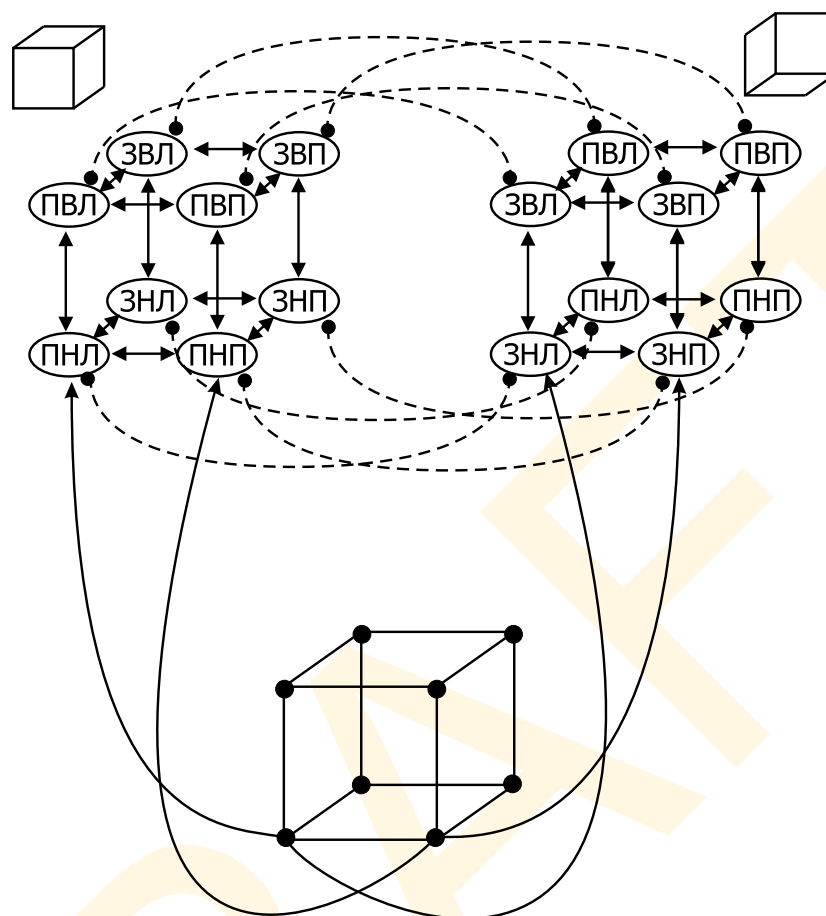
У большинства естественных сцен имеется лишь один вариант перцептивной организации, жесткий и полностью доминирующий. И это одна из причин тому, почему вопрос перцептивной организации не поднимался исследователями вплоть до 1923 г., то есть до того, как о ней заговорил Вертхаймер. Но ряд тщательно подобранных примеров иллюстрирует возможность появления различных перцептов одного и того же ретинального изображения. Первый раз мы продемонстрировали эти неоднозначности в гл. 1 на примере неккерова куба, лица-вазы и уткузайца (см. рис. 1.1.5). Эти фигуры не просто дают несколько вариантов восприятия, но *мультистабильные перцепты*: перцепты, спонтанно меняющихся двух-трех интерпретаций.

Один из наиболее интересных вопросов мультистабильности — это ее происхождение. Самый распространенный ответ на него — это гипотеза нейронного утомления, предложенная много лет назад Вольфгангом Кёхлером (1940), утверждавшим, что перцепт неопределенных фигур меняется между несколькими возможными вариантами интерпретации, за которые отвечают группы нейронов, «утомленных» длительной пульсацией. Теоретическая подоплека сего такова:

1. Различные интерпретации представлены различными паттернами нейронной активности.
2. Перцепт соответствует тому, какой паттерн активнее в данный момент времени.
3. Нейронное утомление — причина того, что различные паттерны активации доминируют в разное время.

### 6.5.1. Сетевые модели

Чтобы проиллюстрировать суть гипотезы нейронного утомления, давайте предположим, что различные паттерны нейронной активности можно идентифицировать по частоте пульсации в двух разных нейронных сетях. Давайте предположим также, что когда одна сеть становится активнее другой, в осознанном восприятии возникает соответствующая интерпретация. Согласно гипотезе нейронного утомления расширенный



**Рис. 6.5.1** Нейронная сеть, моделирующая восприятие неккерового куба. Сплошные линии представляют возбуждающие связи, пунктирные — тормозные (подробнее см. в тексте).

По материалам Rumelhart, Smolensky, McClelland & Hinton, 1986 b).

перцепт данной интерпретации ведет к утомлению нейронов в этой сети — уровень их активности падает из-за истощения биохимических ресурсов непрерывной пульсации нейронов. В итоге, уровень активации в первой сети оказывается ниже, чем во второй. В это время осознанное восприятие переключается с первого перцепта на второй.

На рис. 6.5.1 в общем виде дана сетевая модель восприятия неккерового куба (Rumelhart, Smolensky, McClelland & Hinton, 1986) — две подсети, входящих в состав одной общей (левее и правее каждой из подсетей дана соответствующая интерпретация куба). Левая подсеть представляет перцепт куба как бы при взгляде сверху, правая — снизу. Каждый узелок в каждой из подсетей представляет интерпретацию того или иного вертекса в ретинальном изображении как тот или иной угол в трехмерном кубе. Все узелки помечены согласно их трехмерной интерпретации: передний или задний (П или З), верхний или нижний (В или Н) и правый или левый (П или Л). К примеру «ЗВЛ» означает «Задний Верхний Левый» угол куба, а «ПНП» — «Передний Нижний Правый».

Два различных перцептивных толкования куба появляются в результате неоднозначной интерпретации вертексов по глубине пространства. Сия неоднозначность отражена разницей в маркировке соответствующих узелков в двух подсетях: «П» — пе-

редний, или «3» — задний. Эти пары соответствующих угловых узелков в обеих подсетях соединены с тем же самым вертексом в неккеровом кубе (рис. 6.5.1 внизу), дабы подчеркнуть то, что они обе активированы одной и той же сенсорной информацией, поступившей от соответствующего вертекса.<sup>1</sup>

Поведение данной сети демонстрирует бистабильность: активация динамически разливается по сети и она, в итоге, приходит к одному из возможных стабильных состояний. Иногда все узелки в левой сети становятся активными, а в правой — наоборот. Иногда происходит обратное. Мы говорим, что бистабильность сети проистекает из ее *архитектуры* — картины связей между узелками. В данном случае архитектура такова, что некоторые связи дают т.н. *кооперацию* пар узелков (за счет взаимной активации), тогда как другие — *конкуренцию* (за счет взаимного торможения).

### 6.5.1.1 Кооперация

Кооперация возникает тогда, когда два узелка соединены взаимновозбуждающей связью. В результате активация одного узелка усиливает активацию другого. В примере с неккеровым кубом взаимноактивирующие связи (сплошные линии со стрелками на обоих концах) представлены между соседними узелками внутри одной из подсетей. Сие значит, что, ежели один узелок одной интерпретации активирован, активация эта вначале растекается по ближайшим соседям, а затем по всей подсети. Следовательно, кооперативные связи превращают функцию каждой подсети в единое «фигура — фон», стремящееся перейти в состояние, при котором либо все единицы сети активны, либо нет.

### 6.5.1.2 Конкуренция

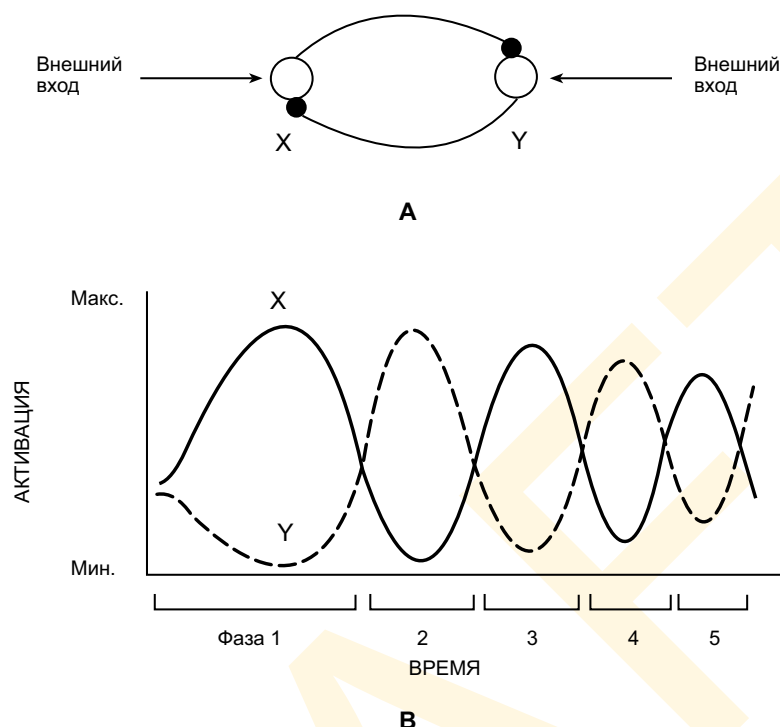
Конкуренция возникает, когда два узелка соединены взаимотормозящей связью. В результате, активация одного склонна подавлять активацию другого.<sup>2</sup> В сетке неккероваго куба взаимотормозящие связи (пунктирные линии со сплошными точками на обоих концах) представлены в различных подсетях между соответствующими узелками. Сие значит, что более активный узелок в такой паре в большей степени подавляет активацию другого узелка, чем тот его. В итоге, более активная единица выдает пульсацию. Таким образом, взаимное торможение рождает конкуренцию между взаимно исключаящими интерпретациями одного и того же вертекса как расположенного либо спереди, либо сзади, но никогда одновременно.

Совместно возбуждающие и тормозящие связи организуют единую сеть, в которой активация стремится обрести одно из стабильных состояний: либо все узелки в левой подсети активируются, а в правой — тормозятся, либо наоборот. Сие взаимоисключающее поведение сети акцентирует наше внимание на одном очень важном аспекте восприятия неккерова куба: в данный момент времени имеет место только один вариант его интерпретации.

Таким образом, сеть, показанная на рис. 6.5.1 учитывает тот факт, что куб Неккера может давать две различных и при этом взаимоисключающих интерпретации: некоторые представления активируют одну подсеть, другие — вторую. Без неких дополни-

<sup>1</sup> Возбуждающие связи, поступившие из ретинального изображения, для простоты показаны только по двум нижним вертексам; каждый вертекс активирует соответствующую единицу в обеих подсетях.

<sup>2</sup> Такую архитектуру называют иногда как «Победителю достается всё».



**Рис. 6.5.2** Динамика взаимоторможения и нейронного утомления. В нейронной сети с двумя взаимоингибирующими единицами (А), когда одна единица исходно чуть активнее другой, более активная единица (X) поднимает свою активность до максимума, а менее активная (Y) — опускает до минимума (В, фаза 1). Нейронное утомление, в итоге, приводит к реверсивной картине нейронной пульсации (фазы 2, 3, 4 и т.д.).

тельных факторов, однако, сеть не явит истинной мультистабильности, перейдя в одно из стабильных состояний восприятия, и будет оставаться в нем до тех пор, пока не изменится сам стимул. Чтобы интерпретации перемежались при одной и той же стимуляции, необходим дополнительный механизм — механизм нейронного утомления.

### 6.5.2 Нейронное утомление

Нейроны склонны терять активность после продолжительной стимуляции. Сие происходит от того, что биохимические ресурсы, требуемые клеткам для продолжения пульсации, истощаются в течение нескольких секунд, понижая частоту этой пульсации в ответ на один и тот же стимул. Для того, чтобы понять, почему сие свойство нейронного поведения порой приводит к мультистабильности, рассмотрим, что происходит между двумя взаимотормозящими (т.е. конкурирующими) единицами в простейшей нейронной сети, показанной на рис. 6.5.2 А. График снизу (рис. 6.5.2 В) иллюстрирует то, как активация узелков меняется во времени.

Положим, единица X исходно слегка активнее единицы Y, поскольку получила слегка бóльшую активацию от единицы управляющей. Следовательно, единица X тормозит единицу Y сильнее, чем единица Y единицу X. В результате, единица X становится относительно активнее, а относительная активность единицы Y — падает. Доминанта X-единицы со временем нарастает и растет до тех пор, пока активация X не



достигнет максимума, а активация единицы  $Y$  — минимума (рис. 6.5.2 В). Такое поведение системы реализуется за счет взаимотормозящих связей между узелками и в целом, независимо от утомления. Однако после того, как единица  $X$  отработала в течение длительного промежутка времени, она начинает утомляться, и частота ее пульсации падает. Снижение активации, в свою очередь, понижает степень тормозного влияния  $X$  на  $Y$ , приводя к тому, что активация  $Y$  начинает расти. В итоге, активность  $X$  падает ниже активности  $Y$  и, в какой-то момент,  $Y$  начинает подавлять  $X$  больше, чем  $X$  подавляла  $Y$ . Исходом явится то, что активация  $Y$  достигнет максимума, а у  $X$  она упадет до минимума, то есть произойдет полное обращение процесса. Чередование доминант таким манером может продолжаться бесконечно (по крайней мере теоретически).

То же происходит в сети неккерова куба: восемь пар согласованных узелков внутри подсетей (рис. 6.5.1) соотносятся через взаимное торможение, подобно двум узелкам на рис. 6.5.2 А; в результате одна из подсетей активируется и достигает максимума активности, но затем начинает нарастать нейронное утомление, понижающее доминанту одной подсети над другой. Процесс идет до тех пор, пока вторая подсеть не возьмет верх над первой и не достигнет максимума активации. Разумеется, тонкости этого взаимодействия намного сложнее, чем в простой двухузелковой нейронной сети на рис. 6.5.2 А (поскольку в процесс вовлечено большее число узелков), но общее мультистабильное поведение такое же.

Нейронные сети описанного типа могут инициировать мультистабильность перцептов неопределенных фигур. Критично значимыми моментами здесь явятся следующие:

**1. Взаимное возбуждение** внутри каждой из подсетей, представляющее альтернативные интерпретации и ведущее к кооперации внутри этих подсетей.

**2. Взаимное торможение** между подсетями, представляющее альтернативные интерпретации и ведущее к конкуренции между подсетями.

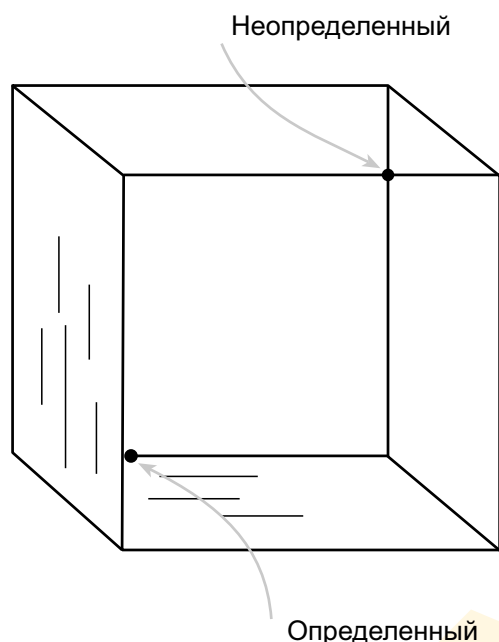
**3. Нейронное утомление**, понижающее со временем активацию высокоактивных единиц.

Разумеется, прочие неопределенные фигуры потребуют иного представления подсетей, но, если подсети эти обладают перечисленными тремя свойствами, они будут демонстрировать мультистабильность.

Данные экспериментов по восприятию неопределенных фигур повышают прогнозистическую ценность теорий нейронного утомления. Один из прогнозов таков, что ежели испытуемые на протяжении длительного времени смотрят на неопределенную фигуру, то, в итоге, они должны склониться к альтернативному перцепту (Hochberg, 1950). Действительно, подобные эффекты наблюдались при самых разнообразных вариантах неопределенной стимуляции (например, Carlson, 1953).

Еще одним предиктором теории нейронного утомления, получившим практическое подтверждение, явилось то, что частота переключения между перцептами растет со временем (Kohler, 1940). Сие ускорение должно иметь место, если утомление, возникшее в период длительного существования одного перцепта, то есть одного варианта интерпретации (назовем его «А»), не полностью рассеивается в период жизни второй интерпретации (назовем ее «В»). Когда перцепт А появится второй раз, его нейронное представление окажется частично утомленным. В результате, уровень активации в нейронном представлении А начнет падать быстрее во второй период в сравнении с первым. В той же логике сходные редукции в А- и В-перцептах появятся в третий раз, четвертый и пр.

## 6.5.3 Фиксация глаз



**Рис. 6.5.3** Стимул для изучения эффектов мультистабильности. Как показал опыт, точка фиксации играет огромную роль в формировании результирующего перцепта.

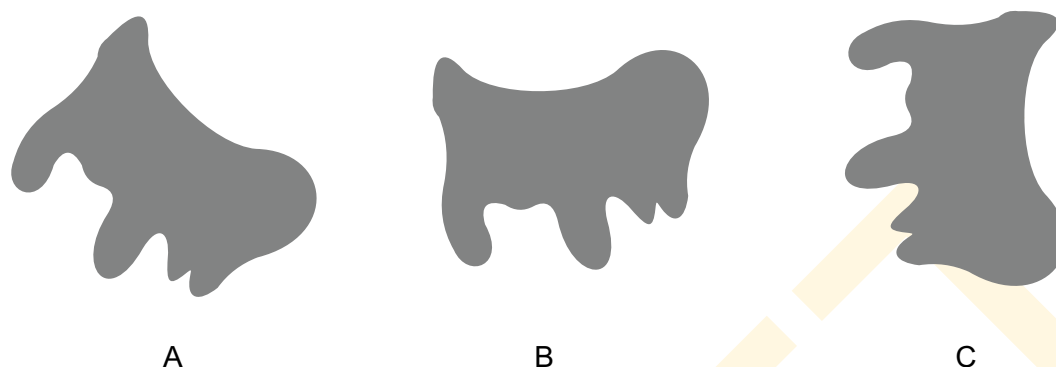
По материалам Peterson & Hochberg, 1983.

Несмотря на всю свою логичность, гипотеза нейронного утомления отнюдь не истина в высшей инстанции. Основным идеологическим конкурентом концепции нейронного утомления является идея о том, что обращение перцептов возникает из-за смены точки фиксации как следствия глазных движений. Например, ежели мы зафиксируем взгляд на верхнем центральном Y-вертексе неккерова куба, то, вероятнее всего, увидим этот куб «сверху». Фиксация на нижнем Y-вертексе дает интерпретацию куба «снизу». В целом, вертекс фиксации склонен восприниматься как ближний, что, собственно, заметил и сам Неккер. Дальнейшие исследования показали, что обращаться могут и постобразы неккерова куба и даже стабилизированные его изображения (Gregory, 1970; Pritchard, 1958). Более того, имеются данные, что причинно-следственные отношения между фиксацией глаз и переменею перцептов могут идти в противополож-

ных направлениях: глазные движения могут стать следствием обращения перцептов, а не его причиной (Noton & Stark, 1971a; Woodworth, 1938).

Однако сравнительно недавние исследования американцев Мэри Питерсон и Джулиана Хочберга (1983) показали, что локальная информация об области, окружающей точку фиксации взгляда оказывает диспропорционально мощное влияние на перцепт. Также был показан эффект наблюдательской установки на восприятие данной интерпретации, и то, как данная установка сочетанно действует с точкой фиксации. Испытуемым предъявлялась видоизмененная версия неккерова куба, имевшего лишь один неопределенный центральный вертекс и один определенный (рис. 6.5.3). Испытуемых просили сосредоточить взгляд либо на одном, либо на другом в течение 30 секунд. Плюс к тому, их просили добиться заданной интерпретации и удерживать ее, глядя либо на горизонтальную, либо на вертикальную полосу возле определенного вертекса. Общее время, в течение которого наблюдатель сообщал о каждом из перцептов, замерялось.

Питерсон и Хочберг обнаружили, что на испытуемых влиял и сам вертекс фиксации, и нацеленность на тот или иной перцепт. Целевой перцепт возникал легче, если взгляд фиксировался на неопределенном вертексе (26 с из 30), и труднее, если вертекс фиксации был определенным (20 с). Соответственно, целевой перцепт достигался с большим трудом и удерживался меньше (13 с), когда взгляд фиксировался на определенном вертексе. Оба результата показали, что локальная информация, исходящая из окружения точки фиксации взгляда, оказывает намного более мощное влияние на результирующее восприятие, нежели информация от периферии другого вертекса. Но



**Рис. 6.5.4** Неопределенные фигуры, использованные в экспериментах по изучению влияния инструкций на мультистабильность. Когда испытуемым предъявляли фигуру А в течение 30 с и без инструкций о потенциальной реверсивности, большинство интерпретировало эту фигуру без обращения. Но стоило только сообщить им о двойственности фигуры (собака — В и парвар — С) — большинство наблюдателей начинали сообщать о реверсивности.

при этом велико значение и нацеленности на тот или иной перцепт. Когда взгляд удерживается на неопределенном вертексе, наблюдатели докладывают о том, что видят целевой образ много дольше чем нецелевой (20 с против 10 с) и то же справедливо в отношении определенного вертекса (19 с против 11 с).

#### 6.5.4 Роль инструкций

Еще одно слабое место в концепции нейронного утомления выявил Ирвин Рок, усомнившийся в том, что такой простой механизм действительно достаточен для объяснения феномена мультистабильности неопределенных фигур (Girgus, Rock & Egatz, 1977; Rock, Gopnik & Hall, 1994; Rock & Mitchener, 1992). Рок заметил, что фактически во всех предыдущих исследованиях экспериментаторы информировали испытуемых о двух возможных интерпретациях *до* предъявления неопределенного стимула. Но что произойдет, если непосвященному студенту ничего заранее не говорить о характере двух различных интерпретаций? Оба ли варианта возникнут, и оба ли будут перемежаться друг с другом согласно гипотезе нейронного утомления? Или же закрепится тот вариант, что возник первым?

Рок и Митченер (1992) предъявляли 18 испытуемым три неопределенных фигуры, аналогичные показанным на рис. 6.5.4 А. Инструкция перед просмотром звучала так: «Скажите, что вы видите? Продолжайте смотреть, но позже я задам вам кое-какие вопросы». После рассматривания трех фигур последовательно в течение 30 с каждая, испытуемых спрашивали, появлялся ли перцепт альтернативный исходному. Чтобы получить сообщение о перцептивном изменении в промежуточный период, экспериментатор показывал однозначный вариант каждой из стимулирующих фигур, в частности такой, что дан на рис. 6.5.4 В и 6.5.4 С. Треть испытуемых сообщила, что *ни разу* не видела альтернативного варианта. Но стоило только дать инструкцию, как сразу началось частое перемежение с альтернативным перцептом.

Результаты эксперимента поставили под сомнение идею о том, что само по себе нейронное утомление достаточно для объяснения феномена мультистабильности. Но

вспомним, что по неккеровому кубу дабы привести нейронную сеть во состояние мультистабильности нужны два дополнительных условия:

- во внутренне кооперативных подсетях должны быть представлены альтернативные интерпретации;

- подсети должны конкурировать друг с другом через взаимное торможение.

Введение этих условий предполагает, что:

- наблюдатели ранее уже видели две разных интерпретации по каждой из альтернативной подсетей;

- испытуемые должны реализовать (на перцептивном уровне) конкурентные, взаимоисключающие отношения между альтернативными интерпретациями.

Разумеется, удовлетворить этим условиям можно только в лаборатории и поэтому сие было сознательно исключено из роковского эксперимента.

В свете теории мультистабильности оказалось вполне закономерным то, что Рок и Митченер (1992) обнаружили множество наблюдателей, которые ничего не говорили о спонтанном обращении. Наоборот, странным оказалось бы обращение перцепта у всех испытуемых. Возможно, они все же каким-то образом были и ранее знакомы с неопределенными фигурами: в конце концов, массмедиа (равно как и литература по восприятию) довольно часто обращаются к этим фигурам, и сложно найти взрослого, который бы никогда их не видел. Поэтому вполне закономерно, что Рок, Гопник и Хэлл (1994) предположили, что, вероятнее всего, дети раннего возраста ничего не знают о такого рода экспериментах. Исследователи обнаружили, что, без явных инструкций об альтернативных интерпретациях, трех-четырехлетние дети *никогда* не сообщают о спонтанном обращении даже в течение длительной экспозиции трех неопределенных фигур (в том числе лица-вазы и уткузайца с рис. 1.1.5). Эти дети также выдали достаточно низкую частоту перемены перцептов после того, как исследователи дали соответствующие инструкции.

Еще одним возможным объяснением спонтанного обращения у непосвященных взрослых наблюдателей является то, что указание пристально всматриваться в одну фигуру в течение 30 с может вызвать у наблюдателя подозрение, что в фигуре этой «что-то не так». В итоге, испытуемый начинает сам искать альтернативную интерпретацию нестандартным путем. В любом случае, до сих пор неясны условия, в которых зрительная система может спонтанно дать альтернативную интерпретацию ретинального изображения, по которому до этого дала уже вполне жизнеспособный перцепт. Сие сложный и очень глубокий вопрос, которому почти не уделяется внимания.

## 6.6 ОНТОГЕНЕЗ

Вспомним, что текущую главу мы начали с тезиса эмпиристов о том, будто окружающий мир в глазах новорожденных совершенно дезорганизован и что организация постепенно формируется на базе растущего опыта визуального общения с этим миром. Гештальт-доктрина, напротив — утверждает, что большинство базовых организационных процессов (если вообще не все) работают с самого рождения и, следовательно, не требуют обучения. Однако же существует и третий вариант рассуждений: организационные процессы хоть и не представлены при рождении, но *развиваются* предопределенным путем по мере роста малыша. Теперь, когда мы кое-что знаем о видах перцептивной организации у взрослых, мы рассмотрим то, насколько все три точки зрения адекватны процессам, лежащим в основе перцептивной организации у детей раннего возраста.

### 6.6.1 Парадигма привыкания

Ответ на поставленный выше вопрос далеко не так прост, как может показаться на первый взгляд, поскольку новорожденные не могут словесно описать свои зрительные ощущения (мо меньшей мере напрямую). Диагностика этих ощущений, требует разработки невербальных психофизических методик, некоторые из которых поставляют ценнейшую информацию о том, каков мир глазами младенцев.

Вопрос зрительной группировки у новорожденных активно изучался американскими психологами Филиппом Кельманом и Элизабет Спелке на базе т.н. *парадигмы привыкания*. В экспериментах Кельмана — Спелке, раскрывших множество ранее не-раскрываемых секретов детского восприятия, детям предъявляли никогда не виденный ими стимул и позволяли глядеть на него, сколько им захочется. К примеру, в роли стимула мог выступать большой красный квадрат. Время, которое малыш разглядывал фигуру, замерялось несколько раз (в серии попыток), а затем усреднялось. Как и следовало ожидать, период разглядывания с каждым разом (т.е. по мере привыкания) снижался: малыши все меньше интересовались стимулом, если в нем не появлялось ничего нового.

Когда время разглядывания падало до некоей критической отметки, скажем, на 80% от первоначальной величины, детям предъявляли уже два стимула: первый — тот же самый, привычный (большой красный квадрат), а другой — новый (например, маленький зеленый круг). Оказалось, что даже самые юные дети в такой стимуляции дольше глядят на новый предмет. Сие феномен, известный под названием *отвыкания*.

В исследовании вопросов зрительной группировки у новорожденных, описанная выше методика применялась Кельманом и Спелке (1983) следующим образом: исходный стимул, к которому новорожденный привык (адаптирующий стимул), был неопределенным — его центр прикрыли четырехугольником (рис. 6.6.1). Вопрос состоял в том, способны ли малыши воспринять отрезки как цельную линию позади четырехугольника или нет? В поисках ответа исследователи проверили свои предположения с помощью двух открытых стимулов — рис. 6.6.1 В и 6.6.1 С: по их логике, если ребенок воспринимает исходный «закрытый» стимул как цельный объект, покрытый плоскостью, то фрагментированному показу он должен уделить больше времени, поскольку прерывный стимул окажется новым для малыша; но если ребенок воспринимает исходный стимул как вставку между двумя несоединенными отрезками, то он больше потратит времени на цельную линию, потому что она выступит для него новым стимулом.

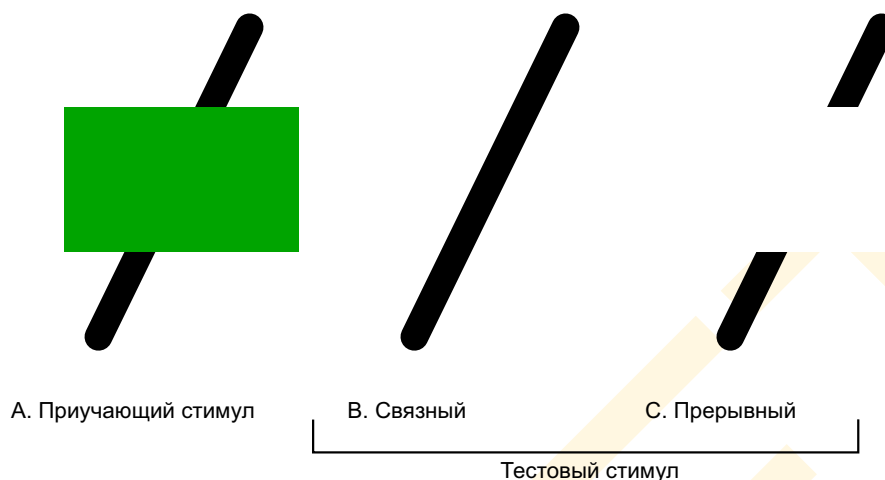
### 6.6.2 Онтогенез группировки

Поведение детей в описанной экспериментальной ситуации оказалось весьма интересным. То, как долго малыши разглядывали отучающий стимул, оказалось критично зависимым как от определенных группирующих факторов, представленных в приучающем показе, так и от возраста ребенка:

— если четырехмесячному младенцу предъявляли статический стимул (рис. 6.6.1), то он не тратил на него особого времени (картинка быстро надоедала ему) и не отдавал никаких предпочтений фрагментированной палочке, когда ее предъявляли после приучения к тестовому стимулу (время добровольного просмотра по данной демонстрации дано на левом графике рис. 6.6.2);

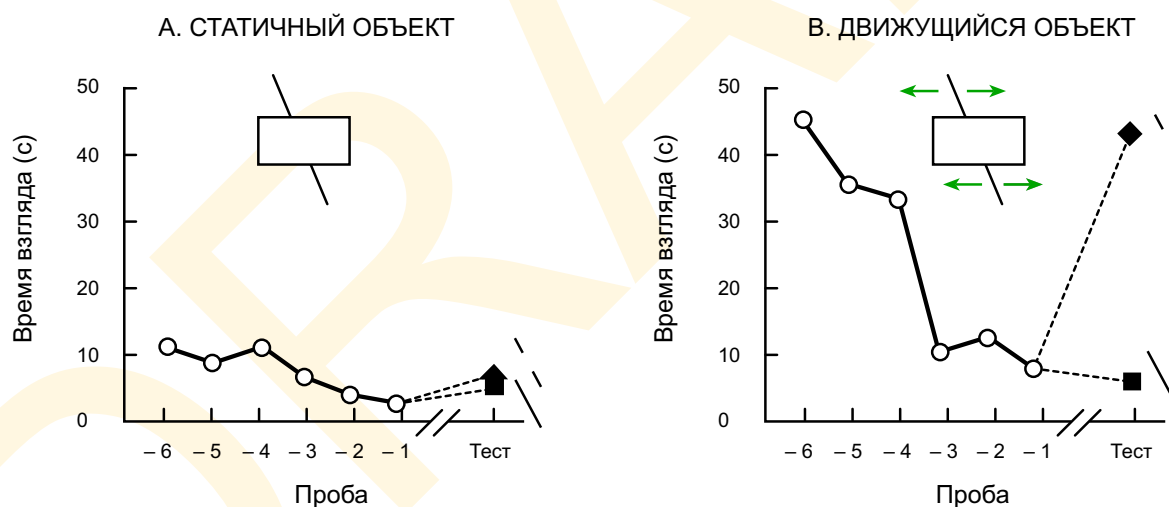
— однако же если тому же четырехмесячному малышу предъявляли *движущийся* адаптирующий стимул, то налицо была радикально иная картина отклика (показана на правом графике рис. 6.6.2) — вначале они активно интересовались приучающей





**Рис. 6.6.1** Парадигма привыкания в изучении онтогенеза перцептивной организации. Детей вначале адаптировали к неопределенному стимулу (А), в котором один фрагмент частично закрывал другой. После сеанса адаптации предъявлялся незакрытый — тестовый стимул, который был либо связным (В), либо прерывным (С). После чего оценивалась мера отвыкания.

По материалам Kellman & Spelke, 1983.



**Рис. 6.6.2** Группировка по цельности и общему току. Когда адаптирующий стимул статичен, четырехмесячный малыш не отвыкает под тестовой стимуляцией. Если адаптирующий стимул движется, дети отучаются двумя фрагментированными объектами, но не одним связным. Результат свидетельствует о том, что дети этого возраста воспринимают только движущийся стимул как единый связный предмет.

По материалам Kellman & Spelke, 1983.

стимуляцией (движение всегда интереснее), а после привыкания, явно больше внимания уделяли фрагментированному стимулу.

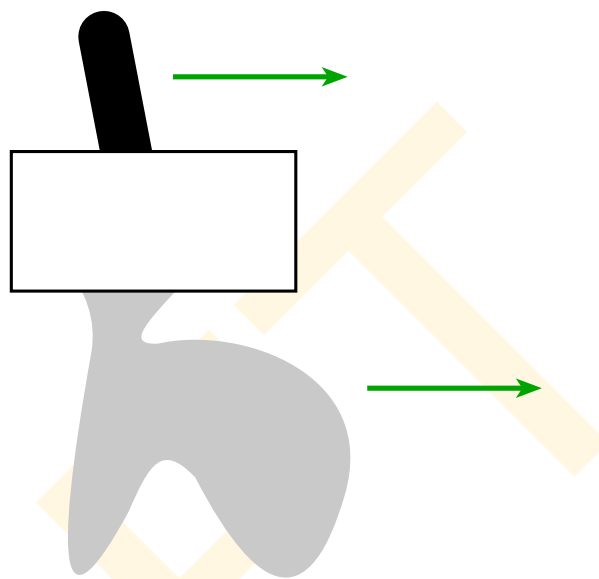
Из результирующей картины откликов мы делаем вывод, что «закрытый» адаптирующий стимул как единый, связный объект дети воспринимают только в движении,

но не в статике. Интересно и то, что любое поступательное движение дает такой результат. Более того, он возникает даже если два закрытых фрагмента стимула разнятся и по форме, и по окраске, и по размеру, и по текстуре (рис. 6.6.3).

Таким образом, мы говорим, что вертхаймеров общий ток (1923/1950) — это наиважнейший фактор перцептивной организации в зрении четырехмесячных младенцев. Правда, сие не значит, что группировка посредством общего тока является непременно врожденным признаком, поскольку может статься и плодом обучения в течение первых четырех месяцев жизни. Но даже в этом случае, она очевидно присутствует в очень раннем возрасте, намного более раннем, чем было принято считать.

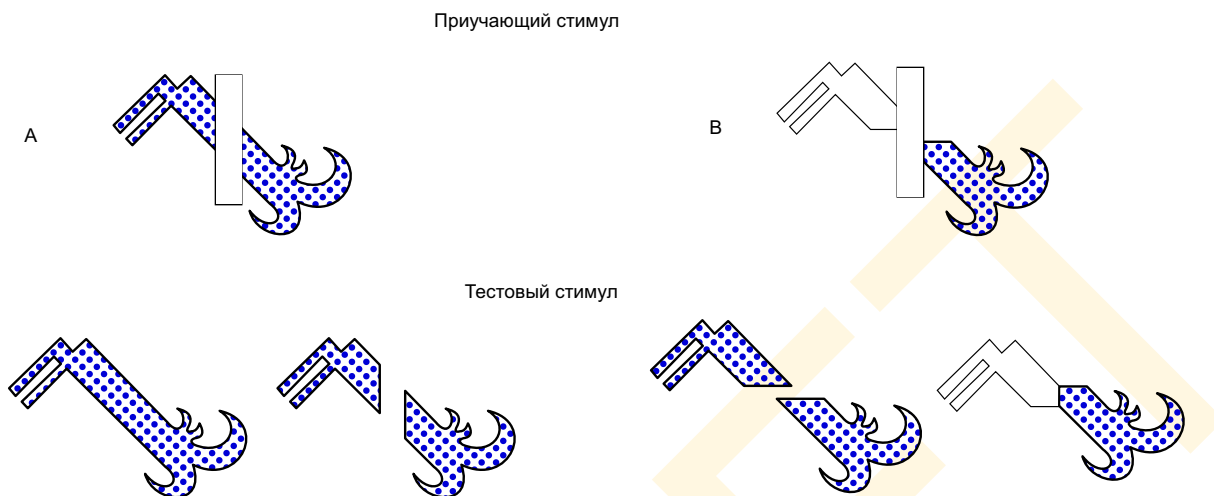
Прочие классические факторы группировки у маленьких детей не дают «взрослой» картины откликов. Если участвующий стимул состоит из двух стационарных частей, сходных по окраске, размеру, ориентации и текстуре, а края его явно продолжают через закрытый контур (см. рис. 6.6.4), дети не отдают заметного предпочтения чему-то конкретно. Сие указывает на то, что четырехмесячные малыши не используют статические факторы для завершения двух видимых концов стимула. Вопреки положениям гештальтизма мы наблюдаем то, что статическая группировка по подобию отнюдь не врожденная способность, но она либо приобретается через обучение при взаимодействии с окружающим миром, либо спонтанно развивается по мере взросления ребенка.

В каком возрасте статические факторы группировки начинают управлять зрительным восприятием объектов? Спелке (1990) докладывает о том, что когда пятимесячным детям предъявляли два адаптирующих стимула, показанных на рис. 6.6.4, они равновелико отвыкали в ответ на два тестовых, предъявленных после (внизу рис. 6.6.4). В семь месяцев, однако, дети, адаптированные к стимулу гомогенной окраски и с ровными краями, активнее переключались на фрагментированный тестовый стимул, нежели на цельный. Таким образом, способность пользоваться цельностью и подобием по окраске как информацией, влияющей на формирование перцептивной единицы, судя по всему, развивается между пятью и семью месяцами внеутробной жизни. Полученные результаты невозможно объяснить якобы неспособностью детей различать эти свойства, поскольку другие исследования показали, что младенцы вполне могут отличить гомогенную окраску поверхности от гетерогенной (Fantz, Fagan & Miranda, 1975), легко видят цельность контуров (Van Giffen & Haith, 1984) и даже симметрию (Bornstein, Ferdinandsen & Gross, 1981). Следовательно, дети младших возрастов еще не могут опираться на эти свойства в восприятии частично закрытых объектов.



**Рис. 6.6.3** Общий ток, взявший верх над статическими факторами группировки. Картина результатов по движущимся стимулам рис. 6.6.2 была получена даже тогда, когда оба участка радикально отличались друг от друга в своих статических свойствах, таких как окраска, форма и цельность.

По материалам Kellman & Spelke, 1983.



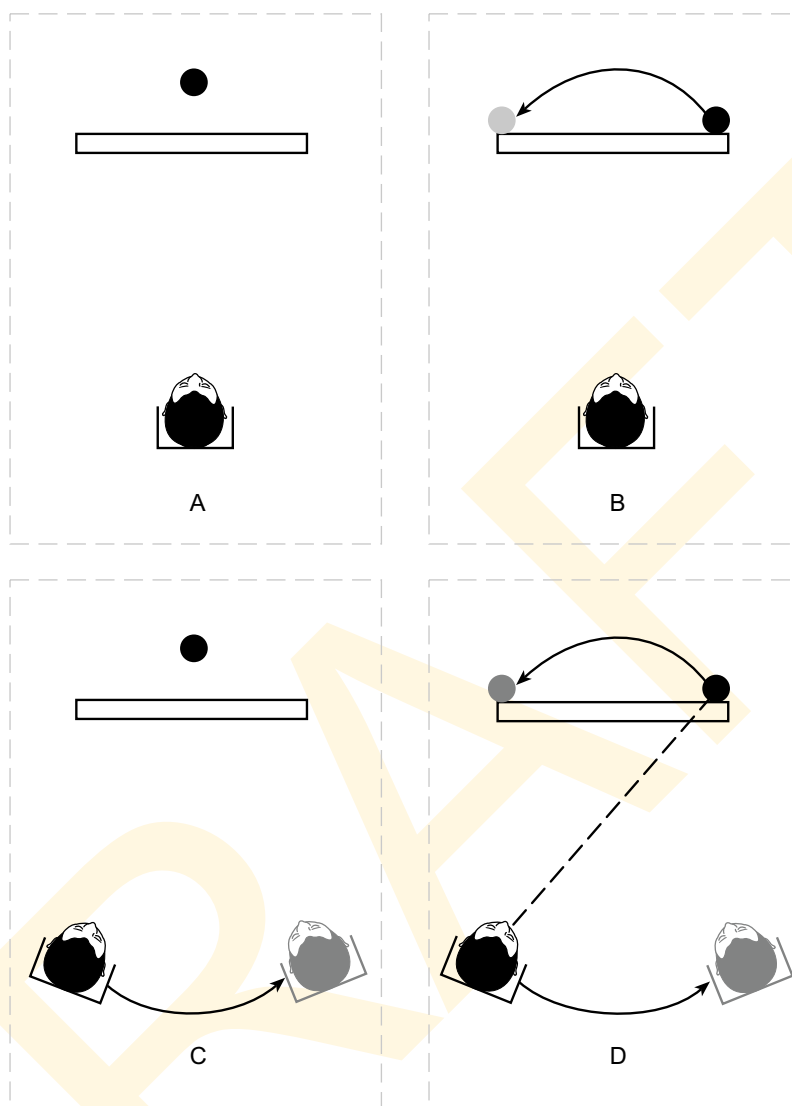
**Рис. 6.6.4** Статические факторы группировки в процессе онтогенеза обретают значимость позже динамических. Пятимесячные малыши не отучаются двумя фрагментированными участками после приучения к закрытым участкам сходным по окраске, размеру, ориентации, текстуре и цельности. Семимесячные дети — отучаются, что указывает на перцепт единого объекта позади кроющего.

По материалам Spelke, 1990.

Приняв то, что общий ток — это пожалуй, самый эффективный фактор формирования перцептивной единицы в раннем возрасте, мы задаем вполне законный вопрос: Фактор этот основан на движении *ретиальной проекции* объекта или все же на движении *самого объекта* в сцене?

Для ответа на этот вопрос Кельман, Глейтман и Спелке (1987) поставили ловкий эксперимент с т.н. центрально-закрытой стимуляцией (рис. 6.6.5): они посадили малыша в кресло на колесах, которое могло оставаться неподвижным, а могло двигаться вместе с объектом. На рис. 6.6.5 даны четыре варианта эксперимента: неподвижный объект — неподвижный наблюдатель (А); движущийся объект — неподвижный наблюдатель (В); неподвижный объект — движущийся наблюдатель (С); движущийся объект — движущийся наблюдатель (D). Результаты показали, что фрагментированный показ снимает эффект привыкания только тогда, когда закрытый объект начинает двигаться (В и D). Эффект привыкания не снимался, когда объект оставался неподвижен, даже если собственное движение ребенка в кресле приводило в движение ретиальную проекцию объекта (С). Таким образом, вероятно, даже в четырехмесячном возрасте группировка по общему току базируется на относительно сложном анализе движения объекта, но не на простом смещении его ретиальной проекции.

Этим экспериментом Спелке (1990) доказала, что ни строго эмпирическая, ни строго нэйтивистская точки зрения не верны в отношении зрительного восприятия объектов маленькими детьми. Она показала, что дети обладают примитивным врожденным механизмом группировки, исключительно зависимым от общего тока. Сие позволяет младенцам корректно организовывать их зрительные ощущения в объекты, как минимум в подходящих для этого условиях. В результате перцептивной организации движущихся предметов малыши ловят их коррелирующие свойства, полезные при идентификации этих предметов как самостоятельных единиц. Эти прочие свойства,



**Рис. 6.6.5** Ретинальное движение и движение объекта в восприятии цельных объектов маленькими детьми. Ребенок отвлекает только тогда, когда общее движение адаптирующего стимула вызвано движением самого объекта (В и D), но не наблюдателя (С).

По материалам Kellman & Spelke, 1987.

суть: гладкость, цельность контуров, гомогенность окраски и текстуры, симметрия и подобие. На базе этого знания малыш укрепляет веридикальность организации зрительного перцепта. И так продолжается до тех пор, пока ребенок достигнет взрослого уровня зрительной осведомленности о внешнем мире.

Предложенная теория проста и логична, но недостаточна для объяснения развития группировки в раннем возрасте. Когда новорожденных тестировали с движущимся стимулом, подобным тому, что показан на рис. 6.6.1, они больше смотрели на сплошную палочку, чем на прерывную, что диаметрально противоположно прогнозам, основанным на теории Спелке (Slater, Johnson, Brown & Badenoach, 1996). Однако в двухмесячном возрасте дети предпочитают прерывную палочку после приучения к движущему-

муса стимулу, при условии, что кроющий прямоугольник не слишком велик (Johnson & Aslin, 1995).

Наконец отметим, что несмотря на серьезный прогресс в изучении онтогенеза перцептивной организации, то как дети обретают способность воспринимать объекты под укрывшей их поверхностью — остается загадкой. Сегодня мы знаем, что процессы организации представлены в возрасте намного более раннем, чем предполагалось, но вряд ли они являются врожденными. Возможно, что апелляции некоторых исследователей к приобретаемому визуальному опыту отчасти имеют право на существование.



**П**олноценные трехмерные объекты — суть базовые единицы наших зрительных ощущений. Результирующий перцепт внешнего мира в нашем сознании — отнюдь не двумерная матрица цветных пятен, отнюдь не опорный скетч из примитивных краев, линий и точек, и даже не 2.5D-скетч из поверхностей разного удаления и разной ориентации — это массив цельных сущностей, состоящих из разных частей и расположенных относительно нас и относительно друг друга в трехмерном гравитационном пространстве. В предстоящих трех главах мы будем говорить о процессах, реализующих наиболее информативное восприятие свойств трехмерных объектов и их частей, т.е. отвечающих на вопрос о том, каковы предметы вокруг нас и как это знание может послужить нам в достижении наших повседневных целей?

Объекты внешнего мира мы различаем по размеру, позиции, ориентации, форме, текстуре, окраске, движению и составляющим эти объекты частям. Большинство из этих параметров сами по себе весьма сложны и обладают собственной структурой. К примеру, цветовое ощущение от поверхности слагаемо из трех компонент — цветового тона, насыщенности и светлоты (см. гл. 3). Пространственные свойства объекта тоже требуют анализа по составляющим: перцептивное расположение и ориентация предмета, к примеру, для представления реальных физических расположения и ориентации в трехмерном пространстве, должны оцениваться как минимум по трем размерностям. Более того, физическую протяженность объекта обычно рассматривают как специфически организованную совокупность его фрагментов.

В текущей главе разговор о зрительном восприятии объектов внешнего мира мы начнем с обсуждения того, как наблюдатель реконструирует действительные свойства этих объектов из свойств их ретинальных проекций. Затем поговорим о том, как перцепты объектов анализируются по составляющим.

**Константность и иллюзии.** Возможно самым главным фактом осознанного зрительного восприятия объектов и их свойств является то, что имеет место более плотная корреляция перцепта с действительными свойствами *дистального стимула* (объекты внешнего мира), нежели со свойствами стимула *проксимального* (ретинальная проекция объектов). Сие столь очевидно, что мы не придаем этому факту осознанного значения. Однако, объяснить то, чем обеспечивается такая корреляция — не так-то просто.

Вообразим, что прямо сейчас мы видим собаку, свернувшуюся калачиком на краю дивана. Ее местоположение мы оценим отнюдь не в логике ретинальной локали (хотя животное и занимает, скажем, правый верхний квадрант поля зрения) — мы воспримем ее размер в реальных физических показателях (примерно 0.5 м в диаметре), но никак не ретинальных размерностях (примерно 6° зрительного угла от точки наблюдения). Как же сие возможно? Мы помним, что свет, отраженный от объектов, проецируется на сетчатку наших глаз в виде пары двумерных оптических изображений, что, собственно, и является отправной точкой зрительного акта. Но каким же образом из этих двух проекций мы умудряемся восстановить действительные свойства внешних объектов?

Как было показано в 1-й главе, факт веридикального зрительного восприятия константных физических свойств объектов внешнего мира (а не зыбких свойств их ретинальных проекций) именуется *перцептивной константностью*. Один из примеров мы

уже разобрали в гл. 3, когда говорили о константности цветовой: суть разговора сводилась к тому, что способность нашего зрения восстанавливать информацию об отражательной способности поверхности (невзирая на подвижность освещения, меняющего спектральное распределение потоков, стимулирующих сетчатку) объясняется совокупным действием серии очень сложных процессов. Нечто подобное происходит и при восприятии пространственных свойств объектов, в частности: размера, формы, ориентации и расположения. Веридикальность восприятия действительных свойств объектов достигается как через бессознательные, так и осознанные процессы (Epstein & Broota, 1986; Epstein & Lovitts, 1985).

К вопросу перцептивной константности существуют два основных научных подхода.

Первый классический подход — гелмгольцев (который иногда называют *непрямым*) — это количественная интеграция ретинальных проекций объектов с иными источниками информации, способными выявить действительные свойства объектов внешнего мира. Например, размер предмета можно высчитать из размера его ретинальной проекции и удаления от наблюдателя (подробнее см. гл. 5). Такой подход назван «непрямым» поскольку, с одной стороны, предполагает, что константность отчасти управляема и другими атрибутами восприятия (в случае размера — перцептивным расстоянием), с другой — потому что контрастирует с прямым подходом, на котором настаивал Джеймс Гибсон (1950, 1966).

Гибсон (см. гл. 5) показал, что свойства объектов, которые остаются неизменными в разных просмотровых условиях, часто коррелируют с постоянными высокого уровня в оптической структуре изображения. К примеру, когда мы видим предмет, лежащий на копланарной текстурированной плоскости земли, количество текстурных элементов, закрытых этим предметом, неизменно и независимо от просмотровых расстояний и перспективы. Следовательно, переменные высокого уровня достаточны для обеспечения константности размера без обращения к неоптическим (т.е., эвристическим) источникам информации. Позже мы вернемся к этому разговору.

Наконец, несколько неожиданно прозвучит то, что те же самые процессы, которые обычно реализуют константность перцепта (то есть его веридикальность), иногда порождают иллюзии. Как мы показали в гл. 1 и 2, причина может стать в том, что веридикальность порой требует привлечения эвристических механизмов, догадок, которые обычно хоть и верны, однако же — не всегда. Когда догадки эти правильны (чаще всего) — перцепт веридикален. Но когда ошибочны — мы воспринимаем ситуацию, систематически отличной от реальности — то есть зрим иллюзию. В некоторых случаях обнаруживается, что константность и иллюзия — это стороны одной перцептивной «медали».

**Проксимальный и дистальный режимы восприятия.** Давеча мы показали, что перцептивная константность относится к восприятию, согласованному с действительными свойствами объектов внешнего мира, но не со свойствами ретинальных проекций этих объектов. Точнее, либо-либо: либо перцепт относится к предметам (дистальная стимуляция), либо к оптическим свойствам проекции этого предмета на сетчатку (проксимальная стимуляция). И хотя так чаще всего и происходит, иногда все же бывает, что наше ощущение относится одновременно к обоим ситуациям. Например, при восприятии светлоты (рис. 7.0.1): теневой участок белой стены храма мы видим и как более темный, и, в то же время, воспринимаем его обладающим той же отражательной способностью, что и освещенная часть стены. Аналогично при восприятии формы: экран компьютера на рис. 7.0.2 мы воспринимаем как прямоугольный, но при этом четко понимаем, что его ретинальная проекция — трапеция. Как же толковать сии очевидные противоречия?

Сегодня ответ на этот вопрос звучит так: множество исследований свидетельствуют о существовании двух возможных режимов зрительного восприятия — *проксимального* и *дистального*. Режим восприятия, который мы будем называть «проксимальным», отражает основные свойства ретинального изображения, то есть — проксимальной стимуляции. Режим, который мы назовем «дистальным» — свойства объектов внешнего мира, т.е. стимуляции дистальной.<sup>1</sup>

Результирующее зрительное ощущение можно рассматривать как некую перцептивную смесь обоих компонентов: один может доминировать над другим в разных ситуациях, в зависимости от задач и намерений наблюдателя.

Перцептивная константность — это по определению признак дистального режима восприятия, поскольку относится к свойствам дистальной стимуляции, остающимся неизменными в разных условиях просмотра. В самом деле, дистальный режим доминирует в большинстве заурядных ситуаций, таких, в частности, как движение в пространстве, координация взаимодействия тела с объектами и стандартное сравнение этих объектов друг с другом. Есть прямой смысл оставаться строго в дистальном режиме, когда на чьем-нибудь дне рождения нужно схватить кусок торта побольше: проксимальный режим может сыграть здесь злую шутку и подсунуть нам меньший из-за того, что он попросту оказался ближе и, следовательно, его ретинальный размер больше.

Бывают ситуации, однако, когда очевидно предпочтителен проксимальный режим зрительного восприятия. Предположим, мы пытаемся нарисовать реалистичную картину некоей сцены. Чтобы выполнить сие максимально точно, нам нужно создать на холсте изображение, которое будет стимулировать сетчатку тем же образом, что и оригинальная сцена. Следовательно, мы должны изобразить укрупненную (и неинвертированную) версию оптического изображения, стимулирующего нашу сетчатку. В данной ситуации критично важно соблюсти относительный размер проекций значимых предметов и перенести это соотношение на полотно. То же касается и прочих свойств изображенных объектов, таких как форма, окраска, ориентация и позиция. Опытные



**Рис. 7.0.1.** Константность светлоты. Стены храма воспринимаются равномерно окрашенными (т.е. имеющими неизменный коэффициент отражения), несмотря на колоссальный разброс фотометрической яркости по разным участкам.

<sup>1</sup> Джеймс Гибсон (1950) режимы эти называл соответственно режимом «поля зрения» и режимом «зримого мира»; Эрнст Мах (1978) окрестил их «проксимальным» и «константным» режимами; Ирвин Рок (1983) — «проксимальным» и «внешним». Мы же будем говорить о «проксимальном» и «дистальном» режимах зрительного восприятия, дабы определить их максимально четко и подчеркнуть их взаимнопротивоположность.



**Рис. 7.0.2.** Константность формы. Экран компьютера в данном ракурсе воспринимается совершенно прямоугольным, даже несмотря на то, что его проекция на сетчатку — неправильная трапеция.

то мы, включив дистальный режим восприятия, быстро поймем, что диаметры одинаковы. Но если затем нас попросят оценить видимые диаметры от точки стояния, то мы быстро переключимся на режим проксимальный и сообщим, что ближний диск выглядит крупнее.

Существование проксимального и дистального режимов существенно усложняет изучение вопроса константности восприятия. Коль скоро осознанная перцепция двуединна, перцептивная оценка может легко пойти либо одним, либо другим путем, в зависимости от полученных инструкций. Возможно, по этой причине перцептивная константность редко бывает полной. Действительно, большинство экспериментов демонстрируют систематическое смещение от точного восприятия объективных свойств предмета к свойствам его проксимального изображения на сетчатке. Феномен сего перцептивного компромисса между проксимальным и дистальным соответствиями именуется *гипоконстантностью*. Мы будем много говорить о ней на страницах текущей главы.

## 7.1 ПРОСТРАНСТВЕННЫЙ РАЗМЕР

Итак, первый вопрос, который нам должно обсудить, — это каким образом наше зрение выводит размер реального объекта из его ретинальной проекции? Разумеется, реальные габариты объекта определяемы просто из размера его ретинального изображения: при прочих равных объекты больших габаритов дадут больший размер своих проекций на сетчатку. Из материала 4-й главы нам уже известно, что в зрительной коре имеются т.н. пространственно-частотные клетки, т.е. клетки, чувствительные к разным размерам стимула. Различие в активности этих клеток может стать базовым механизмом восприятия размера объектов сцены.

живописцы активно пользуются тем, что мы назвали «проксимальным режимом восприятия», благодаря чему им и удается реалистично и точно передать свойства оптических проекций на холст. Когда мы глядим на эти полотна, наше зрение использует эти свойства для реконструкции исходной сцены (заполненной предметами соответствующих размеров, формы, окраски и т.д.).

То, в каком режиме мы находимся в данный момент, может целиком зависеть от внешних инструкций (Carlson, 1960, 1977). Если, к примеру, предъявить нам два DVD-диска, расположив их на разном удалении от нас, и попросить оценить их диаметры,



Однако сходу понятно, что ретинальные габариты не могут стать основой перцепции размеров фактических, поскольку ретинальный размер в том числе зависит от расстояния между наблюдателем и объектом. То есть, при прочих равных, ретинальная проекция ближних предметов окажется крупнее, чем проекция дальних: если мы поставим ладони в нескольких десятках сантиметров перед лицом, они целиком заполнят поле зрения; если же мы вытянем руки вперед, ладони займут лишь малую его часть. Следовательно, из габаритов ретинальной проекции зрительная система должна каким-то образом вывести и размер объекта, и, в том числе, его удаление от наблюдателя.

### 7.1.1 Константность размера

Константность размера относится к способности нашего зрения адекватно воспринимать действительные габариты объектов, несмотря на изменчивость их ретинальных проекций (в результате изменчивости просмотрового расстояния). Повседневный зрительный опыт свидетельствует, что способность эта вполне адекватна большинству объектов и в очень большом диапазоне просмотровых расстояний. То есть, мы крайне редко ошибаемся в выборе самого большого куска торта, причем вне зависимости от расстояния до него. В некоторых случаях, правда, константность размеров дает сильные сбои: автомобиль с самолета смотрится не крупнее муравья, а дома — детской игрушкой, макетом. Как понимать и как толковать эти факты?

#### 7.1.1.1 Отношения размера и расстояния

Давайте начнем с простого количественного описания нашего восприятия габаритов объекта. Как давеча было показано, базой перцепции размера является то, что габариты проекции объекта на сетчатку определяют два разных фактора: фактический размер этого объекта и его удаление от наблюдателя. Характер отношений между фактическим размером, размером ретинальным и расстоянием можно выразить математически.<sup>1</sup> В текущем контексте ключевым для нас явится реверс этих отношений: если расстояние до объекта известно, то, в связке с ретинальным размером, эти данные можно использовать для «вычисления» габаритов реального объекта сцены.

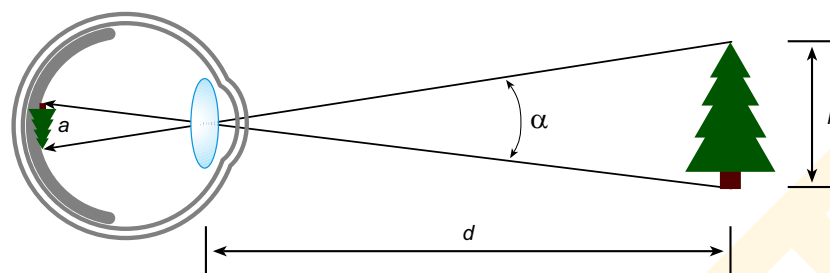
Уравнение, описывающее отношения «размер — расстояние», то же, что и в гл. 5, хотя и несколько иначе:

$$h = 2d \times \operatorname{tg}(a/2) \quad (7.1),$$

где  $h$  — это линейная высота реального объекта,  $d$  — расстояние до этого объекта,  $a$  — ретинальный размер в градусах зрительного угла (см. рис. 7.1.1). Отметим, что сейчас мы выразили размер объекта через отношение расстояния к ретинальному углу, тогда как в гл. 5 расстояние до объекта было выражено через отношение размеров объекта к размерам его ретинальной проекции. Из приведенного здесь уравнения следует, что, ежели нам известно фактическое расстояние до объекта ( $d$ ) — скажем, из данных бинокулярной конвергенции — то фактические габариты этого объекта ( $h$ ) можно вывести из этой величины и величины ретинального размера ( $a$ ).

<sup>1</sup> См. раздел 5.5.4, где мы говорили об этом в контексте восприятия глубины пространства, акцентировав внимание на том факте, что, зная размер ретинальной проекции объекта и, сделав некоторые допущения о его габаритах, можно получить информацию об удаленности этого объекта от наблюдателя.





**Рис. 7.1.1.** Соотношение размера и расстояния. Расстояние до объекта можно определить из высоты этого объекта ( $h$ ) и тангенса его зрительного угла ( $\alpha$ ).

### 7.1.1.2 Иллюстрация константности размера

Феномен константности размера мы можем продемонстрировать с помощью пальцевого эксперимента, описанного в гл. 5. Поставим указательный палец одной руки на длину этой руки, а палец второй руки — на половину расстояния. Теперь начнем постепенно подносить ближний палец к лицу, пока он не закроет собой дальний. Откроем оба глаза, то есть дадим доступ бинокулярной информации и, переводя взгляд с одного пальца на другой и обратно, сравним их размеры как объектов. Разумеется, восприниматься они будут примерно одинаковыми — что и есть на самом деле.

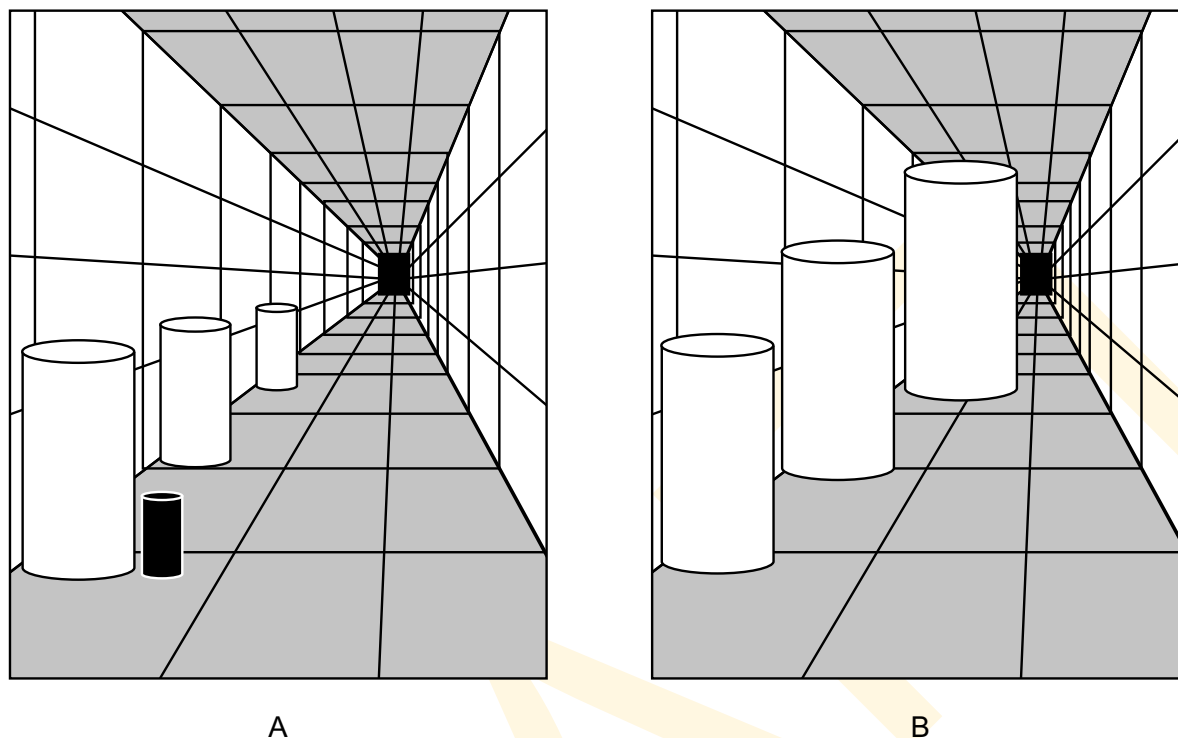
Сие пример константности размера, поскольку габариты обоих пальцев в нашем сознании остаются неизменными, даже несмотря на радикальное и впечатляющее отличие в размерах их ретинальных проекций: закроем один глаз и совместим кончики пальцев — мы обнаружим, что ретинальная высота ближнего минимум втрое больше высоты дальнего (который в аккурат уместается в одну фалангу ближнего).

Итак, нам становятся понятными условия, в которых проксимальный режим восприятия походит более на внедренную и даже доминирующую осведомленность. Мы сильнее осознаем отличие в размерах проекций двух идентичных объектов при двух разных расстояниях, когда:

- отличие в ретинальных размерах велико;
- проекции расположены рядом;
- стереоинформация редуцирована (например, за счет монокулярности);
- мы *сознательно* пытаемся найти эти отличия (т.е. перейти в проксимальный режим).

Константность размера имеет место (хотя бы отчасти), когда информация о глубине пространства поступает исключительно по пикториальным факторам. К примеру, рис. 7.1.2 А — перспективная информация дает устойчивое впечатление глубины коридора. Три цилиндра при этом воспринимаются одинаковыми по реальному размеру: ретинально больший видится ближним, меньший — дальним. Однако, несмотря на отличия по перцептивной глубине мы знаем (или можем узнать), что имеют место отличия в ретинальных изображениях этих цилиндров. При этом отличия эти могут стать весьма неожиданными: у черного цилиндра сравнения тот же ретинальный размер, что и у дальнего белого, то есть составляет лишь треть линейного размера ближнего цилиндра.

Рис. 7.1.2 В вызывает иллюзию, в основе которой лежит тот же феномен: ретинальные габариты цилиндров одинаковы, но при этом задний воспринимается как огром-



**Рис. 7.1.2** Коридорная иллюзия. А — несмотря на то, что ретинальный размер «дальнего» цилиндра оценивается нами всего лишь как немногим меньший ретинального размера ближнего, мы будем сильно удивлены отличием фактическим: размер черного цилиндра сравнения точно такой же, что и у дальнего белого. В — ретинальные габариты всех трех цилиндров одинаковы, но «дальний» при этом выглядит много крупнее двух других.

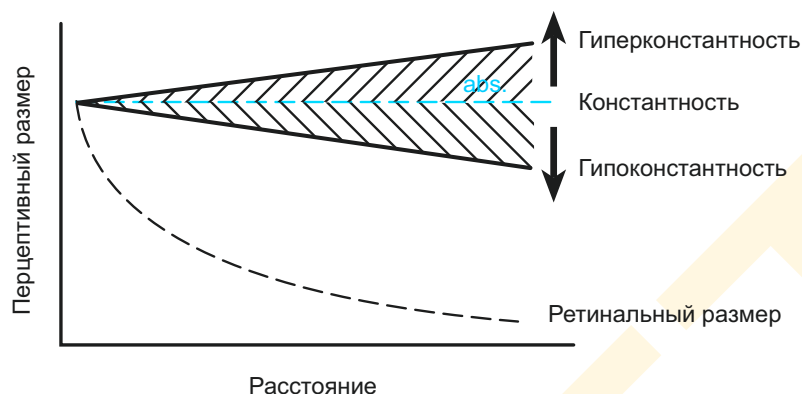
ный, намного больший переднего. Иллюзия эта во многом схожа с иллюзией Понцо (рис. 7.1.10), но с более полными данными о глубине пространства.

Многие отмечают, что как ни старайся, не получается оценить ретинальные размеры объектов, расположенных в реальной сцене на разном удалении. Сие яркое свидетельство того, что механизмы константности размера продолжают работать даже когда мы пытаемся подавить их, сознательно переходя в проксимальный режим восприятия. Поэтому живописцы стараются масштабировать объект на полотне, закрыв один глаз и оценив размер проекции этого объекта в сравнении с большим пальцем или кистью на вытянутой руке. Сие дает объективный размер ретинальной проекции, не искаженный тенденцией нашего зрения учитывать расстояние при восприятии габаритов того или иного предмета.

### 7.1.1.3 Уход от константности

Читатель уже наверняка заметил, что многое нами сказанное о константности размера звучит весьма неопределенно: встречаются такие слова и сочетания как «приблизительно», «в определенных условиях» и пр. Причина тому — несовершенство исходной экспериментальной базы рассуждений.

В опытах по константности размера испытуемым обычно показывают стандартный объект заданного габарита на заданном расстоянии (скажем, линию в 60 см длиной и на расстоянии 2 м от наблюдателя), а затем просят выбрать объект того же размера из



**Рис. 7.1.3** Уровни константности размера. При идеальной константности перцептивный размер не меняется с изменением расстояния (горизонтальный пунктир). Фактическая же константность несколько отклоняется от идеальной в зависимости от внешних инструкций, но при этом *никогда* не соответствует ретинальному размеру.

нескольких, расположенных на ином удалении, скажем, на 3 м дальше. Процедуру повторяют по многим вариантам удаления. Когда наблюдатели выбирают предмет сравнения, размер которого объективно совпадает с размером тестового объекта (абсолютная константность), то график выборки являет собой прямую, помеченную на рис. 7.1.3 как «abs.»; когда они выбирают объект сравнения по равенству размера ретинальных проекций, график превращается в «Ретинальный размер» (пунктир на рис. 7.1.3).

Результаты обычно ложатся где-то между крайними вариантами, но ближе к «abs.». Наблюдатели очень точно попадают в объективный размер, когда поступает добротная бинокулярная информация, а также когда они проинструктированы о том, каков действительный размер объекта. Когда качество информации о глубине пространства падает, например, когда наблюдатели смотрят одним глазом и/или через маленькое отверстие — выбор склоняется к ретинальному равенству (Holway & Boring, 1941). Итак, судя по всему, испытуемые идут на компромисс между константностью (дистальный режим восприятия) и размером ретинальной проекции (проксимальный режим).

⇒ Когда объекты воспринимаются расположенными ближе, чем на самом деле, мы говорим о *гипокопстантности* восприятия размера.

На константность размера можно повлиять с помощью различных инструкций о том, как следует сравнивать габариты объектов. Когда наблюдатели самостоятельно говорят о том, насколько крупным они *видят* объект, они обычно устойчиво держат константность по размеру, с легким-легким уходом в гипокопстантность. Но стоит только сказать им, каков объект на самом деле — то есть, дать, как говорят, объективную инструкцию — сразу налицо гиперкопстантность (Carlson, 1960; Gilinsky, 1955).

⇒ О *гиперкопстантности* восприятия размера мы говорим как о тенденции уравнивать стандартный объект с дальним объектом сравнения, то есть с бóльшим.

Гиперкопстантность возникает от того, что наблюдатель *знает* о том, что визуальный размер уменьшается с расстоянием и *гиперкомпенсирует* этот факт, выбирая объект, который выглядит несколько крупнее.

Однако в большинстве случаев гипо- или гиперконстантность дают веридикальные перцепты. Но константность размера может-таки рухнуть, например в крайних ситуациях, когда мы рассматриваем людей, машины и дома с больших расстояний, скажем, с крыши небоскреба или из иллюминатора взлетающего самолета. Будь константность размера абсолютной — всё воспринималось бы в подлинный размер (т.е., как если бы мы стояли рядом с этими предметами). Но на деле всё иначе. Действительно, предметы с высоты воспринимаются даже еще меньше, чем следовало бы ждать при гипоконстантности: мы их видим совсем мелкими, буквально игрушечными. Отчего же так происходит? Чем объяснить одновременные триумф и провал константности восприятия размера?

#### 7.1.1.4 Учет расстояния

Было предложено несколько вариантов, толкования феномена перцептивной константности размера. Классическое объяснение лежит в рамках гельмгольцевой идеи бессознательной инференции, предполагая, что зрительная система эффективно вычисляет размер объекта из размера его ретинальной проекции, а также перцептивного расстояния, используя при этом нечто сходное с уравнением 7.1. Сия, как говорят, *непрямая* концепция часто звучит как теория *учета расстояния* (Epstein, 1973; Rock, 1975). Идея ценна своей универсальностью: она применима к любой ситуации, в которой наблюдатель располагает информацией об удалении объекта. Следовательно, механизм никак не зависит от присутствия других объектов, которые являются неотъемлемыми компонентами прямых концепций перцептивной константности размера (см. ниже).

Казалось бы, на базе соотношения «размер/расстояние», концепция бессознательной инференции должна давать абсолютную константность по перцептивным габаритам объектов. Из-за чего же тогда возникают отклонения от нее? Ответ прост: очевидно, что погрешности в перцепте расстояния ведут к погрешностям в восприятии размера. В частности, к гипоконстантности, когда большие расстояния видятся меньше, чем они есть на самом деле. Если наблюдатель воспринимает объект ближе, чем он в реальности, расчет размера этого объекта по вышеприведенному уравнению непременно приведет к тому, что размер воспримется меньше реального.

Отметим, что теория способна прогнозировать и гиперконстантность: если объект видится расположенным дальше, чем на самом деле, он воспримется крупнее, чем в реальности.

То есть, в логике бессознательных инференций отклонения от константности восприятия размера объясняются погрешностями в оценке расстояния до целевого объекта.

Изящная теория, но верная ли? Увы, ответ далеко не очевиден и совсем не прост. Несмотря на многие «за» — многое «против»: к сожалению, имеется ряд обескураживающих примеров, в которых размер объекта систематически *переоценивается*, но при этом расстояние до него *недооценивается* (Gruber, 1954). Сие диаметрально противоположно прогнозам не прямой теории. Возможно, компромиссом явится следующая мысль: обстоятельства эксперимента могут оказать большее влияние на результат, нежели собственно перцептивное расстояние, т.е. — условия просмотра, указания экспериментатора, требуемые виды оценки, а также индивидуальные особенности наблюдателей определяют исход опытов (см. Epstein, Park & Casey, 1961).

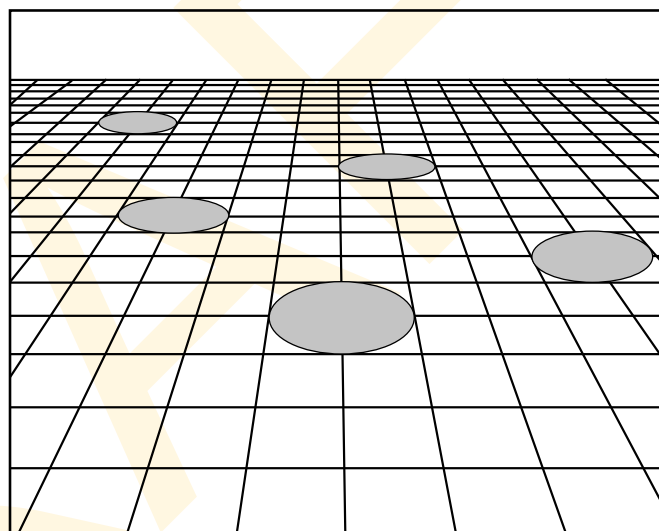
### 7.1.1.5 Текстурная окклюзия

Принципиально иной теоретический подход к зрительной оценке размера предметов предложил Гибсон в своей знаменитой работе по прямому восприятию (см. раздел 2.1.3). Он предположил, что потенциально значимым источником информации о размере объекта является текстура, поскольку текстурные элементы, закрытые объектом, создают некую вполне равномерную шкалу, относительно которой можно перцептивно оценить размер этого объекта. Как показано на рис. 7.1.4 А, если два плоских объекта расположены на общей текстурной поверхности, но при этом на разных расстояниях от наблюдателя, на относительные размеры этих объектов укажет нам число закрытых ими текстурных элементов. Если, скажем, предметы закрывают примерно одно и то же число элементов, то и габариты этих предметов должны быть примерно одинаковы. Сие объяснение вполне согласуется с доктриной прямого восприятия, поскольку идентифицирует перцептивный размер по высокоуровневому параметру стимуляции — числу закрытых элементов текстуры. И ценно то, что исключаются какие-либо промежуточные перцепты, в частности визуальное расстояние до объекта.

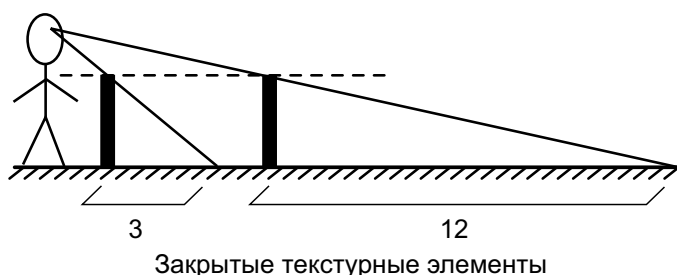
Бесспорно, текстурная окклюзия может дать нам сведения об относительном размере объекта, но при этом важно понимать, что ее применимость ограничивается определенными условиями:

1. Объекты должны сравниваться на одной и той же текстуре поверхности. Если они расположены на разных текстурах или на общей нетекстурной поверхности, то, рассуждая в логике Гибсона, мы ничего не сможем сказать об их относительном размере.

2. Текстурные элементы фоновой поверхности должны быть примерно одинаковы. Если их размер переменчив, то есть если элементы систематически меньше или больше, чем закрытые



А



В

Закрытые текстурные элементы

**Рис. 7.1.4** Текстуральные факторы константности восприятия размера. В определенных условиях объекты одного размера закрывают одно и то же число текстурных элементов (А). Если же эти объекты приподнять над текстурной плоскостью, связь исчезнет (В).

Подробнее в тексте.



объектом сравнения, информация, извлекаемая зрением из текстурной окклюзии, не даст веридикального перцепта размера.<sup>1</sup>

3. Объекты должны быть копланарны текстурной поверхности (Gillam, 1981). Если объекты возвышаются над нею, скажем, как деревья над травой или как мяч над паркетом, соотношение распадается (рис. 7.1.4 В).

Итак, мы видим, что установленные рамки умаляют ценность текстурной окклюзии как индикатора размера объектов естественных сцен. Более того, в отличие от концепции бессознательных инференций, идея текстурной окклюзии не может дать общего объяснения восприятию размера, поскольку оперирует информацией только об относительном размере объектов. Действительно, текстурные элементы организуют устойчивую шкалу, относительно которой можно оценивать габариты любого копланарного объекта, но эта информация ничего не говорит наблюдателю о том, каков *абсолютный* размер объекта, если неизвестен абсолютный размер самих текстурных элементов. Поскольку мы способны воспринимать абсолютные размеры предметов в самых разнообразных ситуациях, в деле очевидно участвует некий дополнительный фактор.

При этом стоит отметить, что в контексте восприятия размера текстура может стать ценной не только за счет окклюзии ее элементов интересующим предметом. Как уже обсуждалось в 5-й главе, текстурные градиенты могут дать количественную информацию о глубине пространства. Следовательно, сие позволяет определить размер объекта через бессознательную инференцию на базе перцептивного расстояния до объекта. Однако такое объяснение весьма отлично от гибсоновой идеи текстурной окклюзии, поскольку логика строится на том, что перцептивный размер предмета *опосредован* перцептом расстояния до него.

#### 7.1.1.6 Относительный размер

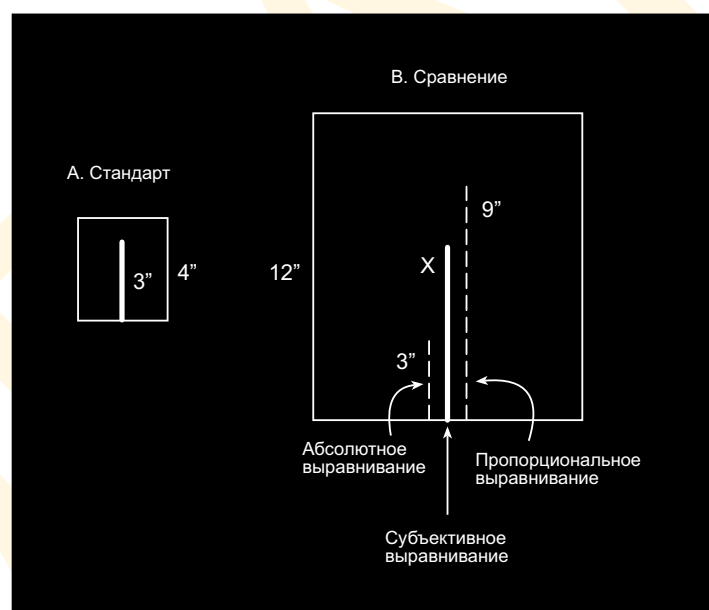
Еще одно толкование, согласуемое с прямыми теориями восприятия размера и его константностью, можно вывести из инвариантности относительных размеров соседних объектов. Согласно сей гипотезе — *гипотезе пропорциональности*, — константность размера весьма легко достижима, поскольку соотношение габаритов ретинальных проекций объектов, расположенных на одном и том же расстоянии друг от друга, сохраняется в широчайшем диапазоне удалений этих объектов от наблюдателя (Rock & Ebenholtz, 1959). Если, к примеру, человек в два с половиной раза выше собаки, то соотношение это остается неизменным, на каком бы расстоянии они не находились от фотографа — 3 метра или 30 (рис. 7.1.5).

Такой подход к восприятию размера аналогичен концепции константности светлоты (см. гл. 3); он также схож и с гибсоновым концептом текстурной окклюзии, поскольку базируется на жестком постоянстве соотношений размеров кроющего объекта и текстурных элементов. И сие радикально отлично от концепта бессознательных инференций, поскольку не требует учета расстояния: фактор расстояния оказывается попросту излишним, поскольку предполагается, что объекты — о соотношении размеров которых идет речь — равноудалены от наблюдателя. Отметим, что сие весьма и весьма серьезное ограничение, существенно сужающее рамки применимости прин-

<sup>1</sup> Если зрительная система не располагает информацией о расстоянии до текстурных элементов, что может пригодиться в оценке их равномерности по размеру, она будет вынуждена предположить, что все текстурные элементы одинаковы. Иными словами, сие толкование механизмов константности априори предполагает неизменность размера текстурных элементов.



**Рис. 7.1.5** Относительные размеры статичных объектов не меняются при движении наблюдателя или же при синхронном движении самих объектов. Ретинальная проекция человека останется в два с половиной раза больше проекции собаки независимо от расстояния до фотографа. Но соотношение изменится, если собака бросится навстречу к нему.



**Рис. 7.1.6** Влияние относительности размера на его восприятие. Наблюдателям предъявлялся некий отрезок внутри стандартного окна (А), после чего испытуемых просили так подобрать длину Х-отрезка во втором окне (В), чтобы уравнивать ее с длиной исходного. Оказалось, что в выборе длины испытуемые в большей мере ориентируются на относительный размер, нежели абсолютный.

ципа пропорциональности. Если собака неожиданно рванет от хозяина к фотографу, соотношение 2.5:1 моментально исчезнет.

Роль относительного в восприятии габаритов изучали Ирвин Рок и Шелдон Эбенгольц (1959). В строго заданных условиях, т.е. в затемненном помещении, иссле-

дователи предъявляли испытуемым светящуюся линию в 3 дюйма длиной, расположенную внутри четырехугольника в 4 дюйма высотой (рис. 7.1.6 А). После чего наблюдателями показывали второй прямоугольник, но уже 12 дюймов в высоту (рис. 7.1.6 В) и просили выставить относительную длину линии, не глядя при этом на исходный образец. Идея опыта состояла в том, что ежели восприятие размера базируется сугубо на фактических (абсолютных) длинах в сцене, испытуемые должны вытянуть линию лишь до 3 дюймов. Ежели оно привязано исключительно к относительному размеру, испытуемые растянут ее до 9 дюймов, дабы сохранить соотношение 3:4, то есть — 9:12. На практике же наблюдатели вытянули линию до 7 дюймов, что примерно по середине между прогнозами по абсолютному размеру и относительному. То есть, очевидно, что относительный размер предмета оказывает определенное влияние на восприятие абсолютного.

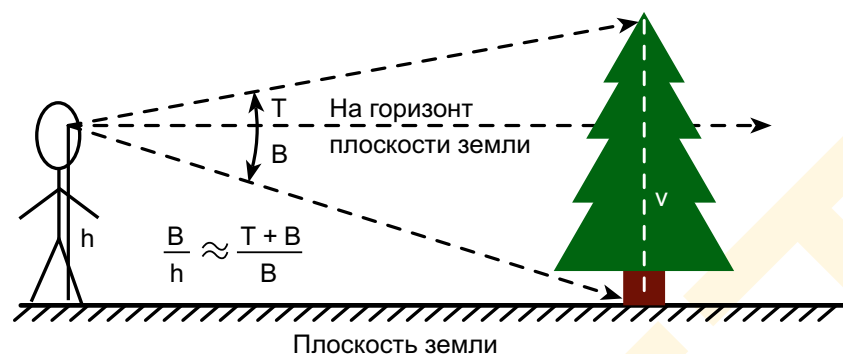
Важно отметить, однако, что относительностью нельзя истолковать все феномены восприятия размера. Подобно текстурной окклюзии она может претендовать лишь на роль шкалирующего фактора, который сам по себе не способен определить абсолютный размер объекта. В нормальных условиях, однако, восприятие габаритов объекта на базе его относительного размера конвергирует с перцептом, базированным на удаленности от объекта, что, в итоге, и дает нормальный уровень константности (Rock & Ebenholtz, 1959). Гипотеза пропорциональности хоть и с трудом, но объясняет и то, почему константность уходит по мере роста расстояния. «С трудом», потому что соотношение размеров постоянно, на каком бы расстоянии от объекта не оказался наблюдатель. Например, соотношение габаритов домов и машин остается неизменным, что с дальнего холма, что через улицу.

Аргумент, который чаще всего звучит как антитеза прямому учету расстояния при восприятии размера единичного объекта, принимает во внимание эффекты окулярной информации о глубине пространства. Мы помним (см. гл. 5), что информация эта относится к состоянию глаз: аккомодации (вариабельности формы хрусталика) и бинокулярной конвергенции (углу между зрительными осями левого и правого глаз, когда они зафиксированы на объекте), но не к оптическим изображениям, стимулирующим их сетчатки. Если наблюдателю стереоскопически предъявить круг на разных углах конвергенции, но при этом так, что у круга этого во всех случаях окажутся одни и те же ретинальные размер и позиция, то перцептивные габариты круга будут радикально меняться от малых (когда оси конвергированы по малому расстоянию) до огромного (когда по большому).

Сказанное часто подают как контраргумент прямым теориям восприятия размера объектов. И сие наглядная демонстрация того, что прямые теории не в силах объяснить восприятие размера во всех возможных условиях наблюдения. Но при этом они могут стать полезными в объяснении многих других феноменов, в которых задействованы сложные оптические взаимоотношения. Судя по всему, зрение использует любую доступную информацию о расстоянии до объекта, чтобы рассчитать его действительный размер. Во многих случаях используются как окулярный, так и оптический источники.

#### *7.1.1.7 Склонение от горизонта*

Еще одним источником информации о размере объекта на плоскости является его положение относительно горизонта этой плоскости. На рис. 7.1.7 показано, что ежели горизонт плоскости виден (или читается из доступной зрительной информации) и длина отвеса от точки наблюдения до плоскости известна, размер объекта на этой плоско-



**Рис. 7.1.7** Склонение от горизонта. Перпендикулярную высоту объекта ( $v$ ), расположенного на плоскости (в данном случае на плоскости земли) можно определить из зрительных углов его вершины ( $T$ ) и основания ( $B$ ) относительно зрительной оси на горизонт и перпендикулярно расстоянию до поверхности ( $h$ ). Когда углы  $T$  и  $B$  относительно малы, отношение плотно аппроксимируется простым отношением:  $v/h = (T + B)/B$ .

сти можно высчитать геометрически (Sedgwick, 1986).<sup>1</sup> Но даже если высота нормали к плоскости неизвестна, можно определить *относительные* высоты объектов, расположенных на этой плоскости.

#### 7.1.1.8 Онтогенез

Ранее было показано, что мы можем вывести действительные размеры объектов из размера ретинальных проекций этих объектов и из удаленности их от наблюдателя. Возникает вопрос: эта способность приобретается через взаимодействие с окружающим миром (например, путем сопряжения зрительных признаков с тактильными данными), или же дана каждому из нас при рождении?

Недавние исследования показали, что константность размера является врожденным признаком. Американцы Алан Слэйтер, Анни Матток и Элизабет Браун (1990) изучали сей вопрос на основе т.н. парадигмы приучения (см. раздел 6.6.1). В адаптирующей фазе опыта каждому малышу по нескольку раз с разных расстояний показывали либо большой, либо маленький кубики. Собственно тест проводили затем либо с тем же самым кубиком, либо с точно таким же, но отличным лишь по размеру. После чего оценивали поведение ребенка. Сеансы стимуляции строились весьма хитроумно: ретинальные габариты двух кубиков в двух тестовых сеансах были идентичными, то есть, крупный кубик предъявлялся с большего расстояния. Новорожденные отдавали явное предпочтение новому кубiku, что устойчиво указывает на их способность видеть отличия в действительном размере предметов.

В других экспериментах использовались несколько иные методики, но которые дали те же самые результаты (Granrud, 1987).

Поскольку в этих опытах детям была доступна информация о глубине пространства, так и осталось неясным, какие из механизмов достаются при рождении. Правда, многое исключала сама экспериментальная ситуация, в частности, вариации проек-

<sup>1</sup> Сей факт плотно соотносится с информацией о глубине, полученной из позиции объекта относительно линии горизонта.

тивного размера (неизменный), текстурную окклюзию (не было фоновой текстуры), относительность размеров (объект был единичным). Но что-то все же оставалось.

Возможно, для младенца наиболее значимый источник информации о глубине пространства — бинокулярная конвергенция. Однако безотносительно к тому, какие источники используются от рождения, их наличие позволяет детям в дальнейшем запоминать прочие признаки, указывающие на действительный размер объектов, согласуя эти признаки с врожденными способностями.

### 7.1.2 Иллюзии размера

Любопытно, что механизмы, отвечающие за константность размера в нормальных условиях зрения, в условиях специфических могут дать определенные виды иллюзий. Сие вполне понятно, поскольку восприятие размера систематически неверидикально, когда систематически неверидикальны перцепты расстояния. Неконстантность габаритов на больших расстояниях — лишь один из примеров.

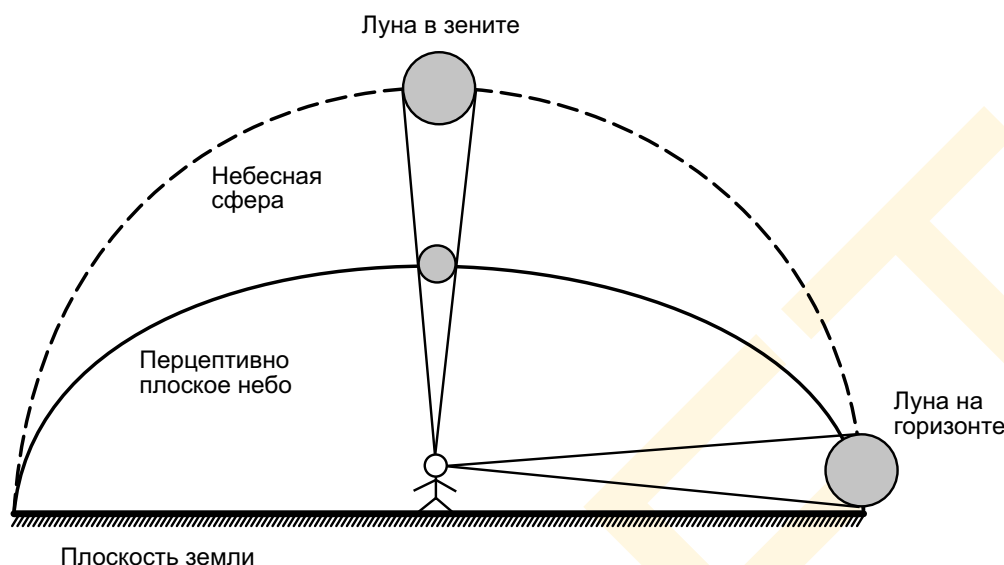
Еще один пример — эймсова комната (см. раздел 5.5.10): геометрия этой комнаты, напомним, такова, что с определенной точки помещение смотрится идеально прямоугольным, но при этом одна сторона помещения фактически намного ближе к наблюдателю, чем другая (см. рис. 5.5.32). Таким образом, имеет место различие между перцептивным расстоянием до двух предметов и расстоянием фактическим. Согласно гипотезе о том, что зрительная система рассчитывает размер объекта из ретинального размера его проекции и перцепта расстояния, ближестоящий персонаж воспринимается крупнее дальнего, даже несмотря на то, что оба примерно одного роста (рис. 5.5.32).

#### 7.1.2.1 Иллюзия Луны

Систематические сбои в восприятии расстояния — это главная составляющая толкования, возможно, старейшей из всех зрительных иллюзий — иллюзии Луны. Лунная иллюзия — это кажущееся укрупнение нашего ночного светила, когда оно над горизонтом, в сравнении с тем, когда в зените. Иллюзия сия эффектна и повсеместна, и не дает покоя ученым с древности. Действительно, многие из нас никогда не задумывались над тем, что «огромная» Луна над горизонтом — это отклонение в восприятии, и полагали, что эффекту этому существует некое физическое объяснение. Разумеется, никому и в голову придти не может, что Луна мельчает, двигаясь по небосводу, но при этом многие полагают, что эффект связан с удалением светила от Земли, или с неким оптическим искажением его размеров атмосферой. Однако ни одно из этих объяснений неверно: фотографии Луны свидетельствуют, что ее размер неизменен по всей траектории. Следовательно, объяснение сего тяжелого нарушения в восприятии строится на том, как мы оцениваем саму зрительную ситуацию.

Согласно наиболее распространенной версии, иллюзия Луны возникает из-за погрешности в восприятии расстояния. Название концепции звучит как «теория видимого расстояния». Изначально ее выдвинул Птолемей, а спустя многие столетия подтвердили Ллойд Кауфман и Ирвин Рок (1962, Rock & Kaufman, 1962), показавшие, что Луна потому выглядит крупнее возле горизонта, что воспринимается более удаленной (рис. 7.1.8). Расстояние от Земли до Луны около 384400 км — это слишком далеко для того, чтобы веридикально оценить абсолютное расстояние до этого объекта (если бы мы смогли достичь константности размера и иллюзия Луны не возникла бы, то планета воспринималась бы нами как объект 3476 км в диаметре!) Вспомним, что





**Рис. 7.1.8** Теория перцептивного расстояния в толковании иллюзии Луны. Светило выглядит крупнее над горизонтом, нежели в зените, потому что воспринимается более удаленным от наблюдателя, т.к. небо мы ощущаем плоским, но не сферическим.

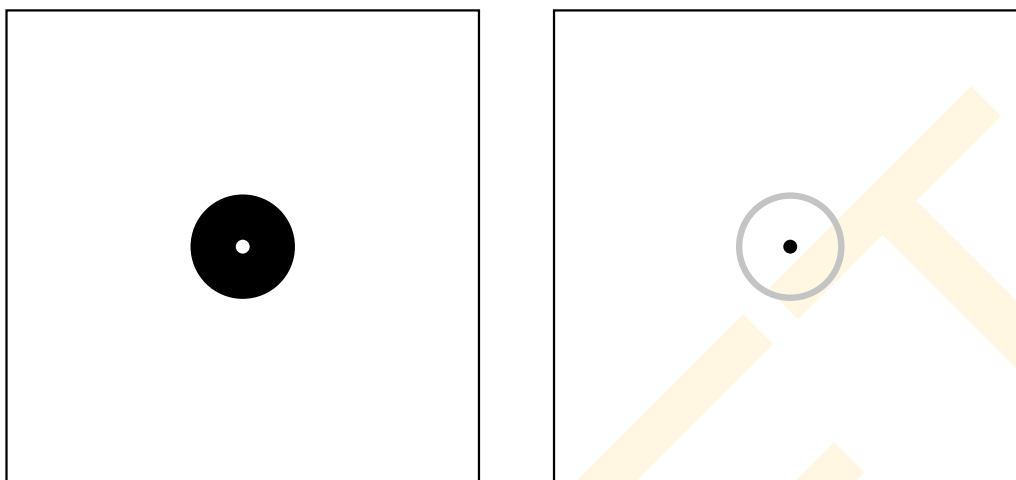
мы непременно должны разобраться с пространственной информацией, поступающей из внешнего мира. Сие и вводит нас в заблуждение.

Ежели исходить из концепции перцептивного расстояния, то одним из факторов неверидикальности в восприятии размеров Луны является пикториальная информация о глубине пространства, извлекаемая зрением из пространственного контекста: когда Луна на горизонте, мы видим ее вместе с протяженным участком земли; когда светило в зените — мы видим его лишь среди звезд и пустого пространства. Теория перцептивного расстояния предполагает, что впечатление большой удаленности, вызванное текстурированной поверхностью земли, понуждает Луну восприниматься нами как более удаленную от горизонта, нежели когда она в зените, поскольку небо мы ощущаем скорее плоским, нежели сферическим. Когда ретиная проекция одного и того же размера толкуется как предмет, расположенный дальше (горизонт), то воспринимается он крупнее, чем когда интерпретируется расположенным ближе (зенит). По этой причине восходящая Луна всегда крупнее Луны в зените.

Тот факт, что две идентичных ретинальных проекции (как, скажем, проекции Луны в двух положениях) воспринимаются как объект разных размеров — это пример т.н. *закона Эммерта*, гласящего, что:

- ⇒ Перцептивный размер объекта, проекция которого на сетчатку постоянна по габаритам, пропорционален перцептивному расстоянию до этого объекта.

Это отношение, выраженное уравнением 7.1, можно эффектно продемонстрировать с помощью зрительных постобrazов. При ярком свете на полминуты — минуту сосредоточим взгляд на белой точке в центре черного круга рис. 7.1.9, что вызовет светлый постобраз круга с фиксированным ретинальным размером. Затем, мы в буквальном смысле можем увидеть, как перцептивные габариты пятна меняются с расстоянием просто за счет того, что мы переносим этот образ на разноудаленные от нас поверхности: вначале давайте посмотрим на дальнюю стену комнаты — постобраз «сел» на стену



**Рис. 7.1.9** Закон Эммерта. Для формирования устойчивого постобраза сосредоточим взгляд на черной точке левого квадрата примерно на 30 с, а затем переведем его на белую стену и сравним перцептивный размер образа с белым кругом в правом квадрате. Пятно будет смотреться ощутимо крупнее на фоне стены, даже при том, что его ретинальный размер ровно такой же.

и смотрится теперь весьма крупным в сравнении со своим прародителем на странице книги, достигая нескольких десятков сантиметров в диаметре (если стенка достаточно далеко от нас);<sup>1</sup> затем посмотрим на квадрат справа в книге и отметим, что теперь пятно воспринимается много меньшим, чем на стене — примерно в исходный размер. Описанное — наглядная демонстрация закона Эммерта, то есть, мы убедились в том, что одно и то же ретинальное изображение дает перцепты разных габаритов при разных перцептивных расстояниях.

Очевидно слабым местом толкования лунной иллюзии на базе концепции перцептивного расстояния является то, что на вопрос: когда Луна ближе — над горизонтом или в зените? — почти все уверенно отвечают, что над горизонтом, тогда как теория прогнозирует обратное. Сей конфликт часто называют *дистанционно-габаритным парадоксом*. Рок и Кауфман объяснили сие очевидное противоречие, предположив, что *осознанно* наблюдатели оценивают расстояние до Луны, исходя из его перцептивного размера, то есть в направлении обратном тому, что происходит *бессознательно*. Если сие верно, то мы можем говорить о двухстадийном процессе, основанном на двух разных вариантах перцепции расстояния — бессознательном и осознанном:

1. Информация о расстоянии, полученная из текстурных градиентов, регистрируется бессознательно, заставляя Луну восприниматься дальше, чем когда она в зените.

2. Когда требуется эксплицитная, осознанная оценка перцептивного расстояния до Луны в двух ее позициях, перцептивные габариты светила в первом и втором положениях берутся как фактор относительного расстояния до нее.<sup>2</sup> В результате, крупная Луна у горизонта оценивается нами как расположенная ближе, чем в зените.

И, хотя озвученная концепция выглядит неуклюже и путано, Кауфману и Року удалось-таки собрать пакет ее доказательств. Четыре наиболее значимых следующие:

<sup>1</sup> Чтобы освежить постобраз рекомендуется немного поморгать.

<sup>2</sup> Относительный размер — это фактор перцептивного расстояния (см. гл. 5).

1. Когда наблюдателей просят взглянуть на ночное безлунное небо и увидеть его как некую поверхность, они оценивают горизонт расположенным дальше зенита. Сие повышает рейтинг версии об уплощенном восприятии небосвода.

2. Когда наблюдатели видят искусственную Луну (размером которой можно управлять оптически) через шторку, закрывающую землю и предметы на ней, классическая иллюзия Луны исчезает. Сходные эффекты возникают, когда наблюдатели оценивают размер светила в полностью затемненной планетарии. И первое, и второе повышает нашу уверенность в том, что именно земля и объекты на ней дают отличия в перцептивных расстояниях до горизонта и до «поверхности» неба в зените.

3. Посредством разных оптических ухищрений исследователи сумели реверсировать зрительный контекст двух лун, проектируя Луну над горизонтом (со всем ее земляным контекстом) вверх, в зенит, а зенитную Луну (с окружающим ее пустым небом) — вниз, на горизонт. Теория перцептивного расстояния прогнозирует реверс иллюзии в этом случае.

4. Кауфман и Рок показали, что наблюдатели всегда докладывают о том, что перцептивно меньшая искусственная Луна, вне зависимости от того у горизонта она или в зените, воспринимается ими как расположенная дальше. Сие аргумент в пользу того, что перцептивный размер Луны управляется перцептивным расстоянием.

Однако несмотря на то, что концепция толкования иллюзии Луны на базе перцептивного расстояния обрела мощную экспериментальную поддержку — это не значит, что ее повсеместно приняли: в последнее время вышло множество публикаций, в которых авторы предлагают совершенно иные объяснения причин лунной иллюзии (Hershenson, 1989). Словом, единственно верное толкование еще предстоит найти, но идея перцептивного расстояния главенствует уже не одно десятилетие.

### 7.1.2.2 Иллюзия Понцо

Иллюзии размера возникают даже тогда, когда наблюдатель, казалось бы, обладает всей полнотой бинокулярной информации, т.е. информации о расстоянии до объектов сцены. Среди классических геометрических иллюзий (см. гл. 1) тому имеются несколько примеров (рис. 1.1.4). Один из них — знаменитая иллюзия Понцо (рис. 7.1.10): нижний горизонтальный отрезок воспринимается короче верхнего, несмотря на их фактическое равенство и на исчерпывающую информацию об их абсолютном расстоянии от точки наблюдения (получена из данных аккомодации и конвергенции). Чем же объяснить появление иллюзии в отсутствие систематических ошибок восприятия глубины пространства?

**Гипотеза бессознательной иллюзии глубины пространства.** Одна из наиболее известных и влиятельных гипотез опирается на идею некорректной интерпретации расстояния, несмотря на отсутствие истинной иллюзии глубины (Gregory, 1966). Концепция гласит, что даже тогда, когда глубина пространства не воспринята эксплицитно, сходящиеся линии активируют механизмы восприятия глубины, то есть провоцируют отклик зрения на пикториальную перспективную информацию (см. раздел 5.5). Мнимого ухода верхнего отрезка за плоскость страницы явно нет, поскольку схождение линий не может в полном объеме «раскошегарить» механизмы восприятия глубины пространства, по крайней мере до такой степени, чтобы их результаты попали в область осознанного зрительного опыта. Линии скорее бессознательно интерпретируются параллельными (как, например, рельсы железнодорожного полотна). В результате, бессознательное восприятие глубины приводит к осознанному перцепту различия в размерах: верхний отрезок оказывается визу-

ально длиннее нижнего, поскольку ближе подходит к сходящимся линиям. Таким образом, некоторые иллюзии размера все-таки могут появиться в процессе обработки информации о глубине, автоматически и неосознанно запускающейся в тех случаях, когда осознанной иллюзии глубины нет.

Однако сколь бы интригующе не звучала в контексте понцовой иллюзии версия «неосознанной глубины», она не очень-то согласуется с фактами. Одна из сложностей в том, что концепция дает некорректные прогнозы: если два тестовых отрезка бессознательно истолкованы расположенными на разных расстояниях от наблюдателя, не должно иметь значения то, как они ориентированы по отношению к наводящим линиям. Фактически же поворот отрезков на  $90^\circ$  разрушает иллюзию — рис. 7.1.11 (Gillam, 1973). Чтобы объяснить сей факт, концепции неосознанной глубины требуется серьезная доработка.

**Гипотеза низкой пространственной частоты.** Существуют однако и другие толкования иллюзии Понцо, которые успешно справляются с эффектом поворота отрезков. Одно из них состоит в том, что зрительная система определяет перцептивные дли-

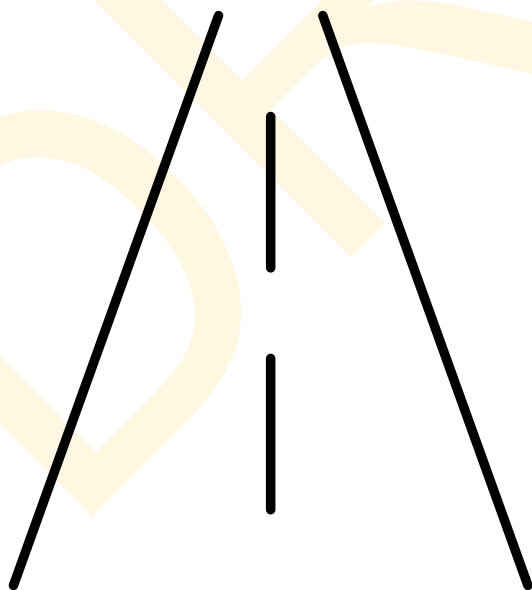


Рис. 7.1.11 Пропадание иллюзии Понцо при повороте отрезков на  $90^\circ$ .

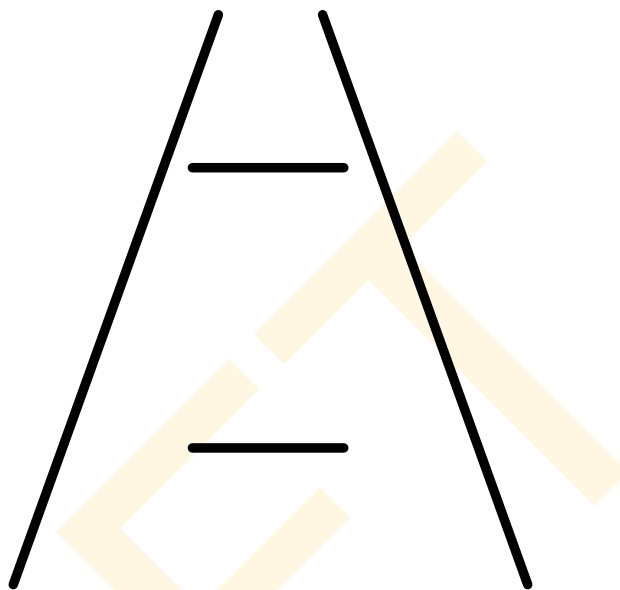
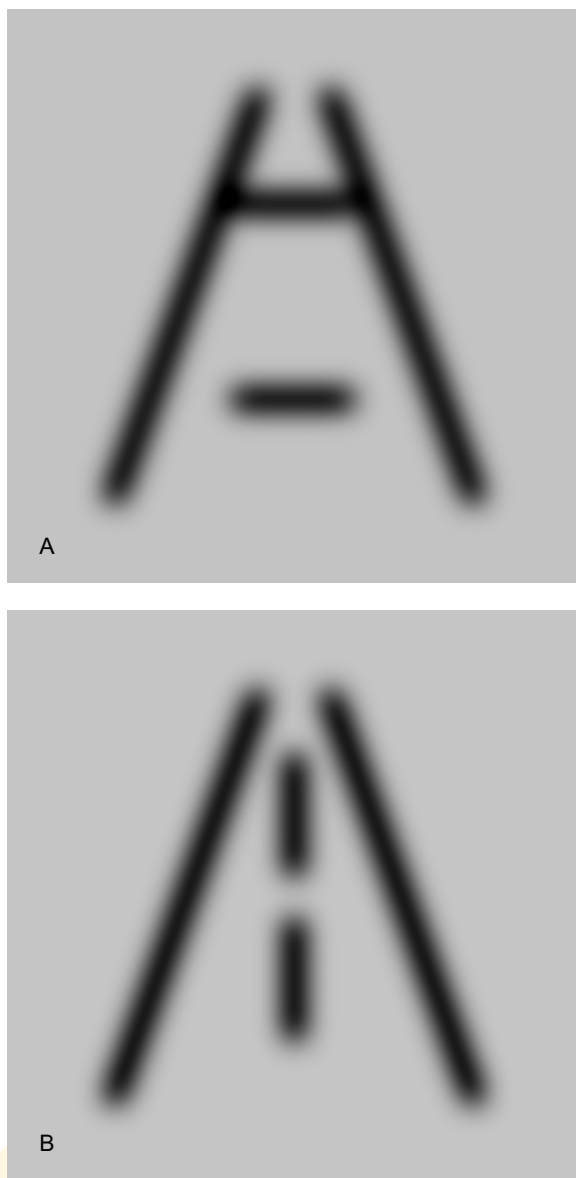


Рис. 7.1.10 Иллюзия Понцо. У горизонтальных отрезков фактически одинаковая длина, но сходящиеся линии заставляют восприниматься верхний отрезок длиннее нижнего.

ны тестовых отрезков, в первую очередь, из пространственно-низкочастотного компонента в стимуляции. На рис. 7.1.12 А дана исключительно низкочастотная составляющая стимула, то есть изображение размыто, и верхний отрезок слился с боковыми линиями — что и сделало его перцептивно длиннее (Ginsberg, 1986). Поскольку с нижним тестовым отрезком этого не произошло, он и выглядит короче верхнего. Версия объясняет, как возникновение самой иллюзии (рис. 7.1.12 А), так и ее пропадание при повороте тестовых отрезков (рис. 7.1.12 В). Но даже такая расшифровка не в силах истолковать все наблюдаемые нами факты. К примеру, когда низкочастотная информация селективно удале-



**Рис. 7.1.12** Понцова иллюзия на низких пространственных частотах. Отфильтровка высоких пространственных частот дает размытый вариант стимуляции, в которой верхний тестовый отрезок становится длиннее фактически за счет того, что сливается с наводящими линиями (А). Если отрезки повернуть на  $90^\circ$  (В), размытие уже не может повлиять на их длину.

объяснение, основанное на том, что размер объекта берется относительно окружающих его объектов (то есть, почти как при пропорциональном толковании константности размера).

Существуют и другие гипотезы, толкующие появление геометрических иллюзий, но ни одна из них сегодня не дает удовлетворительного объяснения тому, почему в

на из стимуляции (рис. 7.1.13), иллюзия Понцо все равно имеет место.

#### **Гипотеза относительного размера.**

Еще одно объяснение понцовой иллюзии звучит в логике отличий в относительном размере объектов. Мы справедливо утверждаем, что верхний отрезок воспринимается длиннее, поскольку он и на самом деле длиннее, если размер каждого из тестовых отрезков выразить в относительном расстоянии между боковыми сходящимися линиями. Иными словами, верхний отрезок пропорционально больше заполняет собой пространство между линиями, нежели нижний. Сие объяснение основано именно на *относительности размера*, то есть на том, насколько верхний объект дотягивается до соседних объектов и насколько нижний. Следовательно, понцова иллюзия относится к тому типу эффектов, которые можно прогнозировать на базе пропорциональности и в контексте константности размера, обсуждавшейся ранее.

#### **7.1.2.3 Иллюзия Эббингауза**

Еще одна классическая геометрическая иллюзия — иллюзия Эббингауза. Суть которой в том, что тестовый круг в обрамлении крупных кругов (рис. 7.1.14) выглядит мельче, чем точно такой же среди мелких. В принципе, мы можем ограничиться дистанционным объяснением, сказав, что круг слева выглядит крупнее, потому что его окружение из мелких кружков, создает перцепт большей удаленности от наблюдателя, чем картинка справа. Однако левый круг смотрится крупнее просто потому, что воспринимается относительно мелких кружков его окружающих, и, аналогично, правый круг выглядит мельче, потому что окружен кругами крупными. Сие очевидное



определенных условиях эти иллюзии либо сходят к минимуму, либо вовсе пропадают. Сие касается классических иллюзий, подобных тем, что возникает при рассматривании рис. 1.1. Существует множество теорий, объясняющих эту эббингаузову иллюзию, но ни одну из них мы не можем назвать удовлетворительной.

#### 7.1.2.4 Оклюзионная иллюзия

Прежде чем закончить разговор о мнимых размерах, стоит сказать пару слов о еще одной группе иллюзий, возникающих, когда объекты зрительно завершаются позади кроющего объекта. В таких ситуациях одновременно возникают две иллюзии, которые влияют на результирующий перцепт в диаметрально противоположных направлениях:

1. Имеет место увеличение габаритов видимой порции частично закрытого предмета (Kanisza, 1979).

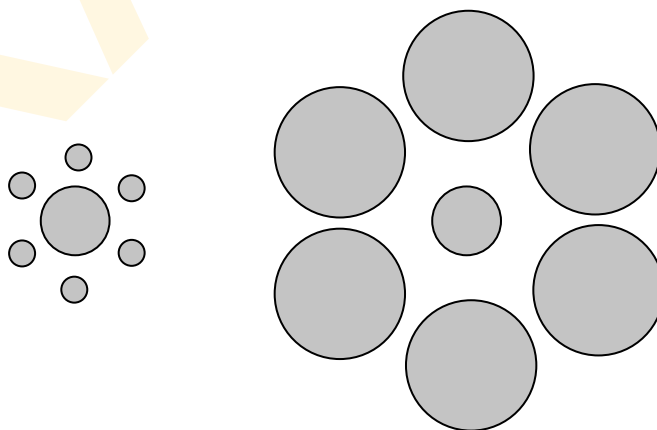
Как показано на рис. 7.1.15 А, видимая часть перцептивно завершенного круга слева воспринимается несколько крупнее, чем физически идентичный ему, но при этом полностью видимый полукруг справа. Вероятно, иллюзия возникает в результате экспансии видимой доли завершенного объекта по кроющей границе, как будто бы зрительная система добавила ему тонкую полосу позади кроющей фигуры.

2. Имеет место перцептивное уменьшение размера невидимой (амодально завершенной) доли кроемого объекта.

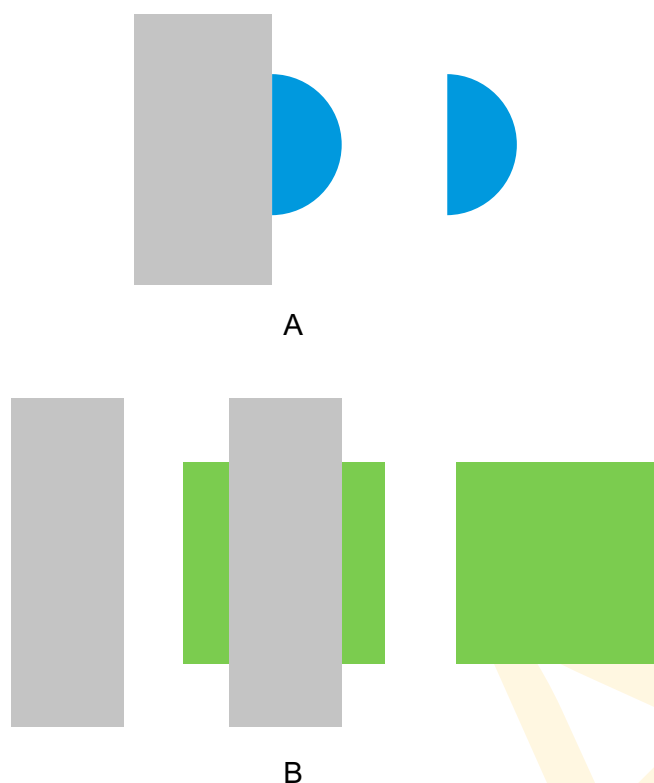
Закрытая часть фигуры на рис. 7.1.15 А не видна, но обратимся к рис. 7.1.15 В, на котором фрагменты кроемой фигуры выходят за границы кроющей полосы. Размер кроемого квадрата физически идентичен незакрытому квадрату спра-



**Рис. 7.1.13** Понцова иллюзия на высоких частотах. Когда пространственно-низкочастотная информация удалена из классического понцова стимула, вопреки прогнозу гипотезы низких пространственных частот, иллюзия все равно появляется.



**Рис. 7.1.14** Иллюзия Эббингауза. У обоих центральных кругов одинаковый физический диаметр, однако диаметр перцептивный у левого тестового круга очевидно больше, поскольку расположен он в обрамлении мелких кругов.



**Рис. 7.1.15** Окклюзионная иллюзия. Несмотря на то, что два полукруга физически идентичны (А), «закрытый» смотрится очевидно крупнее. Также, несмотря на то, что оба квадрата (В) идентичны, закрытый выглядит мельче.

ва, но при этом кроемый квадрат слева смотрится явно уже (Kanizsa, 1979). То есть, амодально завершенная доля кроемого объекта феноменально суживается.

Стоит сказать, что у окклюзионной иллюзии имеется ряд весьма любопытных свойств. Во-первых отметим, что хотя оба эффекта взаимнопротивоположны (рост перцептивного размера видимой доли и его уменьшение у доли закрытой), они не взаимоисключающи. То есть, квадрат на рис. 7.1.15 В будет вероятнее всего выглядеть таким же по толщине как и квадрат справа. Путаницу вносит, однако, тот факт, что хотя закрытая доля квадрата позади четырехугольника перцептивно уже, кроющий прямоугольник сам по себе не воспринимается узким (в чем легко убедиться, сравнив его с физически идентичной полоской справа). Иллюзии размера невозможно истолковать с помощью дистанционных концепций и, следовательно, требуется какое-то иное объяснение. Однако по сей день не существует удовлетворительной версии, способной учесть все аспекты этого феномена. Наиболее подходящее объяснение

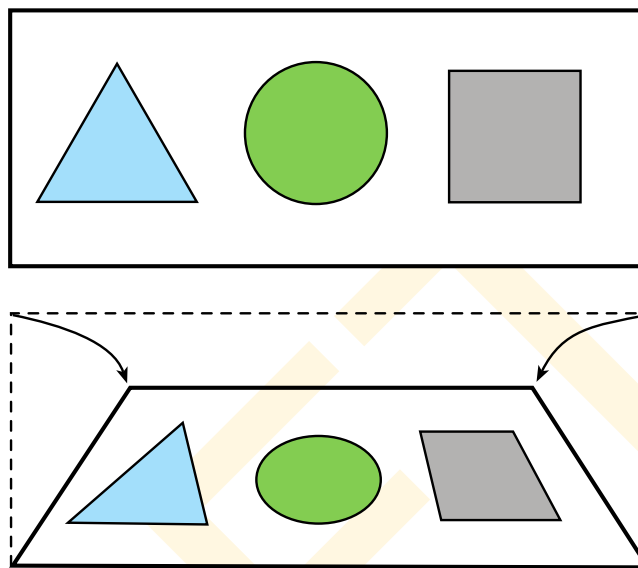
таково, что иллюзии возникают в результате процесса заполнения, частично заливающего закрытую область зрительным ощущением и частично — амодальным знанием (см. раздел 6.4.1).

## 7.2 ФОРМА

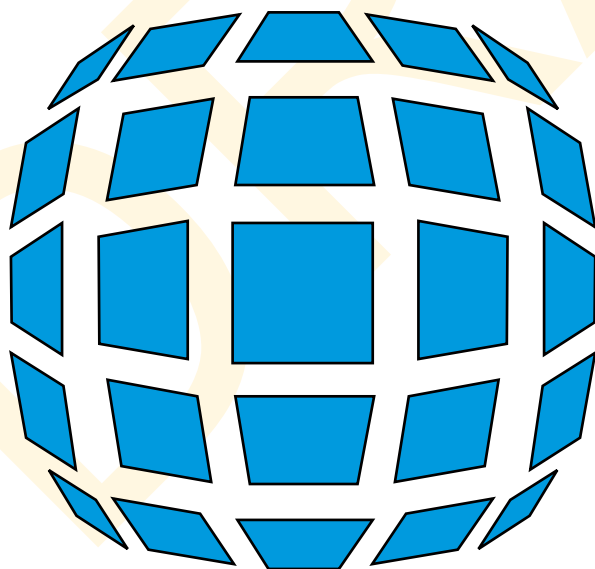
Форма — это одно из наиболее сложных зрительно воспринимаемых свойств объектов. Сложных потому, что представляет собой комбинацию множества разных атрибутов. Действительно, форма объекта имплицитно несет в себе информацию обо всех остальных пространственных свойствах этого объекта: размере, ориентации и позиции. Дело в том, что форма критично зависима от долевого строения объекта и от того, как различные его части соотносятся друг с другом в логике их относительного расположения, относительных ориентаций, относительных размеров и т.д. Характер восприятия формы сложен и загадочен, и по сей день нет общепринятой теории того, что есть форма и как формируется ее перцепт. Углубленно мы рассмотрим этот вопрос в гл. 8, а сейчас поговорим о восприятии формы с несколько упрощенной позиции ее константности.

### 7.2.1 Константность формы

Форма (равно как и размер) — это свойство объектов, которое мы обычно воспринимаем константным, невзирая при этом на изменчивость зрительной перспективы. Отправной точкой концепции восприятия формы является, разумеется, ретинальная проекция объекта. Когда некий *двухмерный* объект располагается на фронтальной плоскости, форма его ретинальной проекции точно повторяет форму реального объекта. Но во всех остальных случаях (либо же когда двухмерные объекты рассматриваются под наклоном — рис. 7.2.1) между формой проекции объекта на сетчатку и реальной формой самого объекта имеются существенные отличия. Следовательно, тот факт, что мы веридикально воспринимаем форму предмета, разглядывая его в разных ракурсах, принято называть *константностью восприятия формы* или коротко — *константностью формы*. Как и все остальные виды константности, константность формы является перцептивным эффектом, требующим развернутого объяснения.



**Рис. 7.2.1.** Константность формы. При наличии добротной информации о глубине пространства даже при наклоне, перцептивная форма предметов в очень широком диапазоне условий просмотра остается той же, что и во фронтальной плоскости.



**Рис. 7.2.2.** Квадрат в разных перспективах. Обыкновенный квадрат может давать множество проективных форм: трапецию, параллелограм, ромб и всякие неправильные четырехгранники.

#### 7.2.1.1 Перспективные изменения

Основная нагрузка на перцептивную константность формы ложится тогда, когда наблюдатель меняет свое местоположение и рассматривает один и тот же объект в разных ракурсах. В большинстве случаев изменения в геометрии видности приводят к изменениям в форме ретинальной проекции. Если объект обладает плоской планарной формой или сделан из тонкой проволоки, характер его проекции можно описать математически как простой мэппинг трехмерного пространства в двухмерное, т.к. все части этого объекта вид-



**Рис. 7.2.3** Три проекции трехмерного объекта как иллюстрация сложных изменений в форме ретиальной проекции при разных ракурсах просмотра. Действительно, у данного объекта мы не найдем ни одной пары идентичных проекций.

ны (рис. 7.2.2 на примере квадрата продемонстрирует нам, как проективные изменения меняют форму ретиального изображения объекта). Но если предмет целен и трехмерен, изменение ракурса наблюдения меняет форму его ретиальной проекции радикально, поскольку новые участки поверхности становятся видимыми, а другие — исчезают. На рис. 7.2.3 даны несколько проекций цельного объекта, демонстрирующее сложные самоокклюзии при изменении ракурса просмотра.

Какими же должны быть эти изменения, чтобы повлиять на перцептивную константность формы? Во-первых, давайте рассмотрим отдельный случай, когда объекты почти не кроют сами себя, т.е. планарные и проволочные фигурки. В этих ситуациях форму можно (как минимум, в принципе) реконструировать с высокой точностью, если расстояние до каждой точки объекта известно. Мы не станем детально расписывать этот процесс, потому что интуитивно понять сказанное нетрудно. Вообразим лишь, что каждая ретиальная точка дает проекцию *из* глаза по направлению, в котором шла проекция на сетчатку и на перцептивное расстояние от глаза до объекта.<sup>1</sup> Если перцептивное расстояние до каждой точки предмета веридикально, и если все точки видны, результирующее множество точек воспроизведет реальный объект. Если информация о глубине извлечена из абсолютных источников (например, аккомодации и/или конвергенции) — и форма, и размер объекта будут полностью реконструированы. Если же относительное удаление известно из *количественных* источников (например, бинокулярной диспаратности, параллакса движения или прочих метрических поставщиков перспективной информации), форма будет реконструируема, но размер — нет, поскольку объект может находиться на любом абсолютном расстоянии. Если в распоряжении имеется *качественная* информация о глубине, (например, полученная путем интерпретации краев), то ни форму, ни размер невозможно однозначно восстановить без дополнительных предположений. Но при этом определенная степень константности формы все еще остается.

<sup>1</sup> Это та самая информация, которая представлена в марровом 2.5D-скетче.

### 7.2.1.2 Двухмерные фигуры

Как же сей анализ сопоставим с фактической перцептивной константностью формы в реальных условиях? Подавляющее большинство опытов константности формы выполнялись с двухмерными фигурами определенного наклона. Когда в наличии имела точная информация о глубине, константность формы была сильна (Lapin&Preble, 1975; Thouless, 1931).<sup>1</sup> Но слабела по мере наклона фигуры, даже при бинокулярном рассматривании (Massaro, 1973; Thouless, 1931). При очень сильных наклонах наблюдатель начинал воспринимать форму как явный компромисс между фактической дистальной стимуляцией и ретиальной проекцией. Наше зрение при этом демонстрирует устойчивую тенденцию к восприятию симметричных фигур квадратами и кругами, но не овалами и трапециями (King, Meyer, Tangney & Biederman, 1976).

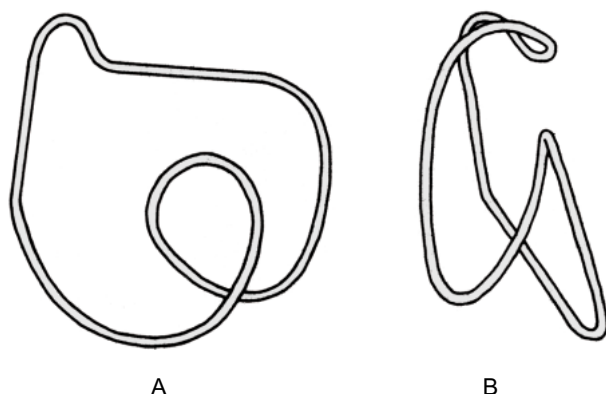
На больших просмотровых расстояниях, на которых восприятие глубины неизбежно ослаблено, уровень константности формы определяется в основном пикториальной информацией о глубине. В отсутствие иных факторов глубины пространства имеет место четкая тенденция видеть форму объекта как если бы он был расположен во фронтальной плоскости (Gogel, 1965, 1978). Однако зрение склонно воспринимать фигуры как более симметричные, то есть согласно их ретиальным проекциям. То есть, константность формы по-разному наклоненных кругов достаточно высока, даже на очень больших расстояниях: то есть, несмотря на то, что проекции кругов на сетчатку — это овалы, они все равно воспринимаются как круги, но наклоненные. Так происходит потому, что, во-первых, круги симметричнее овалов, а во-вторых, потому что круги встречаются в реальности намного чаще.

С другой стороны, константность формы крайне слаба в отношении фигур, чьи проекции должны восприниматься более симметричными и/или знакомыми, чем они есть на самом деле, особенно в отсутствие добротной информации о глубине пространства. К примеру, на больших расстояниях, либо при монокулярном просмотре, слегка наклоненный овал воспримется скорее как круг с иным наклоном. Эффект особо выражен, когда наклон овала таков, что на сетчатку он проецируется как круг. В этом случае тенденции воспринимать форму как симметричную и как лежащую на фронтальной плоскости конвергируют, что ведет к инвариантности перцепта круга на плоскости, но не наклоненного овала. Сие иной пример, в котором процессы, обычно дающие константность формы, напротив — ослабляют эту константность; более того — дают *иллюзию* кругов, когда на самом деле в сцене овалы.

Тенденция воспринимать двухмерные фигуры на больших расстояниях как более симметричные, то есть, согласно их ретиальным проекциям — это закономерный парафраз гештальт-принципа *прягнанц*, утверждающего, что перцепт будет настолько «добротным», насколько позволяют превалирующие условия. Правда, мы до сих пор толком не можем сказать, что делает одну форму «добротнее» или «проще» другой, но симметрия, безусловно, среди наиболее значимых факторов. Плюс к тому, гештальт-концепт *прягнанц* достаточно сложен и требует углубленного анализа. Мы в вернемся к этому вопросу в разделе 8.3.1.

<sup>1</sup> Имеет место типичная тенденция к гипоконстантности (т.е. когда на перцептивную форму предмета влияет некий компромисс между формой фактической и формой двухмерной проекции на сетчатку), но чаще мы воспринимаем форму двухмерных фигур с вполне приемлемой точностью.





**Рис. 7.2.4.** Проволочные фигурки в эксперименте по константности формы. Испытуемые, которым их предъявляли, вначале в одном ракурсе, а потом в ракурсе тестовом, хуже справлялись с поставленной задачей, чем те наблюдатели, которым стимул предъявлялся в одном и том же ракурсе в обоих случаях.

По материалам Rock, DiVita & Barbeito, 1981.

бротной информации о глубине (полученной бинокулярно), по проволочным объектам мы выдаем на удивление слабую константность формы (Rock & DiVita, 1987; Rock, DiVita & Barbieto, 1981).

**Опыт Рока — ДиВиты.** Ирвин Рок и Джозеф ДиВита (1987) предъявляли своим испытуемым серию проволочных фигурок, наподобие тех, что показаны на рис. 7.2.4 А. Затем проверялась память на эти формы в условиях, в которых объект рассматривался либо с той же самой точки (т.е. имел ту же ретинальную проекцию), либо с другой (рис. 7.2.4 В). Когда предъявлялся тестовый объект, распознавание колебалось в пределах от 70 — 75% до 39%. При полной константности в этой ситуации, распознавание было одинаковым в обоих ракурсах.

Рок и ДиВита (1987) решили, что результаты вышли такими потому, что восприятие наблюдателями формы (и память на нее) строго зависимо от качественных изменений в форме ретинальной проекции. Правда, сие вовсе не значит, что перцептивная глубина не играет роли, но лишь то, что перцепт проволочного предмета может жестко зависеть от его ретинальной проекции. Словом, результаты исследования показали, что константность формы таких фигур далеко не столь крепка, как у предметов повседневных.

**Опыт Рока — Пинны — Косты.** Велик соблазн подумать, что результаты опыта Рока — ДиВиты — плод искусственности стимула. Но с непрозрачными трехмерными объектами ситуация даже еще хуже, поскольку ближние к наблюдателю поверхности фактически закрывают другие поверхности этого же объекта. К примеру, на рис. 7.2.5 показан один и тот же глиняный объект в двух разных ракурсах. Выглядит он, явно по-разному и очевидно, что его ретинальные проекции сильно отличны друг от друга. В опытах, проведенных Роком, Пинной и Костой с применением таких фигур, догадка сия полностью подтвердилась — было показано, что константность формы столь же слаба в отношении цельных объектов, как и в отношении проволочных. Итак, очевид-

### 7.2.1.3 Трехмерные объекты

Константность формы проволочных объектов, рассматриваемых бинокулярно с близких расстояний, теоретически должна (также как и у двухмерных фигур) быть довольно высокой. Как было показано ранее, ежели расстояние и направление от наблюдателя к каждой точке предмета известны, то весь проволочный предмет можно реконструировать целиком и с высокой точностью. Однако фактически наше восприятие формы трехмерных проволочных объектов идет труднее, нежели объектов двухмерных. Рок показал, что даже на очень близких расстояниях, то есть при наличии до-



**Рис. 7.2.5** Трехмерный стимул в опыте по исследованию константности формы. В эксперименте Рока — Пинны — Косты (1987) использовались бесформенные глиняные «лепешки». Результаты продемонстрировали весьма низкий уровень константности формы при смене ракурса.

По материалам Rock & DiVita, 1987.

но, что для прояснения взаимоотношений между константностью формы трехмерных проволочных предметов и предметов цельных требуются дальнейшие исследования.

Наконец, нужно пару слов сказать о константности формы трехмерных объектов на больших расстояниях, когда зрению доступна лишь пикториальная информация о глубине пространства. Судя по всему, константность должна слабеть, поскольку включаются такие значимые источники информации о глубине как бинокулярная диспаратность и конвергенция. Аморфные фигуры, в частности, проволочные образцы Рока — ДиВиты, воспринимаются неверидикально при наличии одной лишь пикториальной информации и, соответственно, константность формы слабеет.

Итак, несмотря на все доводы в пользу того, что константность формы слаба, повседневный опыт свидетельствует о том, что на самом деле она достаточно крепка даже в отношении сильно удаленных объектов. И смотрим мы обычно на предметы под разными углами и в разной перспективе, но распознаем их при этом вполне устойчиво. На рис. 7.2.6 дано убедительное доказательство сего тезиса. Почему же данные Рока — ДиВиты свидетельствуют об обратном?

Тому есть несколько причин, и одна из них в том, что наша интуиция в отношении константности формы объекта работает обычно в ситуациях, в которых мы воспринимаем один и тот же объект в разных ракурсах, плавно перемещаясь с одной позиции на другую. В таких условиях перспективные изменения градуальны и легко толкуются в контексте сохранения формы, особенно когда объект наблюдаем нами непрерывно. Таким образом, сии «нормальные» условия принципиально отличны от условий эксперимента Рока — ДиВиты, в котором два ракурса предъявлялись дискретно с большим перерывом во времени между сеансами.

Вторая причина в том, что большую часть времени наше восприятие формы объекта плотно коррелирует с идентичностью. То есть, мы можем быть уверены, что константность формы возникает просто потому, что мы способны распознать один и тот же объект при разных ракурсах просмотра и разными путями. К примеру, мы легко узнаём лицо знакомого человека как анфас, так и в профиль. Если мы опознали его



**Рис. 7.2.6** Несмотря на разные ракурсы, объект устойчиво распознается как один и тот же измерительный прибор — фотоэкспонетр. Сей факт плохо стыкуется с результатами Рока — ДиВиты, показавшими низкий уровень константности при смене ракурса.

как одно лицо, то и его перцептивная форма будет той же. Но окажись перед нами незнакомец или неизвестный ранее предмет — не факт, что мы столь же успешно справились с этой задачей.

Третья причина того, что константность формы сильна в повседневности, состоит в том, что у большинства предметов, нас окружающих, имеются признаки, выделяющие отдельные специфические фрагменты этих предметов (например, «передняя часть» и «задняя часть»), тогда как у роковской проволоки и глиняных «лепешек» таких признаков нет. Плюс к тому, наличие осей симметрии или продолжения позволяет порой воспринимать форму объекта относительно этих осей, позволяя ей оставаться константной в разных ракурсах. Однако аморфные, асимметричные объекты не имеют ни осей, ни плоскостей симметрии и поэтому их форма может восприниматься по-разному с разных точек зрения. Мы продолжим эту мысль в гл. 8.

Наконец, важно отметить, что описанные гипотезы отнюдь не исключают друг друга и могут стать верными одновременно. Хотя по-прежнему неясно, как именно зрительная система преодолевает перспективные отличия в восприятии трехмерных форм, но нет никаких сомнений в ее способности к этому. В противном случае мы бы с двух разных точек зрения видели бы один и тот же объект как два разных. Но такое чрезвычайно редко случается в нашей жизни.

#### 7.2.1.4 Онтогенез

Константностью восприятия формы предметов обладают не только взрослые: имеются данные, что она является врожденным признаком. Алан Слэйтер и Виктория Морисон (1985) обнаружили устойчивое восприятие планарных фигур вне зависимости от их наклона младенцами нескольких дней от роду. Исследователи адаптировали новорожденных к стимулу определенной формы (квадрат или трапеция), предъявляя

затем стимулы тестовые с разным наклоном в глубину. После чего наблюдали за малышами, присматриваясь к предпочтениям ими старых или новых форм. Каждый младенец дольше глядел на предмет новой формы, убеждая ученых в том, что ее константность наличествует при рождении. Однако, хотя опыты Слэйтера и Морисон убедительно демонстрируют врожденный характер константности восприятия формы, имеются кое-какие сведения о том, что в достижении этой константности младенцы пользуются информацией сугубо о глубине пространства. Наиболее вероятный кандидат в источники — бинокулярная конвергенция. Однако гипотеза сия пока еще остается неподтвержденной экспериментально.

### 7.2.2 Мнимые формы

Способность нашего зрения автоматически интерпретировать глубину пространства из двухмерной структуры ретинальных изображений ведет порой к некоторым любопытным иллюзиям в восприятии формы объектов.

#### 7.2.2.1 Иллюзия Эймса

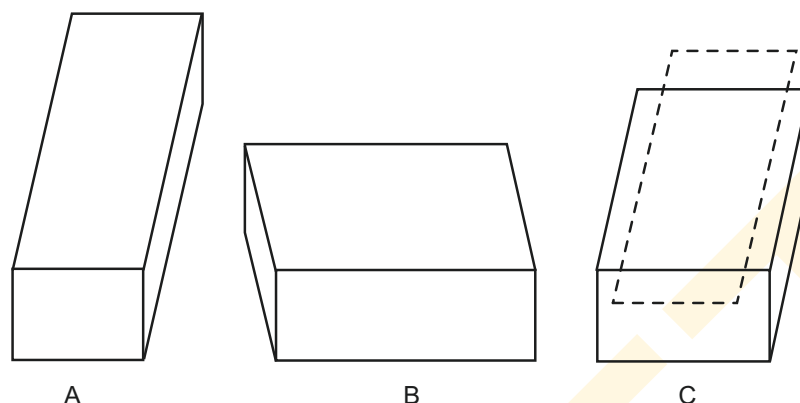
Мы уже говорили о том, что ежели овал рассматривать монокулярно и/или издалека, т.е., когда добротной информации о глубине пространства нет, мы, скорее всего, воспримем его как круг, наклоненный вглубь. Такого рода иллюзии исследовал выдающийся американский психолог Эдельберт Эймс (1951). Схема его опытов была аналогична схеме инициации иллюзии «овал — круг»: когда в темной комнате на фронтальной плоскости монокулярно рассматривалась светящаяся трапеция, она воспринималась при этом как прямоугольник определенного наклона. Сие пример иллюзии формы, возникшей вследствие тех же эвристических предположений, что в нормальных условиях зрения дают как раз константность восприятия, но при этом сделанных в неопределенных условиях и в отсутствие добротной информации о глубине пространства.

На иллюзии, при которой трапеция воспринимается прямоугольником, Эймс построил свою знаменитую комнату. Как мы уже говорили в разделе 5.5.10 эймсова комната — это трапециевидное помещение, которое воспринимается прямоугольным, когда его разглядывают с определенной точки (см. рис. 5.5.32). Одна сторона задней стенки много выше другой, но при этом видится она прямоугольной, поскольку высокая сторона еще и дальше от наблюдателя. Собственно сие и вызывает трехмерную иллюзию формы. Отметим, что эймсова комната — это весьма интересный пример тенденции нашего зрения воспринимать форму объектов (в данном случае самую комнату) согласно простейшему и наиболее типичному сценарию.

#### 7.2.2.2 Иллюзия Шепарда

Еще одна весьма мощная иллюзия формы, относимая нами к восприятию глубины пространства и детально исследованная Роджером Шепардом (Стэндфордский университет, 1981), показана на рис. 7.2.7. Чтобы понять ее, нужно ответить, какой из объектов — слева или справа от центрального — дает двухмерную проекцию той же формы, что и верх объекта в центре? (Желательно попробовать самостоятельно ответить на этот вопрос, а потом продолжить чтение.) Практически все без исключения говорят, что это предмет справа, но фактически — это объект слева! Удивительно, но пунктир, положенный на правый брусок, демонстрирует реальную форму «крышки» среднего бруска, если повернуть его на 90°.





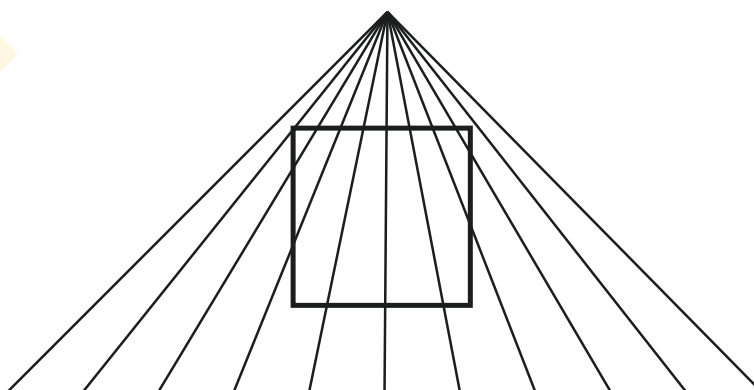
**Рис. 7.2.7** Шепардова иллюзия. Параллелограм, образующий верхнюю поверхность объекта В, очевидно более походит на вершину объекта С, но при этом его ретинальная форма идентична верхней поверхности объекта А, повернутой на  $90^\circ$  (пунктирная линия — это контур верхней плоскости объекта В).

По материалам Shepard, 1981.

Сия весьма впечатляющая иллюзия — плод работы тех же механизмов, что в норме дают константность формы на базе пикториальной информации. Забавно, что двухмерные «крышки» брусков в контексте поставленной задачи воспринимаются неверно, даже если мы приложим усилие: мы упорно зрим их как верх трехмерных объектов. Дело в том, что обработка информации о глубине идет более или менее автоматически, и именно на ее основе мы делаем свой перцептивный выбор. Иллюзия же образуется из-за того, что перцепт этот не нужен при решении поставленной задачи. Но если попросить нас выбрать предмет, у которого форма «крышки» более всего схожа с таковой у центральной фигуры в трехмерной логике — наш ответ будет верным.

### 7.2.2.3 Иллюзия трапеции

Еще одна иллюзия, которую мы можем отнести на счет механизмов интерпретации глубины пространства и константности формы дана на рис. 7.2.8: фактически, это вариант понцовой иллюзии с рис. 7.1.10: четырехугольник на самом деле является идеальным квадратом, но явно склонен восприниматься трапецией основанием кверху. Объяснение эффекта станется ровно тем же, что и у классической иллюзии Понцо: если сходящиеся линии воспринимаются как параллельные на плоскости земли, то верхняя сторона будет казаться длиннее нижней. Согласно этой гипотезе иллюзия особенно выражена, когда линии воспринимаются как задающие плоскость земли,



**Рис. 7.2.8** Мнимая форма. Фигура в центре — это идеальный квадрат, но за счет эффекта сходящихся линий выглядит трапецией основанием кверху.



а четырехугольник как бы лежит на этой плоскости. Если четырехугольник выглядит стоящим, т.е. перпендикулярным плоскости земли, он больше походит на истинный квадрат.

### 7.3 ОРИЕНТАЦИЯ

Следующим зримым свойством объектов внешнего мира является их ориентация относительно других объектов. Мы хорошо знаем, что перцепт ориентации остается веридикальным в постоянно меняющихся условиях: отвесные линии и края мы всегда воспринимаем как вертикальные, а параллельные горизонту — как горизонтальные. В эволюционной логике сей факт исключительно важен, потому что влияние силы тяжести на выживание всеобщно и огромно: вертикальные предметы склонны к относительно устойчивому равновесию, тогда как наклоненные — очевидно, падающие.

Действительно, нашу способность отличать строго вертикальные предметы от наклоненных трудно переоценить. Так чем же объясняется наша способность воспринимать ориентацию предметов из зрительной информации?

#### 7.3.1 Константность ориентации

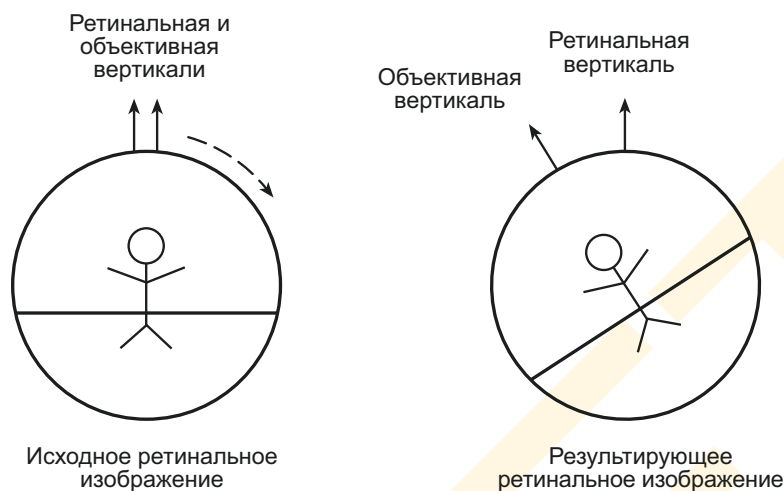
Отправной точкой наших рассуждений вновь явится изображение на сетчатке. В 4-й главе мы говорили о том, что ретинальная ориентация фиксируется корковыми клетками V1, селективно настроенными на линии и края (или локальные пространственные частоты) различных направлений. Возможно, пульсация этих нейронов реализует восприятие положения объектов через ориентацию их составляющих, т.е. линий и краев.

К сожалению, пульсация нейронов V1, чувствительных к ориентации, будет коррелировать с ориентацией перцептивной только тогда, когда голова наблюдателя расположена строго вертикально, и когда линии и края лежат на фронтальной плоскости. Проблема в том, что ретинальная ориентация линий и краев определяется не только ориентацией объекта в реальной сцене, но также положением головы и ориентацией линий глубь пространства. Чтобы упростить разговор, мы ограничимся линиями, расположенными на фронтальной плоскости, а себя оставим зависимыми только от положения головы.

Итак, наша голова на удивление подвижна: мы вращаем ею, оказавшись в тупике, когда заглядываем за угол, когда в задумчивости опираемся головой на руку или когда ложимся спать. При этом, если повернуть голову, скажем, на  $30^\circ$  по часовой стрелке, ретинальные проекции объектов поля зрения повернутся на  $30^\circ$  против нее (см. рис. 7.3.1), но при этом мы не почувствуем какого-либо изменения в ориентации этих объектов: в нашем восприятии они сохраняют свое положение! Как же так получается?

Во-первых, отметим, что сие были примеры *константности ориентации*: мы склонны веридикально воспринимать гравитационное положение объектов несмотря на изменения в положении нашей головы. Если бы константности не было и восприятие ориентации определялось сугубо ориентацией ретинальной, мир вокруг шатался бы из стороны в сторону, ибо голова, наша, вспомним — всё время в движении. И сие оказалось бы перцептивно неприемлемым по несчетному числу причин. Однако на деле зрительная система спокойно определяет ориентацию объектов в реальных сценах, учитывая как положение головы, так и положение ретинальной проекции объекта.

В целом, отношения между ориентацией объекта в реальной сцене ( $O_{\text{объект}}$ ), ориентацией его ретинальной проекции относительно длинника головы ( $O_{\text{проекция}}$ ) и ориентаци-



**Рис. 7.3.1** Константность ориентации. Перцептивная ориентация объектов сцены не меняется при наклоне головы, даже несмотря на то, что ретинальная проекция поворачивается в противоположную сторону.

ей головы наблюдателя относительно вектора силы тяжести — гравитационной вертикали ( $O_{\text{голова}}$ ) можно выразить уравнением:

$$O_{\text{объект}} = O_{\text{проекция}} + O_{\text{голова}} \quad (7.2)$$

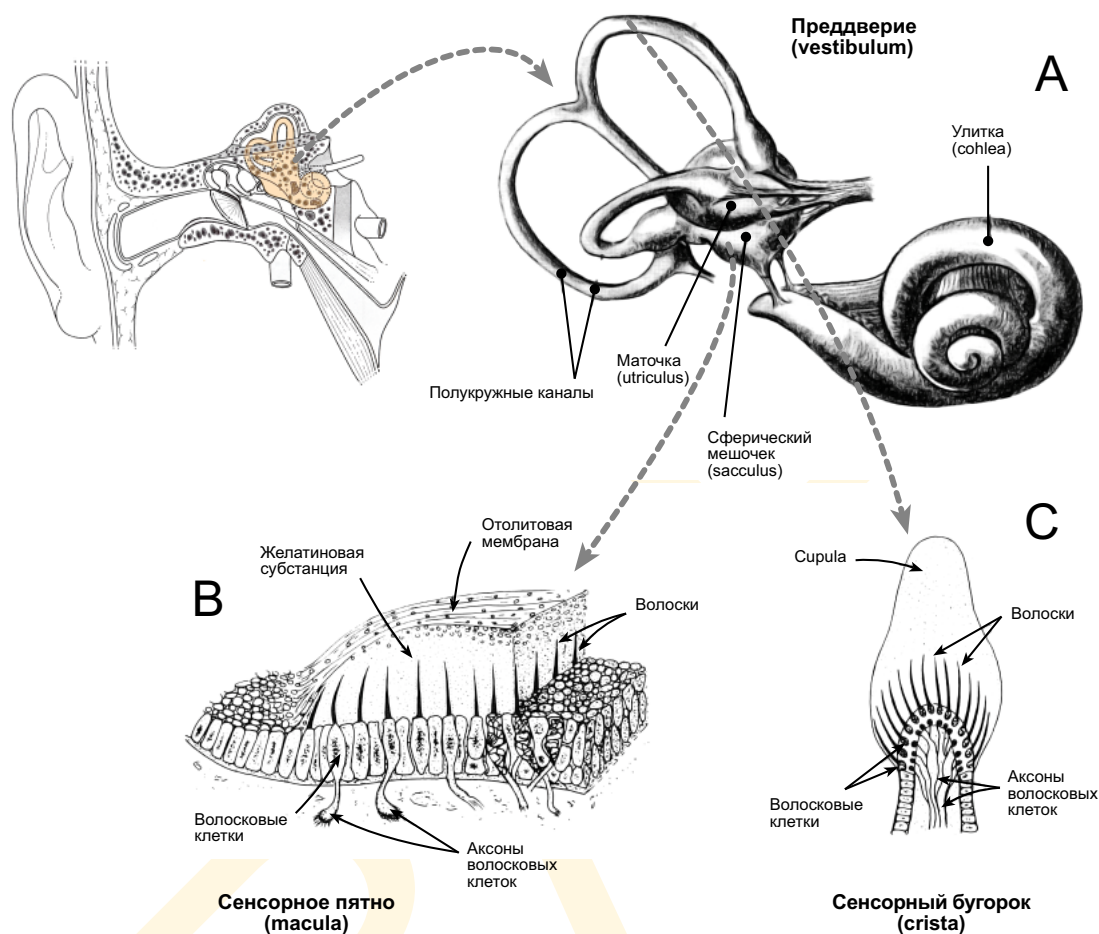
Сие значит, что ориентацию реального объекта относительно вектора центра тяжести ( $O_{\text{объект}}$ ) можно рассчитать, прибавив угол между длинником головы и вектором силы тяжести ( $O_{\text{голова}}$ ) к углу между проекцией объекта на сетчатку и длинником головы ( $O_{\text{проекция}}$ ). Поскольку ориентацию ретинальной проекции относительно головы можно вывести из выходного сигнала пространственно ориентированных клеток — детекторов краев и линий, зрительной системе остается лишь определить положение головы относительно гравитационной вертикали.

Главным источником информации о гравитационных вертикалях и положении головы является т.н. *проприоцептивная система* — биологические структуры, отвечающие, в том числе, за чувство вертикали и равновесие организма внутри поля тяготения Земли.

### 7.3.1.1 Вестибулярный аппарат

За обеспечение пространственного равновесия нашего тела отвечает т.н. *вестибулярный аппарат* — орган, расположенный в т.н. *внутреннем ухе*. Внутреннее ухо расположено внутри сложного комплекса сообщающихся между собой полостей височной кости, образующих т.н. костный лабиринт. Строение вестибулярного аппарата дано на рис. 7.3.2 А: орган состоит из трех соединенных между собой трубочек (т.н. *полукружных каналов*), заполненных жидкостью — *эндолимфой*, и мешочков (т.н. *утрикула* и *саккула*), тоже заполненных эндолимфой. Полукружные каналы, утрикула и саккула посылают в мозг информацию о пространственном положении головы, но при этом информацию разного характера.

Итак, мы говорим, что вестибулярный выходной сигнал способен дать констант-



**Рис. 7.3.2** Вестибулярный аппарат. Внутренне ухо (А) состоит из трех органов, реагирующих на изменение положения головы: маточку (utricle), сферический мешочек (sacculus) и полукружные каналы. Функциональным элементом рецепторов утрикулы и саккулы (В) являются волосковые клетки, «впаянные» в желатиновую массу, покрытую слоем отолинов (твердых частиц). Функциональным элементом рецепторов полукружных каналов (С) являются клетки, «впаянные» желеподобную субстанцию, движущуюся вместе с эндолимфой.

ность ориентации, если объединен с информацией о проекции объекта на сетчатку. Соответствующая комбинация дана в уравнении 7.2.

Утрикула и саккула поставляют в центр информацию о положении головы и тела относительно вертикали. Как уже было сказано, оба представляют собой небольшие капсулы, заполненные жидкостью и выстланные изнутри волосковыми клетками. Волоски эти проникают в некую желатиновую массу, к поверхности которой прилеплено множество плотных частичек. Сила тяжести действует на эти частички, и поэтому, когда мы наклоняем голову, волоски отклоняются на угол, величина которого определяется углом наклона головы. Наклон волосков, в свою очередь, вызывает пульсацию клеток, которая и поставляет в мозг информацию о наклоне наших туловища и головы.

Полукружные каналы при наклоне головы тоже дают отклик, но их основной сиг-

нал возникает при смене положения собственно головы, но не гравитационной ориентации тела. Каждый канал представляет собой трубочку, заполненную эндолимфой с единичным гибким рецепторным органом, прикрепленным к стенке. Когда голова поворачивается, трубочка смещается по отношению к жидкости, сгибая рецептор. Изгиб ведет к пульсации нервных клеток. Когда голова в покое (даже если и не вертикально) относительное движение жидкости внутри каналов прекращается и сигнал в мозг не поступает. То есть, полукружные каналы кодируют информацию о динамике положения головы, но не о ее статической ориентации.

Если вестибулярный аппарат работает нормально и его выходной сигнал четко согласуется с ретинальной ориентацией, константность положения будет близка к идеальной. Но на деле, как правило, всё несколько иначе. К примеру, когда голова наклонена слишком сильно, мы склонны воспринимать вертикаль как слегка наклоненную в направлении головы. Наблюдатели, наклонившие голову на  $90^\circ$  в полной темноте видят светящийся стержень, отклоненный от гравитационной вертикали на  $8^\circ$  в сторону наклона их головы — как строго вертикальный, а строго вертикальный — как отклоненный на  $8^\circ$  в противоположном направлении (Aubert, 1861). Впрочем, сие ослабление полной константности ориентации легко объяснить, если предположить, что наклон головы недооценивается на  $8^\circ$ , то есть воспринимается как  $82^\circ$  вместо  $90^\circ$ . Чем и объясняется то, почему очень сильный наклон головы нарушает константность ориентации как минимум в темноте.

Если в нормально освещенном помещении мы наклоним голову на  $90^\circ$ , константность ориентации сохранится, в чем легко убедиться на личном опыте. Может быть мы затруднимся сказать, какие объекты в сцене *строго* вертикальны, но самая сцена в целом будет выглядеть все же вертикальной, нежели отклоненной в противоположном направлении (как то прогнозирует аубертов опыт со светящимся стержнем в темной комнате). Дело в том, что зрительная система не полностью полагается на вестибулярную информацию, но лишь присовокупляет ее к подходящей зрительной. Что мы и рассмотрим коротко.

### 7.3.1.2 Тактильное и мышечно-суставное чувства

Еще одним проприоцептивным источником информации о гравитационной вертикали является влияние последней на наше тело. В частности, весьма информативно давление на кожу, возникающее при сопротивлении корпуса силе тяжести. Например, если мы, стоя, сильно наклонимся вправо, то почувствуем мощное давление на правую ступню и соответствующее расслабление в левой. То же касается напряжения ягодицы, если мы наклонимся сидя. Большую часть времени тактильные ощущения остаются неосознанными, но при этом все равно фиксируются и постоянно отслеживаются нашим мозгом, обеспечивая равновесие.

Стоит сказать, что давление на кожу формирует вестибулярно независимый источник информации о вертикали. Чтобы источник этот был эффективен в контексте обеспечения константности ориентации, ощущение давления должно дополняться информацией о том, как наша голова расположена по отношению к проприоцептивно информативным частям нашего тела. Эти данные мы получаем через т.н. *мышечно-суставное чувство*,<sup>1</sup> информирующее нас о том, каковы относительные пространствен-

<sup>1</sup> В зарубежных источниках часто говорят не о мышечно-суставном чувстве, а о «кинестетической обратной связи». — Прим. пер.

ные ориентации наших суставов. Совокупно проприоцептивная информация и мышечно-суставное чувство позволяют определять положение головы относительно поля тяготения Земли.

### 7.3.2 Иллюзии ориентации

Как было сказано выше, зрительная информация влияет также и на восприятие ориентации объектов сцены. Эффект на удивление силен — зачастую настолько, что порой взяв верх над проприоцептивной информацией, он вызывает устойчивую зрительную иллюзию.

Одной из наиболее красноречивых демонстраций влияния зрительной информации на перцепцию ориентации мы наблюдаем, когда входим в слегка наклоненное помещение, подобное тем, что бывают в Луна-парках. В первый момент можно заметить некоторый наклон двери, но очень быстро комната начинает восприниматься как строго вертикальная. Стóит появиться такому перцепту, как за ним сразу приходят всевозможные иллюзии ориентации: люстру мы видим как висящую под каким-то странным углом к потолку, а самих себя мы ощущаем «покосившимися» на какую-то одну сторону (несмотря при этом на то, что на самом деле стоим строго вертикально). Если же мы попытаемся выправить свое положение, подстроившись под расположение стен в комнате, то легко можем потерять равновесие и упасть.

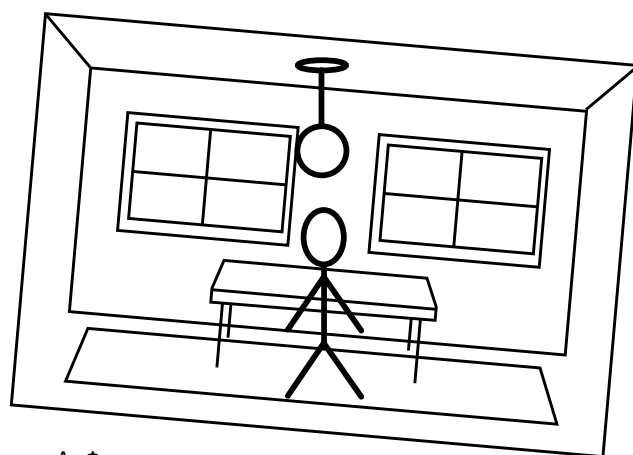
#### 7.3.2.1 Система отсчета

**Косая комната.** В косой комнате Луна-парка наше чувство гравитационной вертикали подчинено ложной системе отсчета, навязанной нам структурой помещения. Система отсчета в восприятии — это совокупность неких эталонов, по отношению к которым воспринимаются свойства объектов. Среди наиболее значимых эталонов — ориентация. Когда мы стоим в обычных, естественных условиях, ориентация нашей системы отсчета совпадает с гравитационной вертикалью, поскольку доминирующие предметы в сцене (стены домов, уличные фонари, люди и т.д.) расположены либо параллельно, либо перпендикулярно вектору силы тяжести.

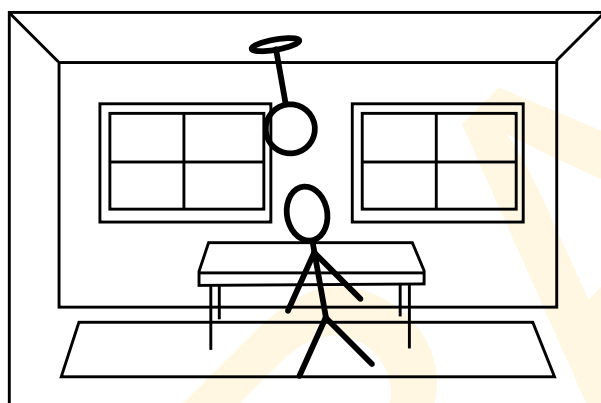
Бессознательное эвристическое предположение о том, что стены строго вертикальны, а пол и потолок — горизонтальны, обычно верное и руководит нашим восприятием ориентации объектов, давая веридикальные перцепты. Но когда мы ходим внутри нарочно косой комнаты, предположения эти оказываются ошибочными, запуская механизм формирования иллюзии ориентации. Зрительная точка отсчета в косой комнате выравнивает наше чувство вертикали так, как показано на рис. 7.3.3. В итоге, мы и чувствуем себя наклоненными, поскольку ориентация нашего собственного тела не совпадает с нашим восприятием вертикали. Как мы увидим далее, концепция системы отсчета во многих ситуациях объясняет возникновение иллюзорных перцептов, наведенных определенными зрительными контекстами.

**Прутикооконный эффект.** Еще одним из давно и хорошо известных контекстуальных феноменов является т.н. *прутикооконный эффект*, активно изучавшийся в свое время психологами Соломоном Эшем и Эндрю Виткиным (1948а, 1948b). Исследователи предъявляли наблюдателям светящийся прутик внутри большого наклоненного светящегося четырехугольного окна и просили испытуемых поставить прутик строго вертикально (см. рис. 7.3.4). Ученые обнаружили при этом, что испытуемые допускают систематическую ошибку (рис. 7.3.4 В), ставя прутик в положение где-то между истинной вертикалью (рис. 7.3.4 А) и полностью окномзависимым вариантом, когда прут





А. Фактически



В. Результат восприятия

**Рис. 7.3.3** Иллюзия, возникающая в косой комнате. Наблюдатель внутри такой комнаты (А) воспринимает ее вертикальной, а себя наклоненным (В).

стабильная во времени, как, скажем в косой комнате.

параллелен стороне окна, наиболее близкой к вертикали (рис. 7.3.4 С).

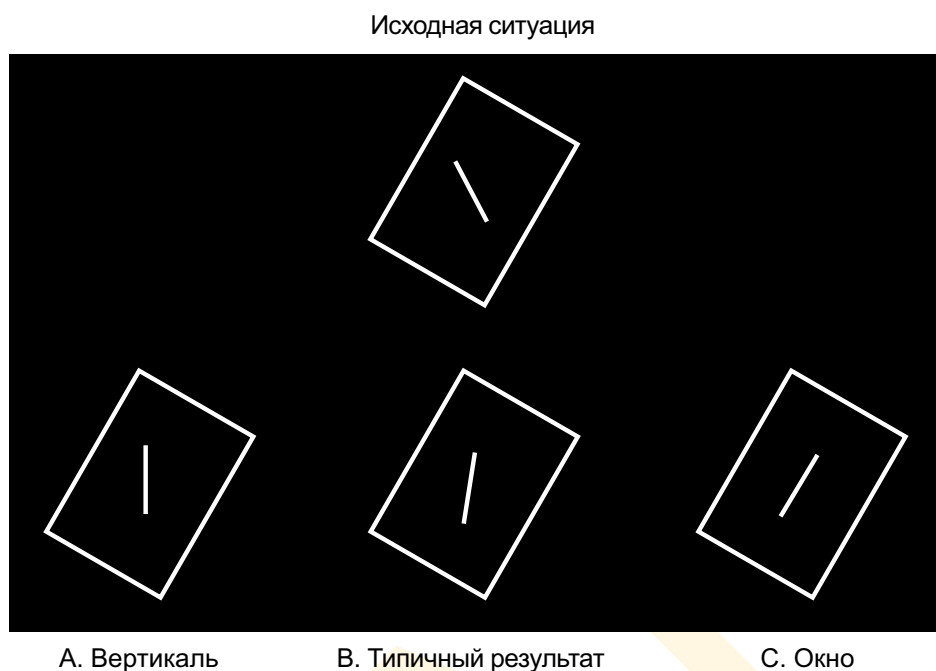
Подобно уходу константности ориентации в косой комнате, прутикооконный эффект можно объяснить, предположив, что окно влияет на наше восприятие вертикали в направлении собственной ориентации. Но в отличие от косой комнаты, влияние ориентации окна неполно. То есть, испытуемые редко ставят прутик параллельно наиболее вертикальной стороне окна, но чаще, как мы уже сказали, выбирают некое компромиссное положение между истинной вертикалью и ориентацией сторон.

Некоторые опыты показали, что эффект тем больше, чем крупнее окно и наоборот (Ebenholtz, 1977; Wenderoth, 1974). В прочих исследованиях обнаружилось, что когда наблюдателю предъявлены два окна — одно внутри другого — крупное окно перцептивно доминирует (DiLorenzo & Rock, 1982). Приведенные факты стыкуются с той идеей, что прямоугольник в прутикооконной задаче влияет на зрительную точку отсчета лишь как «суррогат реальности» (Rock, 1990). То есть, мы говорим, что зрительная структура в большей степени влияет на точку перцептивного отсчета, когда структура эта крупная, всезаполняющая и

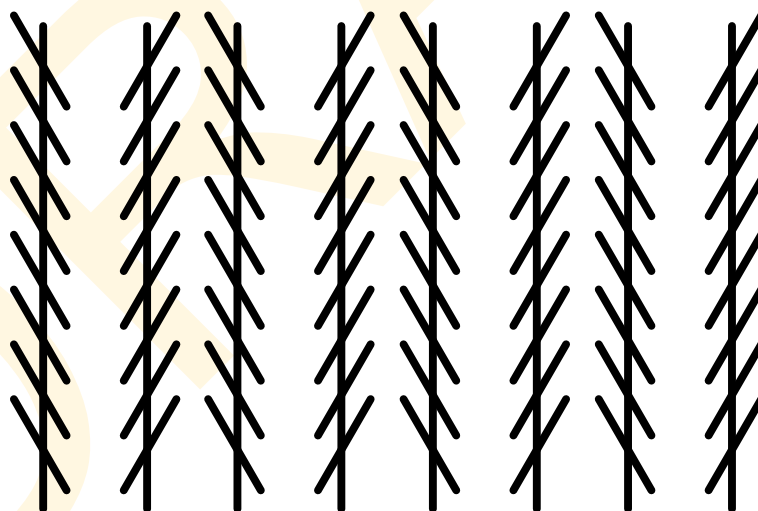
### 7.3.2.2 Геометрические иллюзии

Существует группа геометрических иллюзий ориентации, которые выглядят вполне независимыми от наших погрешностей в восприятии вертикали. Одна из наиболее известных — иллюзия Зольнера, при которой параллельные линии воспринимаются наклоненными, из-за того, что на их восприятие влияет фон из множества косых отрезков — рис. 7.3.5.

И хотя объяснение этой иллюзии весьма туманно, его можно отнести к идее контраста, звучавшей уже в отношении иллюзий размера. В данном случае ориентация длинных вертикальных линий может восприниматься с относительным контрастом к положению коротких отрезков: самая левая линия, например, выглядит слегка повернутой по часовой стрелке, поскольку именно так она и ориентирована по отношению



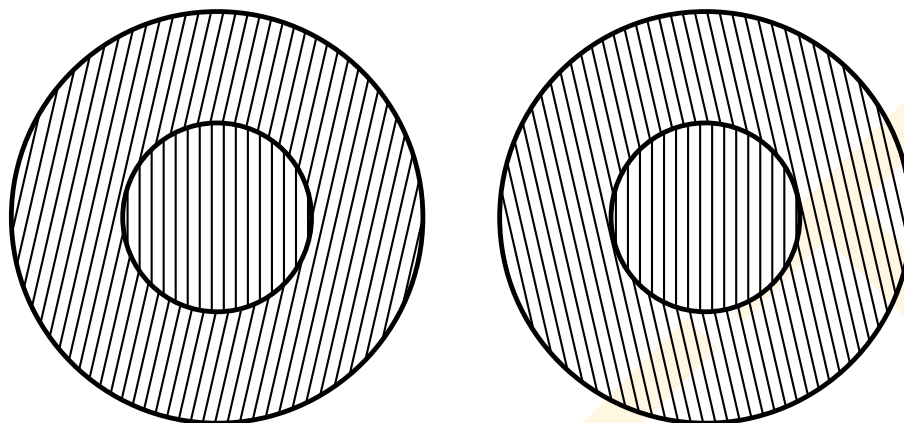
**Рис. 7.3.4** Прутикоокопный эффект. Когда испытуемым в полной темноте предъявляли ярко светящийся прутик внутри светящегося наклоненного окна, а затем просили поставить этот прутик строго вертикально, наблюдатели ориентировали его в направлении длинной стороны окна (В), но не в сторону истинной вертикали (А).



**Рис. 7.3.5** Золльнерова иллюзия. Вертикальные линии воспринимаются наклоненными (сходящимися и расходящимися), хотя фактически строго параллельны.

ко всем отрезкам, пересекающим ее (последние повернуты против часовой стрелки от вертикали); вторая линия смотрится повернутой против часовой стрелки, контрастируя с наводящими отрезками, ориентированными по часовой стрелке, и т.д.

Контрастные иллюзии в восприятии ориентации предметов давно известны: на



**Рис. 7.3.6** Одновременный контраст ориентации. Центральные решетки строго вертикальны в обеих фигурах, но под влиянием окружений выглядят наклоненными в разные стороны.

рис. 7.3.6, к примеру, две решетки в центре кругов фактически идеально вертикальны и параллельны друг другу, но воспринимаются слегка наклоненными в разные стороны и непараллельными, поскольку имеет место контекстное влияние окружающих фоновых решеток. Однако точный физиологический механизм иллюзии до сих пор остается неясным: одно из возможных объяснений опирается на ту гипотезу, что между нейронами имеется некое ингибиторное взаимодействие (например, Blakemore, Carpenter & Georgeson, 1970; Tyler & Nakayama, 1984).

Еще одно возможное объяснение зольнеровой иллюзии звучит так, что наше зрение переоценивает малые углы и недооценивает большие из-за латерального торможения между нейронами, ответившими на линии, образовавшие угол (например, Bekesey, 1967). В результате острые углы между вертикальными линиями и перекрестными штрихами выглядят больше, чем они есть на самом деле. Чтобы объяснить мнимый наклон вертикалей, мы должны сделать еще одно предположение — допустить, что иллюзорное изменение положения больше касается длинных линий, чем коротких. Но верно ли такое предположение — до сих пор неясно. Возможно от того, что даже если вертикали длиннее пересекающих их отрезков, короткие вертикальные сегменты, заключенные между соседними отрезками, короче последних и, следовательно, менее чувствительны к изменениям в их перцептивной ориентации.

## 7.4 РАСПОЛОЖЕНИЕ

Следующим глобальным свойством объектов реального мира является их *расположение* (позиция) в сцене. Существуют как минимум два пути, которыми мы можем воспринять позицию объекта: относительно нас самих, т.е. нашего тела и, второй путь — относительно других объектов в сцене. В текущем разделе мы сосредоточимся на восприятии расположения объектов относительно наблюдателя, что часто называют «положением относительно эгоцентра». В гл. 8 мы рассмотрим прочие схемы извлечения зрением информации о расположении предметов, т.е. информации, зависимой от позиции объекта относительно других объектов в сцене.

Позицию объекта относительно эгоцентра чаще всего характеризуют в показателях *полярных координат* этого объекта, т.е. *радиального направления* от наблюдателя на объект и *расстояния* от наблюдателя до объекта. И таким способом можно охарак-

теризовать позицию всех видимых объектов сцены. Стоит сразу уточнить, что радиальное направление на объект от наблюдателя — это направление от переносицы наблюдателя (т.е. середины между его глазами — т.н. *циклопический глаз*) на центр объекта. Восприятие направления в расположении объектов — главный вопрос данного раздела. При этом о факторах, управляющих перцепцией расстояния, мы не будем говорить, поскольку уже обсудили этот вопрос в гл. 5.

### 7.4.1 Восприятие направления

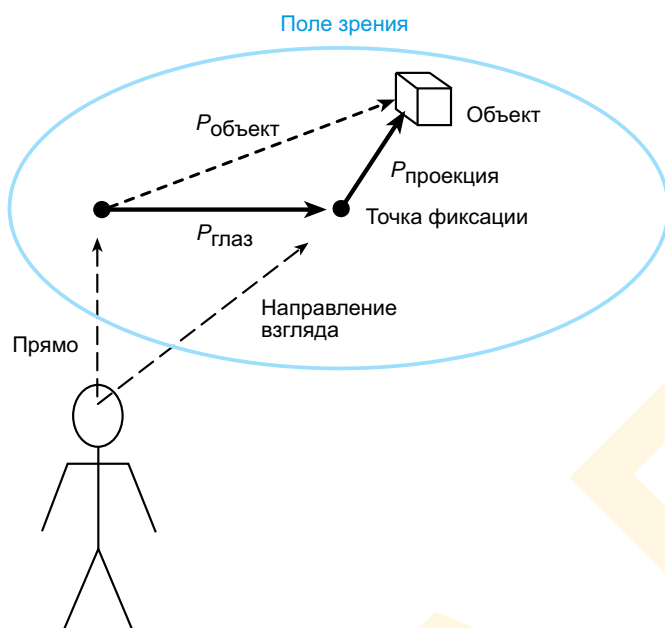
Восприятие радиального направления на объект (т.е. направления от эгоцентра) начинается с оценки информации о позиции проекции этого объекта на сетчатку. Положение проекции объекта на сетчатку относительно центральной ямки сетчатки определяется склонением относительно направления взгляда наблюдателя, поскольку сигнал от каждого ретинального рецептора вызван фотонами, прилетевшими в одном, строго заданном направлении. Мы уже знаем, что информация о положении ретинальной проекции фиксируется на ранних этапах зрительной обработки посредством множества ретинотопических карт, сохраняющих относительное расположение элементов ретинального изображения в двухмерных листках корковых клеток.

Однако может ли сия физиологическая структура отвечать за восприятие направления от эгоцентра на объект? К сожалению, нет — не может. По крайней мере сама по себе. Количественный анализ свидетельствует о том, что направление на объект зависит не только от ретинальной позиции его проекции, но также и от направления, куда смотрят глаза наблюдателя. К примеру, когда мы глядим строго вперед, точка, стимулирующая центр сетчатки, воспринимается как расположенная «прямо по курсу», т.е. строго впереди. Но когда наши глаза направлены, скажем, на  $10^\circ$  вправо, точка, стимулирующая центральную ямку сетчатки расположится на  $10^\circ$  вправо от линии прямой видимости, а точка, располагавшаяся строго прямо, — теперь стимулирует ретинальную локаль на  $10^\circ$  справа от центральной ямки (т.е. на  $10^\circ$  влево от центра поля зрения). То есть, перцептивное направление на объект не зависит от положения наших глаз (Hill, 1972). Сей факт убедительно свидетельствует о том, что наше зрение определяет радиальное направление на объект, интегрируя данные о положении глаз и ретинальной позиции стимуляции.

Каждый из параметров (т.е. позицию объекта по отношению к прямому направлению от эгоцентра [ $P_{\text{объект}}$ ], позицию проекции объекта на сетчатку по отношению к центральной ямке [ $P_{\text{проекция}}$ ] и позицию глаза по отношению прямой видимости от эгоцентра [ $P_{\text{глаз}}$ ]), можно выразить в виде вектора. На рис. 7.4.1 показаны отношения этих величин:

$$P_{\text{объект}} = P_{\text{проекция}} + P_{\text{глаз}} \quad (7.3)$$

Когда голова наблюдателя направлена строго вперед, то направление взгляда от эгоцентра берется непосредственно из позиции глаз относительно головы. Но если голова повернута, положение глаз нужно рассчитывать, добавляя к вектору позиции глаз относительно головы вектор положения головы относительно тела. Если такое уравнение окажется разрешимо зрительной системой, можно достичь константности расположения объектов сцены.



**Рис. 7.4.1** Константность расположения. Положение объекта относительно эгоцентра ( $P_{\text{объект}}$ ) зависит как от направления взгляда относительно эгоцентрически прямого направления ( $P_{\text{глаз}}$ ), так и направления на объект внутри поля зрения ( $P_{\text{проекция}}$ ).

константность в перцептивном направлении объектов реальной сцены от эгоцентра можно рассчитать, сложив вектор смещения проекции объекта на сетчатку с вектором смещения глаза относительно тела (см. рис. 7.4.3). Каждый вектор представляет собой изменение в позиции (DP) глаза, объекта реальной сцены или проекции этого объекта относительно их исходных позиций. Так, нулевой вектор ( $\Delta P = 0$ , отмечен точкой на рис. 7.4.3) представляет неизменность положения, а прочие векторы — его изменение, слагаемое из *расстояния по заданному направлению*. Если мы обозначим изменение в направлении на реальный объект после смещения глаз как  $\Delta P_{\text{объект}}$ , смену положения его ретиальной проекции как  $\Delta P_{\text{проекция}}$ , и смену направления глаз как —  $\Delta P_{\text{глаз}}$ , то отношение между этими векторами можно записать как:

$$\Delta P_{\text{объект}} = \Delta P_{\text{проекция}} + \Delta P_{\text{глаз}} \quad (7.4)$$

Например, если объект неподвижен, а наблюдатель перевел глаза вправо, то ретиальная проекция объекта сдвинулась на ту же величину влево, т.е. с обратным знаком. Перцептивное смещение объекта в сцене в этом случае (нулевой вектор —  $0^\circ$  зрительного угла) — это сумма смещения глаз (скажем,  $30^\circ$  вправо) и результирующего смещения ретиальной проекции объекта ( $30^\circ$  влево). Таким образом, о константности

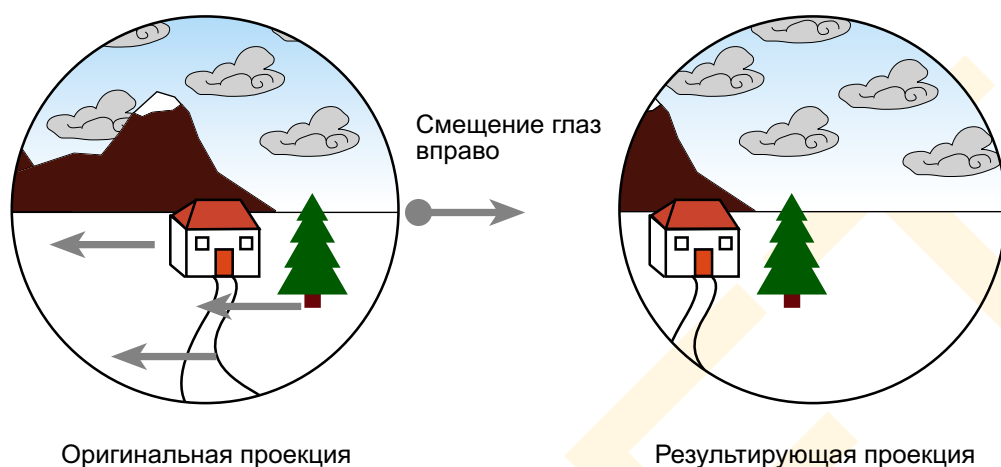
### 7.4.2 Константность расположения

Наши глаза постоянно в движении, и ретиальные проекции предметов сцены непрерывно прыгают по сетчатке. Например, смещение глаз вправо ведет к тому, что проекции всех стационарных объектов сдвигаются влево относительно поля зрения (рис. 7.4.2).<sup>1</sup> Тем не менее, несмотря на замысловатую зигзагообразную траекторию, стационарные объекты воспринимаются прочно стоящими на своих местах в сцене, что получило название *константности расположения от эгоцентра*. То, как зрительная система достигает ее — предмет нашего разговора в текущем разделе.

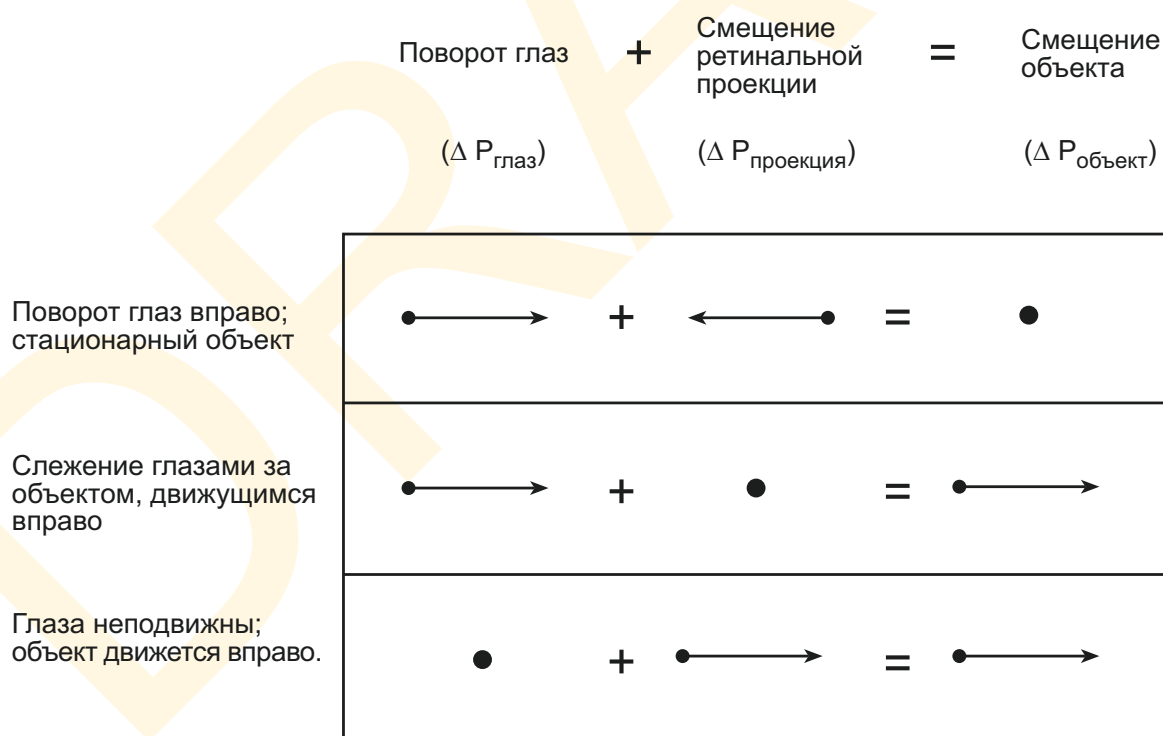
Количественный анализ ситуации убеждает в том, что

<sup>1</sup> Поскольку оптическая проекция объекта на сетчатку полностью инвертирована (включая право-лево), изображения объектов, фактически смещаются на сетчатке вправо, когда глаза поворачиваются вправо. Мы не учитываем сей факт в наших рассуждениях, потому что он неимоверно усложнит изложение материала.





**Рис. 7.4.2** Константность расположения и движение глаз. Когда глазные яблоки сдвигаются в одном направлении, проекции объектов на сетчатке смещаются на соответствующее расстояние в направлении противоположном.



**Рис. 7.4.3.** Векторная поправка к повороту глаз и смещению ретинальной проекции. Перцептивное смещение позиции объекта после поворота глаз — это сумма вектора смещения глазных яблок и вектора смещения ретинальной проекции.

направления докладывает нулевой вектор смещения объекта в реальной сцене. Сие имеет место тогда, когда движение глаз и смещение ретинальной проекции полностью обнуляют друг друга.<sup>1</sup>

#### 7.4.2.1 Непрямые теории константности расположения

Каким же образом достигается перцептивная константность расположения объектов, если глаза всё время в движении? Классическая гельмгольцева идея гласит, что посредством бессознательной инференции, основанной на решении векторного уравнения, приведенного выше: если наблюдатель сенсорным сигналом осведомлен о смещении проекции объекта на сетчатку и о смещении самих глаз, сдвиг объекта в реальной сцене можно рассчитать, попросту сложив векторы согласно уравнению 7.3.

Но откуда поступает информация о смещении глаз? Гельмгольц говорит о двух вариантах ответа:

1. *Афферентная версия* (часто именуемая *версией входного сигнала* или *обратной связью*). Гласит, что после того как мозг дал сигнал глазам на смещение в определенном направлении, сенсоры глазодвигательных мышц регистрируют информацию о мышечном напряжении. Данные отправляются в мозг, реализуя тем самым обратную связь, информирующую его о реальном положении глаз. Мозг объединяет входящую (афферентную) информацию с данными о положении ретинальной проекции объекта, что и позволяет выполнить расчет по уравнению 7.3.

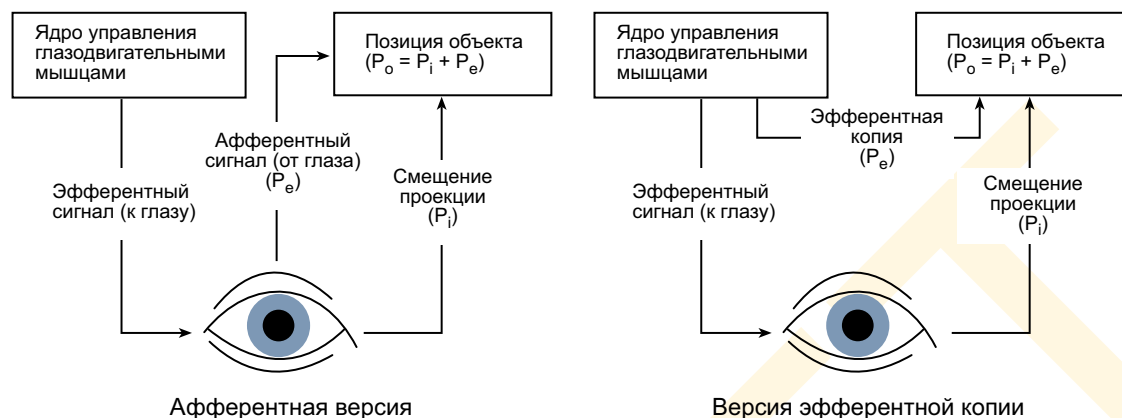
2. *Версия эфферентной копии* (теория выходного сигнала или командная версия), предполагает, что всякий раз, когда мозг посылает выходной (эфферентный) управляющий сигнал глазодвигательным мышцам, он отправляет и копию этого сигнала в те центры, что рассчитывают новые позиции объектов зримой сцены. Этот сигнал иногда называют *дополнительным*. После чего для выполнения расчетов по уравнению 7.3 он объединяется с вектором ретинальной проекции.

Обе концепции предполагают, что при расчете позиции объекта используется информация о глазных движениях, но при этом разнятся они тем, каков источник этой информации: в эфферентной теории — это копия выходного командного сигнала; в афферентной — входящий сигнал обратной связи как результат исполнения команды. На рис. 7.4.4 даны принципиальные схемы каждой из версий.

**Опыт Гельмгольца.** Гельмгольц (1867/1925) полагал, что сделать выбор в пользу одного из вариантов можно, *пассивно* сдвинув глазное яблоко (т.е. внешним усилием, но не центральной командой — командой из мозга). Если механизм константности базирован на афферентном сигнале от глазодвигательных мышц, рассуждал ученый, он «добросовестно» отразит пассивное смещение глазных яблок и константность позиции сохранится. Если же механизм эфферентный, константности не будет, поскольку смещение глаз вызовет изменение ретинальной позиции без какого либо командного сигнала, нейтрализующего это смещение. Если при этом к главному яблоку мы приложим внешнюю силу, сцена перцептивно сдвинется.

Опыт столь прост, что можно провести его прямо сейчас: закроем один глаз и аккуратно надавим пальцем на внешний угол открытого; когда палец сдвинет глазное яблоко, сцена перцептивно качнется в сторону. Разумеется, это качание сугубо иллюзорно. Очевидно, что результат опыта говорит в пользу версии эфферентной копии.

<sup>1</sup> Отметим, что приведенное выше уравнение полностью идентично предыдущему, с той только разницей, что представляет *изменение* в направлении, а не само направление.



**Рис. 7.4.4.** Афферентная версия и версия эфферентной копии, толкующие константность расположения объектов при движении глаз. Афферентная гипотеза предполагает, что сигнал обратной связи, идущий из глаза, суммируется с сигналом смещения ретиальной проекции, что и позволяет высчитать изменение в расположении объекта. Концепция эфферентной копии предполагает, что копия команды на смещение глазных яблок суммируется с сигналом ретиального смещения.

Отметим, что в свое время звучало множество возражений против столь простой демонстрации, но хирургический ее вариант (с применением щипцов и анестезии), дал ровно те же результаты (Brindley & Merten, 1960).

**Опыт Маха.** Еще один тест, позволяющий сделать выбор между описанными двумя толкованиями выполнил Эрнст Мах (1914/1959), исходя при этом из положений, обратным гелмгольцевым. Мах рассудил так, что ежели команда глазодвигательным мышцам была дана, но по каким-то причинам не реализовалась, сцена все равно перцептивно сместится даже если глазное яблоко останется неподвижным. Афферентная концепция прогнозирует, что, наоборот, — перцептивного сдвига сцены не возникнет, поскольку глаза остаются неподвижны, и, следовательно, нет изменений в сигнале обратной связи от глазодвигательных мышц. Мах проделал опыт на себе, зафиксировав глаза оконной замазкой через переносицу. Затем, когда он попытался шевельнуть глазами, сцена перцептивно сдвинулась. То есть, и махов результат прозвучал в защиту концепции эфферентной копии.

**Опыт Стивенса.** Куда более суровую и опасную методику фиксации глаз опробовал на себе психолог Дж. Стивенс: он попросил знакомого врача сделать ему укол кураре — яда, временно парализовавшего глазодвигательные мышцы отважного исследователя (Stevens et al., 1976). Но поскольку кураре действует в том числе и на дыхательные мышцы, Стивенсу пришлось подключить аппарат искусственной вентиляции легких. Коль скоро глазодвигательные мышцы ученого были полностью блокированы, он совершенно не мог двигать глазами, что явилось абсолютной проверкой маховых прогнозов. Однако оставалась одна проблема: герою требовалось каким-то путем сообщить своим соратникам о том, имеется ли перцептивное смещение или нет при попытке шевельнуть глазами. Поэтому одну руку заранее перетянули жгутом, что не позволило кураре повлиять на ее мышцы и что дало возможность сигнализировать ею о том, есть перцептивное движение сцены или нет. Результат опыта полностью подтвердил открытия Маха в опыте с замазкой.

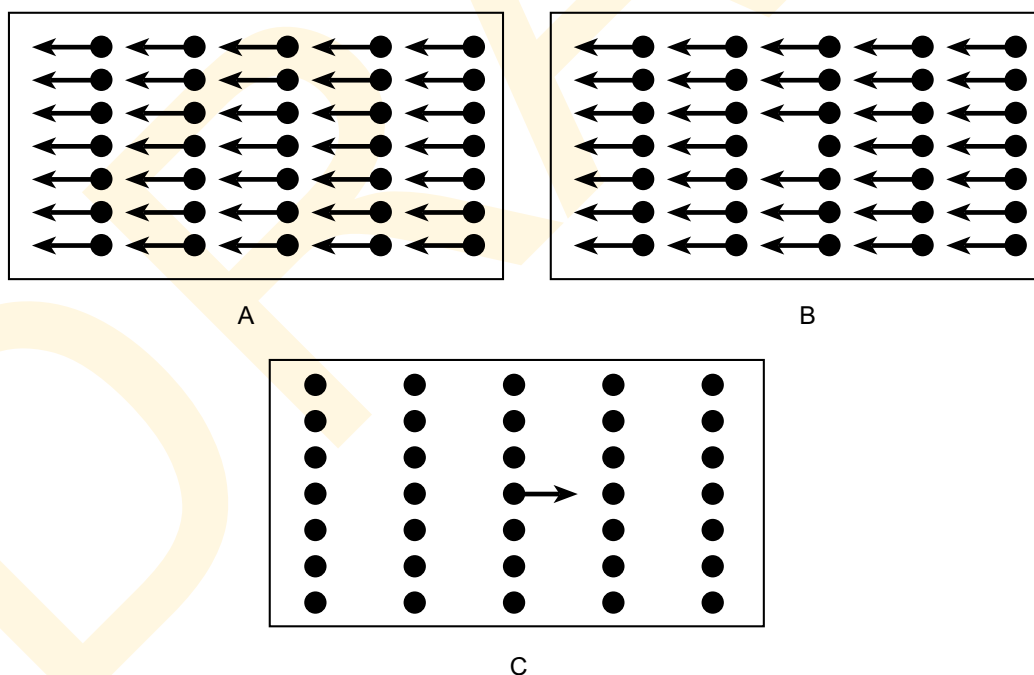
Вывод из всех описанных тестов получается таков, что константность расположения при движении глаз сохраняется мозгом через копию эфферентной глазодвигатель-

ной команды, посылаемой в те мозговые центры, где смещение ретинальной проекции корректируется смещением глазных глаз. Однако система работает безупречно до тех пор, пока команда вызывает соответствующее движение глаз. Если движения нет (как в опытах Маха и Стивенса) или же, если движение это вызвано чем-то другим, но не эфферентной командой (как в опыте Гельмгольца), те же самые механизмы формируют иллюзию движения.

#### 7.4.2.2 Прямые теории константности расположения

Концепция эфферентной копии — это т.н. *непрямая* теория, построенная на множестве опосредованных выводов: например, изменения в расположении объекта от эго-центра выводятся из смещения ретинальной проекции путем учета вектора поворота глаз. Однако существует и альтернативная — прямая версия константности, предложенная Гибсоном (1966), и целиком построенная им на свойствах оптического потока.

Суть гибсонового концепта такова: если мы сместим глаза в каком-то направлении, но голова при этом останется неподвижной, проекция объекта на сетчатку сдвинется в противоположном направлении. Следовательно, поворот глаз вправо даст равномерное поле оптического потока, показанное на рис. 7.4.5 А. Гипотетически та же картина движения ретинальной проекции сцены возникнет, ежели вся сцена вдруг обернется влево вокруг оси, проходящей через глаз наблюдателя. Но понятно, что вероятность такого события исчезающе мала. Таким образом, равномерный оптический поток данного направления информирует о том, что глаз на соответствующую величину повернулся в направлении противоположном. Следовательно, константность расположения объек-



**Рис. 7.4.5.** Оптический поток при константности расположения. Когда мы смещаем глаза вправо, ретинальные проекции равномерно смещаются влево (А). Когда мы следим за объектом, движущимся вправо, его проекция на сетчатку неподвижна, а проекции всех остальных объектов смещаются влево (В). Если вычесть общий вектор движения поля зрения из (В), получим истинную картину — движущийся вправо объект на неподвижном фоне.

тов относительно эгоцентра при движении глаз можно объяснить, предположив, что зрительная система попросту вычитает любой общий вектор движения в поле потока.

На рис. 7.4.5 В показано поле потока, образующееся при слежении наблюдателем за объектом, движущимся вправо. Объект фиксации неподвижен, но всё оставшееся поле движется в противоположном направлении. Если сей общий компонент движения вычесть по каждой точке, результирующее поле потока, отразит подлинное положение дел — единичный объект, движущийся вправо, при движении глаз наблюдателя и неподвижном фоне — рис. 7.4.5 С.

Существует великое множество сложнейших обстоятельств, в которых ретинальные позиции неподвижных объектов меняются, но мы безошибочно воспринимаем их стационарными. К этим обстоятельствам относятся движение головы, всего нашего тела и глаз. Если положение глаз остается фиксированным относительно головы, ее поворот вызывает смещение ретинальной проекции объекта в направлении противоположном этому смещению, то есть так же, как и при глазных движениях. Пока информация о повороте головы доступна мозгу, смещение ретинальной проекции объекта компенсируется.

Когда и голова, и глаза двигаются одновременно, их векторы следует сложить, дабы обнулить результирующий сдвиг ретинальной проекции. Однако и здесь информации, заключенной в поле оптического потока, достаточно, чтобы выполнить нужные расчеты и достичь константности положения (Longuet-Higgins & Prazdny, 1980).

Колебания всего нашего тела в пространстве при движении (ходьбе, беге и т.д.) также вызывают смещение ретинальной проекции стационарных объектов без перцепта их движения. Рассмотрим, что происходит, когда мы просто идем вперед, устремив взгляд, к примеру, на объект, расположенный прямо перед нами. Как показано на рис. 7.4.6, проекции прочих объектов радиально «разбегаются». Но даже в этом случае мы не воспринимаем эти объекты движущимися, но однозначно неподвижными. Однако сложные преобразования, индуцированные движением наблюдателя в сцене, невозможно истолковать как простое компенсаторное смещение, как в случае поворота головы и/или глаз: на базе картины оптического потока зрение должно выполнить сложнейший обсчет и позиции объектов, и движения наблюдателя. Мы отложим этот разговор до 10-й главы, в которой речь пойдет о восприятии движения.

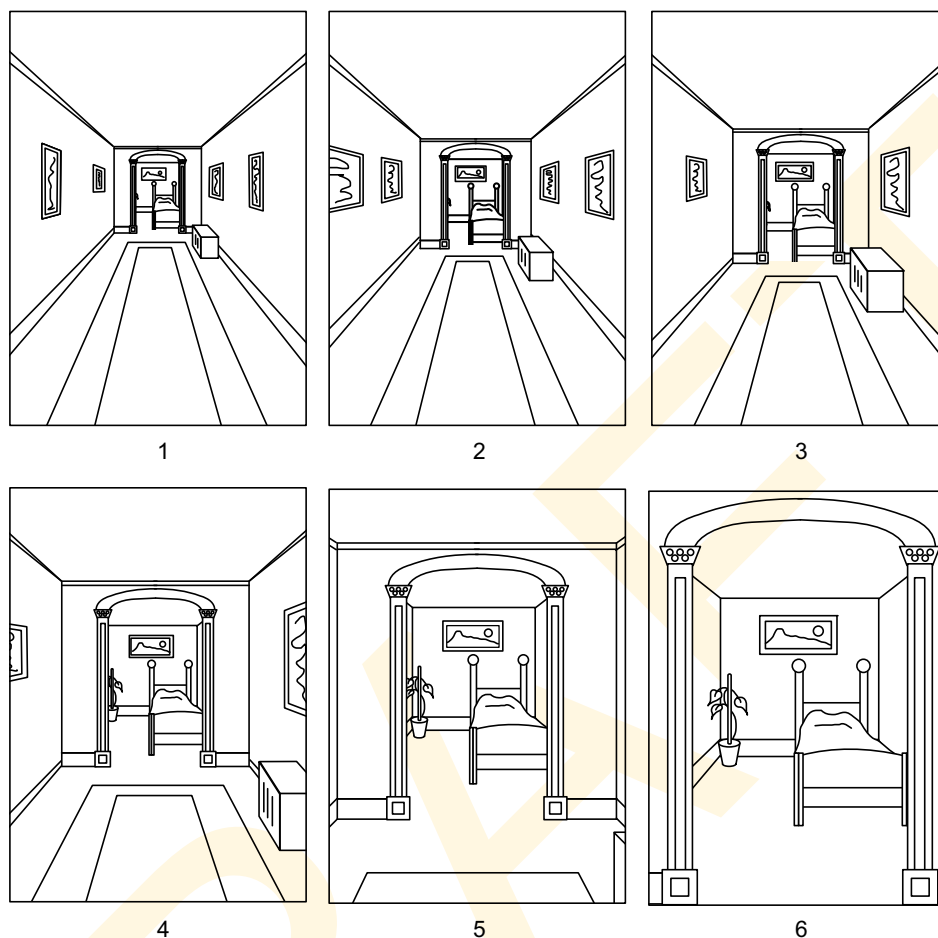
### 7.4.3 Иллюзии расположения

Как было показано ранее, наше восприятие расположения объектов относительно эгоцентра весьма и весьма устойчиво. Отметим, что с эволюционных позиций сие крайне значимо, поскольку активному мобильному организму исключительно важно знать, как расположены предметы относительно него самого, нежели каковы проекции этих предметов на сетчатку. Но при этом все равно нельзя назвать наше восприятие позиции предметов идеально надежным.

На рис. 7.4.7 дан пример иллюзии расположения, возникшей под контекстным влиянием окружения. Профессор Амстердамского университета Отто Роелофс (1935) в затемненном помещении предъявлял своим испытуемым светящуюся прямоугольную рамку и просил их указать направление прямо перед собой. Исследователь обнаружил, что если рамку дать строго перед наблюдателем, восприятие направления будет безупречным. Но когда ее показывают чуть сместив, испытуемые систематически воспринимают прямое направление от эгоцентра сдвинутым на несколько градусов в сторону центра окна. Феномен получил название *роелофсового эффекта*.

Иллюзия прямого направления от эгоцентра — это контекстный эффект, в многом



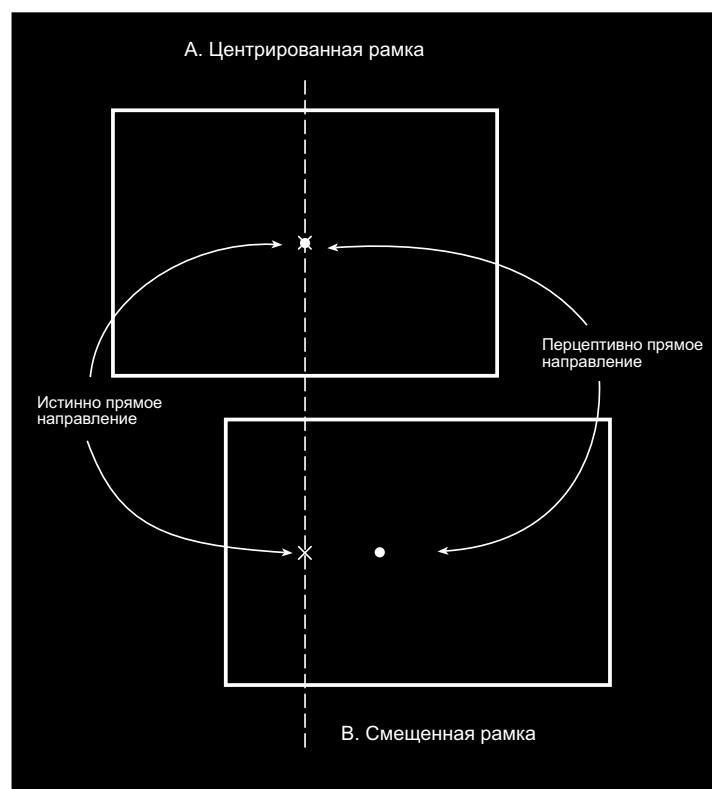


**Рис. 7.4.6.** Константность расположения. Когда мы движемся в сцене, структура оптического потока информирует наше зрение не только о том, где расположены предметы относительно нас, но и о том, как мы перемещаемся относительно окружения (обычно неподвижного). Количественные процессы, реализующие сие, очень сложны. Мы обсудим их в гл. 10.

По материалам Gibson, 1966.

сходный с прутикооконым, но не идентичный ему: в разделе 7.3 мы говорили о том, что наблюдатель порой допускает систематическую ошибку в оценке гравитационной вертикали, если она предъявлена в большом прямоугольном окне, несколько отклоненным от вертикали и горизонтали (рис. 7.3.4); при роелофсовом эффекте мы с трудом определяем направление вперед относительно самих себя, если нам предъявлен объект, несколько смещенный от центра.

Итак, мы видим, что ежели речь идет о таких глобальных зрительных свойствах объектов как размер, форма, ориентация и расположение, зрительная система успешно достигает их веридикальной перцепции во многих и многих условиях наблюдения. Но не во всех: в некоторых экологически редких ситуациях, те же самые механизмы, что в норме дают веридикальный перцепт, вызывают устойчивую иллюзию. Сие случается из-за того, что механизмы, лежащие в основе перцептивной константности — *эвристичны*, т.е. базированы на версиях, которые, хоть и чаще всего, но все же не всегда верны. Когда они оказываются верными — константность имеет место. Когда ошибочными — возникают разного рода иллюзии.



**Рис. 7.4.7** Роелофсов эффект. Когда четырехугольная рамка позиционирована относительно эгоцентрически прямого направления, перцептивное прямое направление веридикально. Когда сдвинута — перцептивно прямое направление смещается к центру четырехугольника.

## 7.5 ПЕРЦЕПТИВНАЯ АДАПТАЦИЯ

Очевидно, что в широчайшем диапазоне нормальных условий зрения мы весьма точно воспринимаем такие перцептивные свойства объектов как размер, форма, расположение и ориентация. Наш дальнейший разговор о том, как свойства эти высчитываются зрительной системой из параметров ретинальных проекций и прочих релевантных ситуации факторов: например, перцепт размера объекта зависит как от габаритов его ретинальной проекции, так и от расстояния до объекта сцены, а перцепт ориентации объекта — от ориентации ретинальной проекции и положения головы относительно гравитационной вертикали. Сие примеры того, что при реконструкции действительных свойств объекта из ретинальных данных, зрительная система использует разного рода сопутствующую информацию (меняющуюся в нормальных условиях зрения), как, например, расстояние до объекта и положение головы. Результатом работы сих компенсаторных процессов является то, что мы называем *перцептивной константностью*. Т.е. мы говорим, что зрительная система способна компенсировать нестабильность ретинального изображения.

Однако вопрос, который нам предстоит обсудить сейчас — способно ли зрение компенсировать те изменения в условиях наблюдения, которые в норме остаются неизменными, но которыми можно принудительно манипулировать оптически. К примеру, оптика глаза пространственно инвертирует ретинальное изображение реальной сцены,

из-за чего свет, пришедший из верхней части сцены, падает в нижнюю часть сетчатки и наоборот. И мы знаем, что аналогичные инверсии происходят по всем направлениям (рис. 1.2.6). В норме инверсия *всегда* имеет место и, следовательно, в норме никогда не требует специальной коррекции или переинтерпретации: наш организм либо при рождении уже наделен способностью к соответствующему пространственному мэппингу, либо моментально обучается ему сразу после рождения. Как бы выглядел окружающий мир, к примеру, если бы ретинальные изображения *не* обращались? И смогли бы мы обучиться их адекватному восприятию спустя какое-то время?

Вопрос влияния различных трансформаций ретинального изображения на перцептивные свойства объектов активно обсуждался в исследованиях т.н. *перцептивной адаптации* (Welch, 1986).

⇒ Перцептивная адаптация — это полупостоянные изменения в восприятии (или в перцептивно-моторной координации), снижающие сенсорные несоответствия, вызванные изменчивостью стимула.

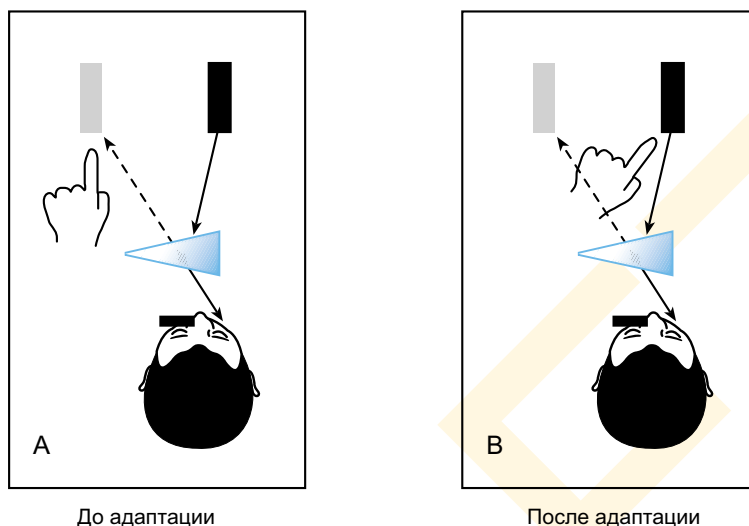
Перцептивная адаптация принципиально отлична от адаптации *сенсорной* (см. раздел 1.1) и, главным образом тем, что сенсорная адаптация — это *краткосрочное* изменение в восприятии, возникающее *автоматически* в результате нормальной, но пролонгированной стимуляции в единичной сенсорной модальности.

**Опыт Гельмгольца.** Гельмгольц (1867/1925) был среди первых, кто занялся систематическими исследованиями перцептивной адаптации. Как-то раз он провел необычный опыт, надев очки со специальными призмами, сдвигавшими в сторону изображение видимого мира примерно на  $11^\circ$ . Сие небольшое латеральное смещение в целом незаметно и все объекты выглядят как обычно. Однако если в такой ситуации мы закроем глаза и попробуем дотянуться до предметов, которые только что видели, то обязательно промахнемся на угол смещения призмой (рис. 7.5.1 А). Ошибка сия вызвана разницей между перцептивным положением объекта и его положением фактическим. Гельмгольц обратил внимание на то, что может быстро справиться с ошибкой, если начнет следить за своими руками (рис. 7.5.1 В). Снижение числа и степени моторных ошибок — первый признак наступления перцептивной адаптации.

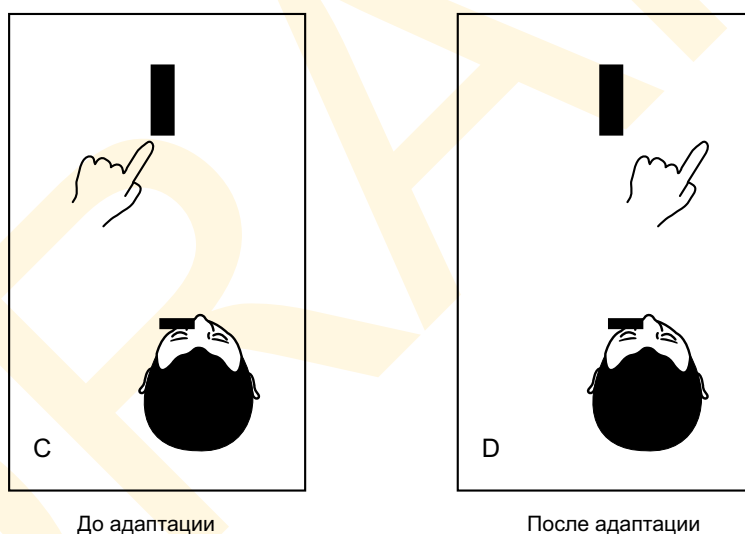
Ученый обнаружил также, что когда адаптация к смещению ретинального изображения наступила, то, если быстро снять призму с глаз, мы начинаем ошибаться в *противоположном* направлении. Сие есть т.н. *обратное последствие* (см. рис. 7.5.1 D) — второй важнейший маркер перцептивной адаптации. Он особенно значим, поскольку очевидно исключает вероятность того, что адаптация — это результат осознанной коррекции наблюдателем. Будь оно так, наблюдатель немедленно и с высокой точностью возвращался бы к норме сразу после снятия очков с призмами.

**Опыт Стрэттона.** Одно из первых и широко известных исследований адаптации к искаженным ретинальным проекциям было посвящено эффектам «реинверсии» ретинального изображения, о чем шла речь выше. Вспомним, что, несмотря на оптически перевернутое изображение на наших сетчатках, мы воспринимаем окружающий мир как прямостоящий. Джордж Малькольм Стрэттон — отец-основатель психологической лаборатории Калифорнийского университета — изучал перцептивную адаптацию к реинвертированным ретинальным проекциям (1896), рассматривая сцены через специальную призму, пространственно обрабатывающую изображение на сетчатке. В течение восьми дней в период бодрствования исследователь носил специальные очки, позволявшие ему видеть окружающий мир только через эту призму, расположенную перед его правым глазом. Ретинальная проекция из-за этого становилась прямо по отношению к

## Опыт с призмой



## Опыт без призмы



**Рис. 7.5.1.** Адаптация к оптическому смещению. До адаптации испытуемые точно показывали пальцем на целевой объект (С), но промахивались, когда на пути света ставили призму (А). При слежении за руками наступала адаптация, в результате которой наблюдатели вновь начинали точно попадать пальцем в целевой предмет (В), но промахиваться уже без призмы (D).

реальной сцене. Левый глаз был полностью закрыт плотным наглазником, поэтому всю зрительную информацию исследователь получал только в реинвертированном виде. На ночь ученый оба глаза прикрывал черной материей, чтобы во время пробуждения случайно не проэкспонировать сетчатку обычным инвертированным изображением.

Неудивительно, что поначалу мир через очки казался перевернутым. В течение восьми дней опыта Стрэттон пытался вести обычный образ жизни: сам одевался, ел, гулял, читал, писал и т.д. Однако все задачи, требовавшие визуально-моторной координации, давались ему поначалу невероятным трудом. Было так тяжело, что по нача-

лу ученому было легче действовать с закрытыми глазами, воображая правильно стоящий мир, т.е. без какой-либо фактической зрительной стимуляции. Но после того, как Стрэттон проносил очки несколько дней, он начал потихоньку адаптироваться к перевернутому стимулу. В итоге, исследователь начал спокойно читать, писать и выполнять обычные, привычные бытовые действия. В его отчете читаем, что к концу опыта, т.е. к восьми дням, Стрэттон уверенно ориентировался в пространстве, быстро и точно дотягивался до нужных ему предметов. Исследователь понял, что адаптировался моторно, но произошла ли адаптация *перцептивная*? Стал ли он *ощущать* внешний мир как расположенный правильно?

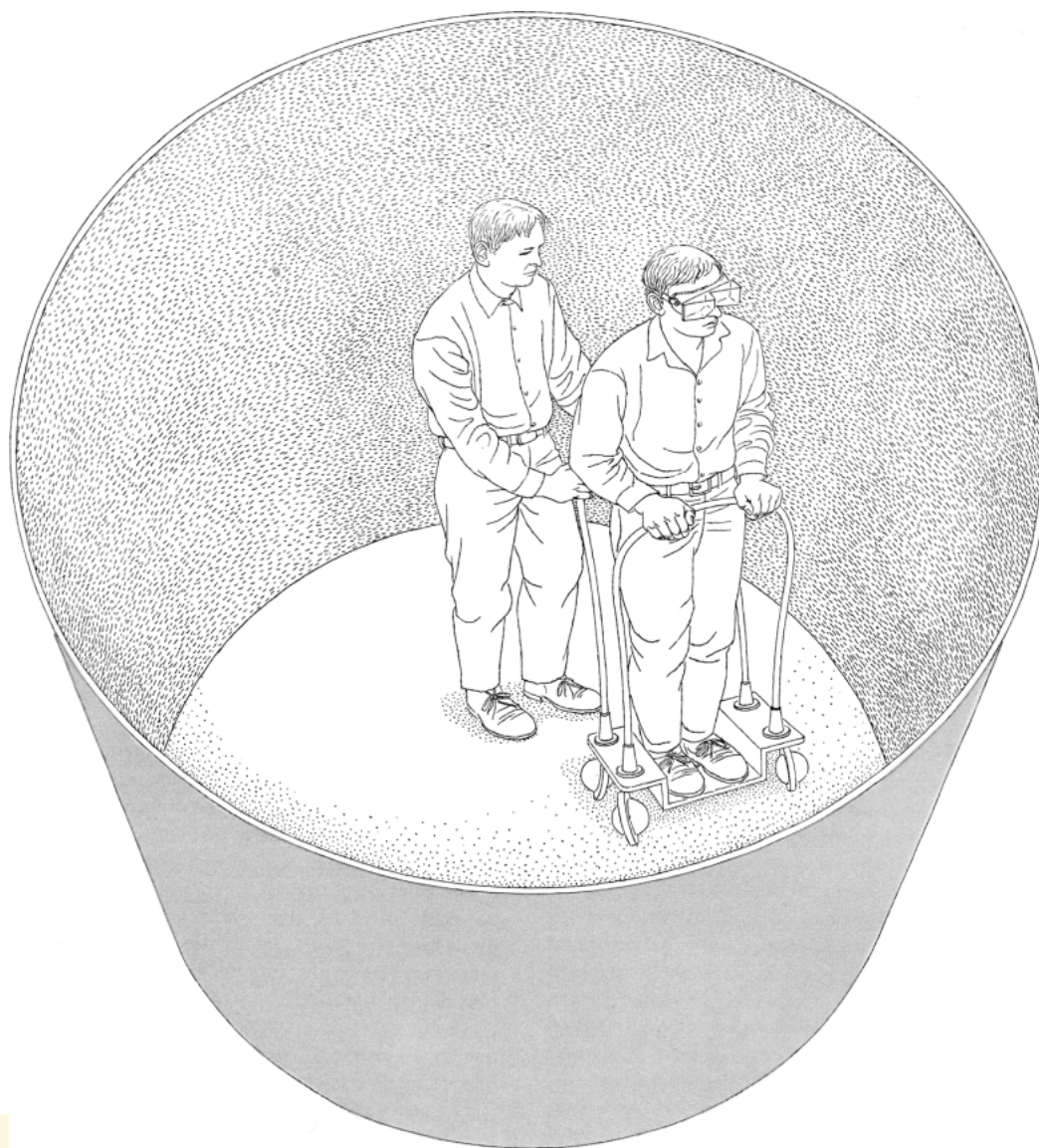
Вопрос прямой и конкретный, но ответ далеко не очевиден. Стрэттон сообщил о нескольких случаях (особенно, если не думал об этом эксплицитно), когда он забывал о том, что реальность перевернута очками, а к концу эксперимента многие предметы уже казались ему прямостоящими. Впрочем, так было не все время: иногда только часть сцены воспринималась прямо, а оставшаяся — перевернутой. В другой раз Стрэттон сцену ощущал прямостоящей, но себя — перевернутым, как бы свешивающимся с потолка. За весь опыт не нашлось ни одного более или менее продолжительного отрезка времени, когда ученый ощущал бы окружающий мир целиком как в обычной жизни. Более того, когда, наконец, к концу восьмого дня, Стрэттон снял свои очки, он не обнаружил отрицательного последствия: мир вокруг не «перевернулся» вверх тормашками (сие отрицательное последствие обычно полагают лакмусовой бумажкой подлинной перцептивной адаптации).

Однако звучат возражения, что, мол, Стрэттон носил очки недостаточно долго, что суженное поле зрения каким-то образом затрудняло адаптацию и т.д. Поэтому несколько психологов повторили стрэттонов опыт, но уже в течение большего времени и с применением усовершенствованных методов пространственной инверсии ретинальной проекции (например, Kohler, 1962). Полученные ими данные в целом подтвердили стрэттоновы выводы о моторной адаптации: Кёхлер, например, в период опыта, то есть с инвертирующей призмой на глазах, мог кататься на велосипеде и на лыжах. Однако в отличие от Стрэттона Кёхлер периодически и устойчиво видел окружающий мир полностью прямостоящим. Причем в течение длительного времени.

Еще один феномен, обнаруженный Стрэттоном, заключался в том, что, надев инвертирующие очки, он утратил константность расположения предметов относительно эгоцентра. Он описал это как «дрожание поля зрения», которое возникало в моменты поворота головы. Дрожание при инверсии возникало по следующей причине: когда мы поворачиваем голову вправо, оптический поток в норме движется влево. Сие движение обнуляется сложением векторов смещения глаз и ретинальных проекций, что дает константность расположения объектов (см. раздел 7.4.2) при смещении глаз. При вращении головой — то же. Однако поскольку призма поменяла право и лево, левонаправленный оптический поток стал правонаправленным. Когда правонаправленный вектор ретинального смещения складывается с правонаправленным вектором поворота головы, то получается перцептивное смещение предметов сцены вправо, причем с удвоенной скоростью. В итоге, ни о какой константности расположения и речи нет, чем и объясняется стрэттоново «дрожание».

В отличие от перевернутости, вызванной инвертирующими очками, оказалось, что к эффекту дрожания можно полностью адаптироваться. После нескольких дней практики, константность расположения вошла в норму, т.е. дрожание поля при поворотах головы исчезло. В конце эксперимента, когда ученый снял наконец очки, он заметил возврат эффекта дрожания, даже несмотря на то, что оптическая проекция внешнего мира на сетчатку стала естественной. Итак, остается неясным, по-





**Рис. 7.5.2.** Адаптация при активном движении и пассивном. Испытуемым первой группы дали возможность обследовать сцену в искажающих очках; наблюдатели второй группы видели ее через эти же очки лишь пассивно. Наблюдатели первой группы (активные) полнее адаптировались к искажениям, нежели наблюдатели пассивные.

По материалам Held, 1965.

чему адаптация к инверсии оказывается слабее и менее полной, чем адаптация к ориентации стимула.

В прочих исследованиях перцептивной адаптации применялись иные оптические искажения, в частности: повороты, увеличения и уменьшения, зеркальность, латеральное смещение, изменение кривизны и т.д. (Rock, 1966; Welch, 1978, 1986). Обнаружилось, что уровень адаптации и ее время варьируют в разных вариантах оптического искажения. По началу наше перцептивно-моторная интеграция совершенного негодная, но, спустя какое-то время — от минут до недель — при условии непрерывно искаженного воздействия, она поднимается до вполне удовлетворительного уровня. В некоторых случаях, например, при небольших латеральных смещениях ретинального

изображения, испытуемые сообщали, что их восприятие мира полностью восстановилось после адаптации. Действительно, они обычно говорили, что *чувствуют* свою руку там, где видят, но не там, где она на самом деле. Сей феномен, при котором зрение берет верх над другими сенсорными модальностями, с которыми конфликтует, получил название *зрительной доминанты*.

Однако сам факт адаптации ничего не говорит о том, какова именно адаптация в данной ситуации: зрительная или моторная? Дальнейшие эксперименты показали, что ответ зависит от специфики оптического искажения, но при этом работа моторной и зрительной систем сочетанна, функции тесно переплетены, а эффекты зачастую аддитивны (Hamilton, 1964; Harris, 1965).

Одно из наиболее впечатляющих открытий, сделанных при исследовании феномена адаптации, таково, что ее величина оказалась зависима от характера деятельности испытуемого. Американский психолог Ричард Хельд (Массачусетский технологический институт) показал, что активное обследование окружающего пространства критично в адаптации к пространственным преобразованиям. Ученый изучал адаптацию у активных и у пассивных наблюдателей (Held & Freedman, 1963). Испытуемым вначале надевали очки, искривлявшие прямые линии (при этом мера этого искривления была заранее измерена). Затем половина из волонтеров активно двигалась внутри большого цилиндра, внутренняя поверхность которого была покрыта мелкими случайно разбросанными точками. Наблюдателей же второй группы катали внутри того же цилиндра на тележке (рис. 7.5.2 на пред. стр.). Проведенные затем измерения показали, что активные наблюдатели лучше адаптировались к пространственным искажениям в стимуле.

Хельдовы данные говорят о том, что контроль движений собственного тела в условиях оптических искажений играет важнейшую роль в адаптации. Важно отметить при этом, что без такого контроля для веридикальной перцепции нужны намного большие контрасты в зрительной стимуляции. Те наблюдатели, что могли контролировать свои движения, с радикально большей эффективностью определяли то, что, во-первых, стимул искажен, во-вторых — как он искажен. Тем, кого катали на тележке, это давалось с намного большим трудом.

## 7.6 ЧАСТИ

У зрительного восприятия объектов имеется еще один важный аспект, о котором нельзя говорить лишь в контексте веридикальной перцепции глобальных свойств этих объектов (размера, формы, ориентации и пространственного расположения): дело в том, что большинство сложных объектов воспринимаются нами как состоящие из отдельных частей.

В термин «часть» мы вкладываем следующий смысл:

- ⇒ Часть объекта — это отграниченный фрагмент этого объекта, имеющий полуавтономный и объектоподобный статус в зрительном восприятии.

К примеру, человеческое тело воспринимается состоящим из головы, туловища, рук и ног; обычное офисное кресло мы видим, как имеющее сидалище, спинку и четыре ножки; у дерева мы различаем ствол, ветви и множество листочков. Вдобавок к этим частям восприятие объектов несет в себе пространственные взаимоотношения между составляющими: скажем, сваленные в кучу ножки, спинка и сиденье никак не воспримутся нами как цельное кресло.

### 7.6.1 Признаки восприятия частей

Зрительный опыт свидетельствует, что восприятие сложного объекта запускает механизм самопроизвольного восприятия его частей. У нас также остается впечатление, что составляющие объекта стабильны и неслучайны. Более того, мы уверены, что и другие люди воспринимают их так же, как и мы. Здоровый наблюдатель, к примеру, никогда не увидит нижнюю часть головы и верхнюю часть туловища как самостоятельную часть тела или сидалище, а одну ножку — как самостоятельную часть кресла.

Доказательства тезиса о том, что большинство объектов мы воспринимаем состоящими из частей, собираются по нескольким направлениям — это: язык, на котором мы говорим об окружающем нас мире; специально подготовленные примеры; данные психологических опытов.

#### 7.6.1.1 Язык

Одним из наиболее очевидных аргументов в пользу перцептивного деления объектов на части являются слова, которыми мы маркируем эти части. Восприятие составляющих крайне важно и перманентно в нашей жизни, поэтому у нас имеются отдельные слова и словосочетания для очевидных частей многих важных для нас предметов: к нашему телу относятся такие термины как «голова», «туловище», «рука», «нога»; к стулу или креслу — «ножки» и «сиденье»; к дереву — «ствол», «крона», «ветви», «листья». Причем несмотря на то, что мы редко когда видим эти части отдельно от самих объектов. То есть, структура нашего языка отражает, лежащую в его основе структуру восприятия.

При этом мы не можем исключить и того, что в какой-то мере восприятие частей определяется речевыми обычаями нашей культуры. Сие еще одна версия гипотезы Сапира — Хорфа.<sup>1</sup>

Однако существуют два основных варианта опровержения этой гипотезы.

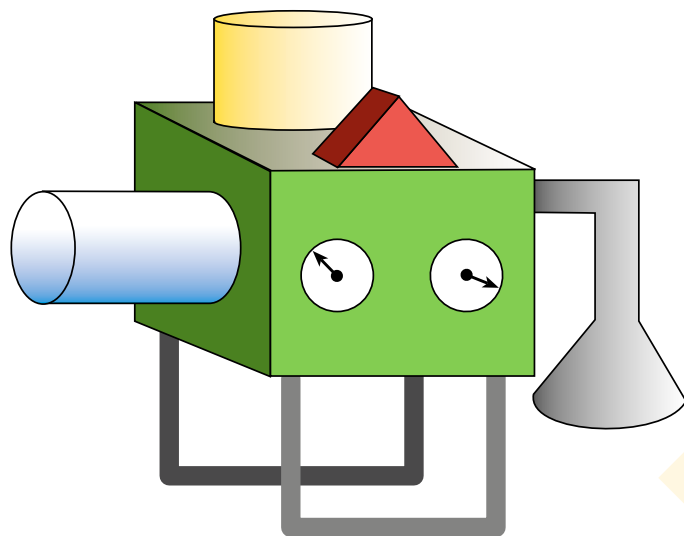
Первый звучит так, что во всех языках есть отдельные слова, относимые к голове, руке, ноге, но нет ни одного слова ни в одном из языков, означающего комбинацию из нижней части головы и верхней части туловища.

Вместе с тем, нельзя сказать, что вербализация видимых частей объектов систематична и универсальна во всех языках. В английском, например, есть отдельные лексические статьи, относимые к кисти и к предплечью, но нет таких, что относятся к верхней части руки.<sup>2</sup> В отношении ног мы оперируем разными словами: переднюю часть низа ноги назовем «голенью», а заднюю — «икрой», но общего слова, именующего низ ноги, у нас нет. В других языках эти части конечностей называются по-другому, но так или иначе, лингвистические структуры, относимые к объектам с устойчиво читаемыми частями, остаются относительно стабильными во всех языках мира, намного стабильнее, чем если бы долевая структура предметов вокруг нас являлась языковым производным.

Вторым контраргументом хорфовой гипотезе восприятия частей объектов является то, что мы самопроизвольно воспринимаем дольчатость, которых ранее никогда не видели и у нас нет слов, относимых к этим долям. Если бы восприятие частей опреде-

<sup>1</sup> Гипотезу Сапира — Хорфа мы уже обсуждали в контексте цветовых категорий (гл. 3): вспомним — идея состояла в том, что восприятие и мышление определены языком, а не наоборот.

<sup>2</sup> На языке анатомов эта часть тела называется «плечо» — *Прим. пер.*



**Рис. 7.6.1.** Неизвестный ранее объект. Несмотря на то, что мы никогда не видели его раньше, воспринимается он состоящим из четко обозначенных частей.

лялось нашим языком, мы понятия бы не имели о том, как делить новый объект на части и не смогли бы этого сделать. Давайте быстро взглянем на предмет, изображенный на рис. 7.6.1: мы сходу увидим, что он состоит из частей и легко выделим эти части. Эту способность можно объяснить только если восприятие частей в комплексном объекте — это природная функция перцепции как таковой, по большей части независимая от языка. Лингвистическая структура, безусловно, может иметь *некоторое* влияние на перцепцию, но основное влияние идет в обратном направлении: восприятие (совместно с использованием) позиционирует части, которым язык дает названия.

### 7.6.1.2 Феноменология

Прочие доказательства дольчатости в восприятии объектов можно найти среди некоторых весьма эффектных феноменологических демонстраций. Американцы Дональд Хоффман и Витман Ричардс (1984) дают ряд примеров, в которых даже очевидно цельные и единые перцепты имеют устойчиво долевою структуру.

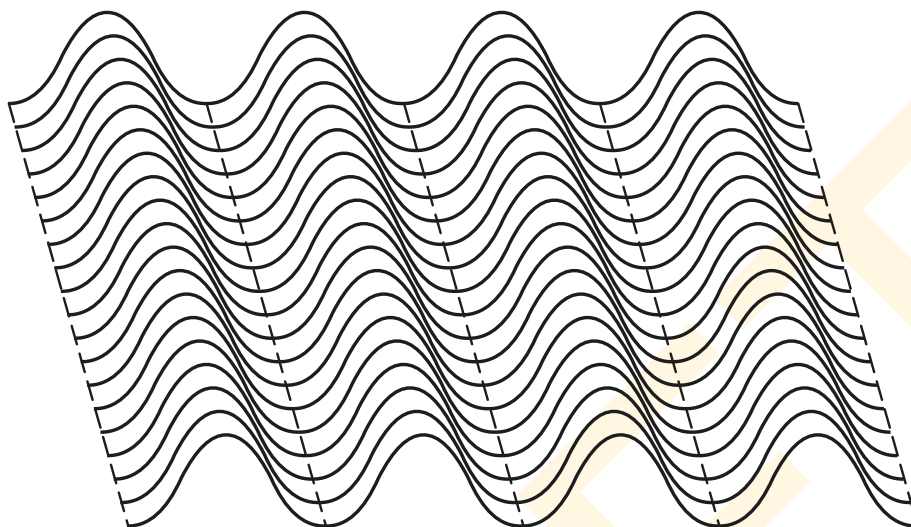
Давайте вначале поглядим на волновую поверхность рис. 7.6.2. Несмотря на то, что она полностью гладкая и непрерывная, выглядит эта поверхность собранной из множества параллельных гребешков (условно отделены друг от друга пунктирами). Если страницу перевернуть вверх ногами, то долевая организация изменится, и пунктиры теперь будут идти строго по середине гребней. Однако и в первом, и во втором случаях фигура упорно будет смотреться собранной из параллельных «гребней» («хребтов»), но никак не из «впадин» («желобов»). То есть, сей пример демонстрирует не только дольчатость структуры объекта, но и то, как эти доли (части) определены.

Еще одна демонстрация Хоффмана — Ричардса (1984) — это т.н. реверсивная лестница (рис. 7.6.3): две черных точки принадлежат одной общей ступеньке; однако фигура перцептивно реверсивна и, если ее перевернуть вверх ногами, то черные кружки окажутся расположенными на разных «ступеньках» (горизонтальная на одной «ступеньке», а вертикальная на «ступеньку» выше). Важно отметить, насколько сильно феноменологическое воздействие: в первом случае точки ложатся на одну общую часть предмета, а в другом — на разные.

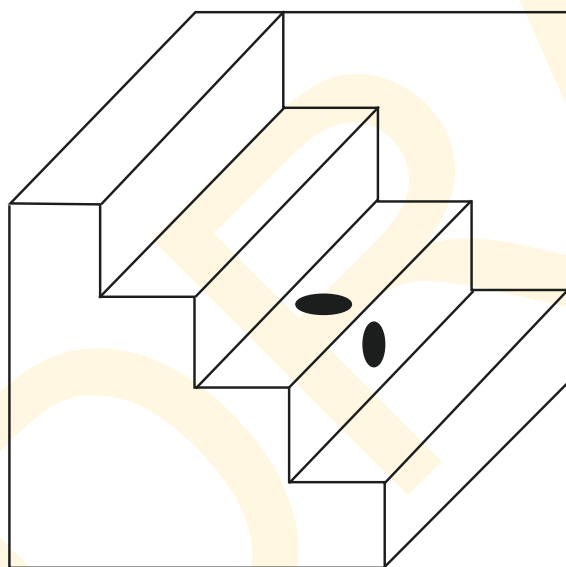
### 7.6.1.3 Перцептивные эксперименты

Следующую характеристику нашего восприятия частей объектов находим в различных психологических опытах. К примеру, Пальмер (1977) выполнил серию экспериментов по изучению восприятия долей простых двухмерных абстрактных фигур.





**Рис. 7.6.2.** Волновая поверхность. Несмотря на гладкость, фигура воспринимается состоящей из параллельных гребней (но не борозд). По крайней мере до тех пор, пока видится как бы сверху. Границы частей отмечены пунктирами.



**Рис. 7.6.3.** Реверсивная лестница Хоффмана — Ричардса. Исходно две черных точки видятся расположенными на одной ступеньке, но если фигуру перевернуть вверх тормашками, точки «окажутся» на разных ступеньках.

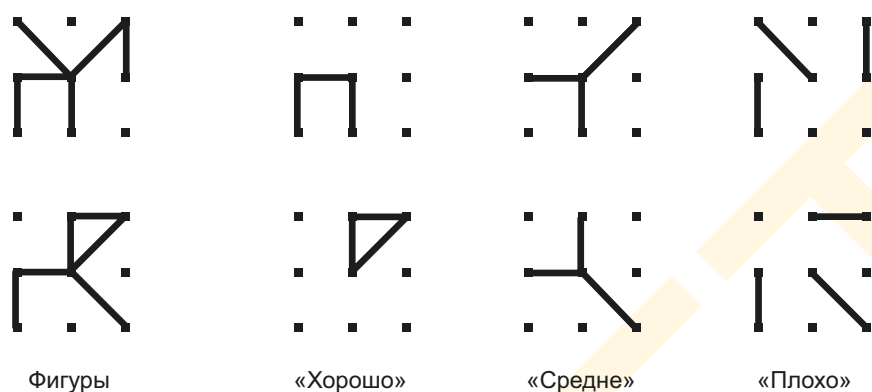
Результаты показали, что одни подмножества в стимуляции воспринимаются как части фигуры, а другие — нет.

Стимулы в опытах Пальмера представляли собой простые планарные фигуры, состоящие из шести линейных сегментов (рис. 7.6.4 «Фигуры»). В первом эксперименте испытуемых просили разделить фигурки, показанные слева на рис. 7.6.4 А, на их естественные части. Во втором — наблюдатели должны были охарактеризовать «натуральность» отдельных подмножеств по каждой фигуре. В третьем эксперименте требовалось максимально быстро решить включены или нет различные трехсегментные доли в состав той или иной шестисегментной фигуры (оценка скорости проводилась по трехбалльной шкале). Результаты всех трех опытов подтвердили гипотезу о том, что даже совершенно новые, ни разу не виденные предметы, воспринимаются нами как имеющие четко определенные части.

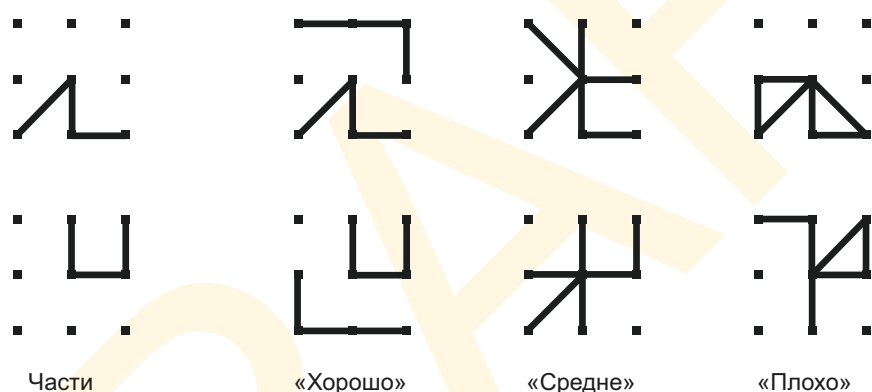
На рис. 7.6.5 А даны результаты пальмера опыта, в котором отдельные части требовалось обнаружить внутри одной и той же фигуры. Оценка велась по трех-

балльной шкале: «хорошо» означало, что испытуемый отыскивал нужную часть быстрее и точнее, чем при оценке «средне» и «плохо». Результат эксперимента четко ука-





А. Стимулы группы «Та же фигура»

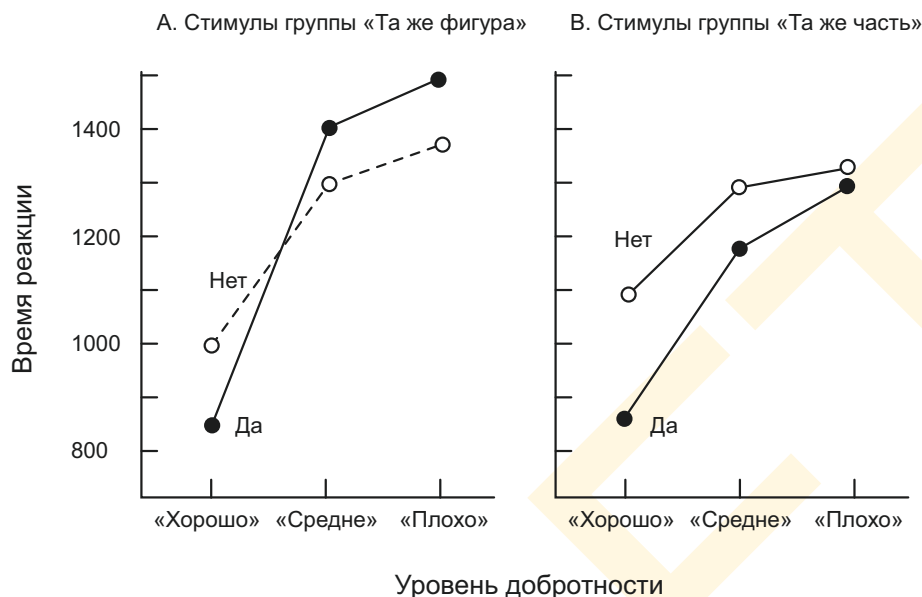


В. Стимулы группы «Та же часть»

**Рис. 7.6.4.** Стимулы эксперимента по детекции частей. А — перед наблюдателями была поставлена задача оценить «добротность» трехсегментных частей внутри шестисегментных фигур. В — задача выявить часть внутри целого.

По материалам Palmer, 1977.

зал на то, что «хорошо» относится к спонтанно воспринимаемым частям, а оценка «плохо» — наоборот (см. также Reed & Johnson, 1975). Возможно, однако, что результаты по условию «Та же фигура» — это следствие внутренних отличий между частями «хорошо» и «плохо», к примеру: части «хорошо» связны и компактны, тогда как «плохо» несвязны и рассредоточены. Чтобы исключить такую интерпретацию результатов и показать, что восприятие частей строго контекстно зависимо, в опыт были включены другие стимулы (точнее, в условие «Та же часть»), в которых ровно те же самые три сегмента вставлены в разные фигуры, что дало «хорошие», «средние» и «плохие» части (рис. 7.6.4 В). Как и следовало ожидать, время идентификации «хороших» частей оказалось меньше, чем у «средних» и «плохих», даже несмотря на то, что они в каком-то смысле были теми же самыми частями (рис. 7.6.5 В). Те же данные были получены и при решении испытуемыми задачи деления на части и их ранжирования частей. В



**Рис. 7.6.5.** Результаты эксперимента по детекции частей. Время верификации наличия частей внутри цельных фигур в обоих случаях растет по мере падения добротности (А и В).

По материалам Palmer, 1977.

итоге, исследователи пришли к выводу, что все новые фигуры «автоматически» воспринимаются нами как состоящие из самостоятельных.<sup>1</sup>

### 7.6.2 Долевая сегментация

Итак, приняв, что восприятие сложных объектов и поверхностей структурировано по частям, резонно задать себе вопрос: каким путем зрительная система выделяет эти части? Из приведенных выше данных следует, что дольчатость перцептивной структуры многих объектов строго закономерна и стабильна среди огромного числа наблюдателей. Каким же образом части объекта читаются нашим зрением из структурных характеристик этого объекта?

На самом общем уровне имеются два варианта ответа: деление объекта на части (как говорят — *парсинг*) по т.н. *формным примитивам* или же *по правилу границ*. После того, как зрение выделит набор частей в предмете (по первому или второму методам) начнется конструирование высокоуровневых частей на основе принципов, изложенных в разделе 6.1.

Первый подход к долевым сегментации требует наличия исходного комплекта простых, неделимых форм — примитивов, составляющих полное множество базовых ча-

<sup>1</sup> Пальмер (1977) смоделировал результаты своего эксперимента, используя линейное средневзвешенное значение нескольких гештальт-принципов группировки (таких как связность, непрерывность и сближенность — см. раздел 6.1), которые оценил количественно с помощью простейших математических действий. Спустя два десятилетия пальмеров подход углубил А. Шимаия (1997), применив много более сложные стимулы, включавшие окклюзию и зрительную завершенность. Модель Шимаия на основе некоей линейной комбинации семи различных факторов (многие из которых плотно соотносились с гештальт-принципами группировки) оказалась способна прогнозировать ответы испытуемых по зрительному выделению частей из фигур.

стей. Более сложные объекты анализируются затем как некие конфигурации этих примитивов. То есть, любой объект можно разделить на составляющие, из которых он собран. Сей подход выглядит аналогичным школьному делению рукописных слов на отдельные буквы: цельная линия слова делится точками между зрительно идентифицированными буквами. Однако удастся сие только тогда, когда имеется относительно небольшое множество примитивных компонент и лишь весьма отдаленно напоминает то, что происходит в повседневных сценах.

Подход на основе *правила границ* предполагает, что объекты делятся на части по границам между этими частями. Фиксированного множества формных примитивов не требуется — работа идет непосредственно с целым объектом. Сие похоже на попытку парсинга печатного слова, скажем, по принципу «разделить по U-образным линиям», но без попытки найти буквы. Теперь подробнее.

### 7.6.2.1 Формные примитивы

Идея, лежащая в основе парсинга по формным примитивам самоочевидна и привычна. Сие исключительно структуралистическая концепция, приложенная к зрительному восприятию объектов и их частей. Ее базовая теза такова, что любой перцепт формы объекта можно декомпонировать на относительно небольшое множество исходных *атомарных форм*. В неявном виде мы уже обсуждали эту идею, когда говорили о марровом опорном скетче, где в роли примитивов выступали отрезки линий, края, пятна и оконцовки (см. гл. 4). И, хотя такие примитивы вполне подходят двумерным проекциям, они с трудом справляются с ролью примитивов трехмерных. Самоочевидно то, что формный примитив трехмерного объекта должен быть именно трехмерным, то есть волюметричным, объемным, а не плоским или паче одномерным.

Отметим, что Марр (1982) активно проводил в жизнь бинфордову идею о том, что сложные формы можно разделить на примитивные составляющие из неких *типизированных* цилиндров. И хотя мы отложим обсуждение этой концепции до следующей главы, здесь все же напомним, что мы касались этого вопроса еще в конце второй главы, когда впервые описали марров количественный подход к зрению. В целом, бинфордова идея состоит в том, что форму можно представить как некую конструкцию из цилиндров соответствующих размеров и видов (рис. 2.4.8). Очевидно, что обычные цилиндры работают плохо, но если мы говорим о цилиндрах *типизированных*, их потенциал как примитивов радикально повышается. На рис. 7.6.6 даны несколько примеров типизированных цилиндров (часть А) и то, как из них можно собрать цельный объект обычной сцены (часть В). Важно отметить, что ежели у нас на руках имеется множество формных примитивов и некий алгоритм их детекции в ретинальном изображении, сложные трехмерные объекты можно сегментировать на составляющие из этих примитивов. При условии, что примитивы типизированы, доленая сегментация становится вполне возможной даже в отношении ранее не виденных объектов (рис. 7.6.1).

Однако у идеи доленой сегментации по формным примитивам есть ряд недостатков. Один из них — контекстуальное влияние на нее. Эксперименты, аналогичные пальмеровому (1977), показали, что одну и ту же часть объекта можно легко детектировать в одном фигуральном контексте и совершенно невозможно — в другом (см. рис. 7.6.4 В). Действительно, ежели доленая сегментация идет через поиск формных примитивов, то почему одну и ту же часть труднее распознать в одних случаях, и легче в других? Один из возможных ответов таков, что у зрения может иметься предпочитаемый им порядок детекции примитивов, то есть зрительная система в первую очередь ищет элементы определенного вида: если в объекте присутствуют несколько конкурирующих,

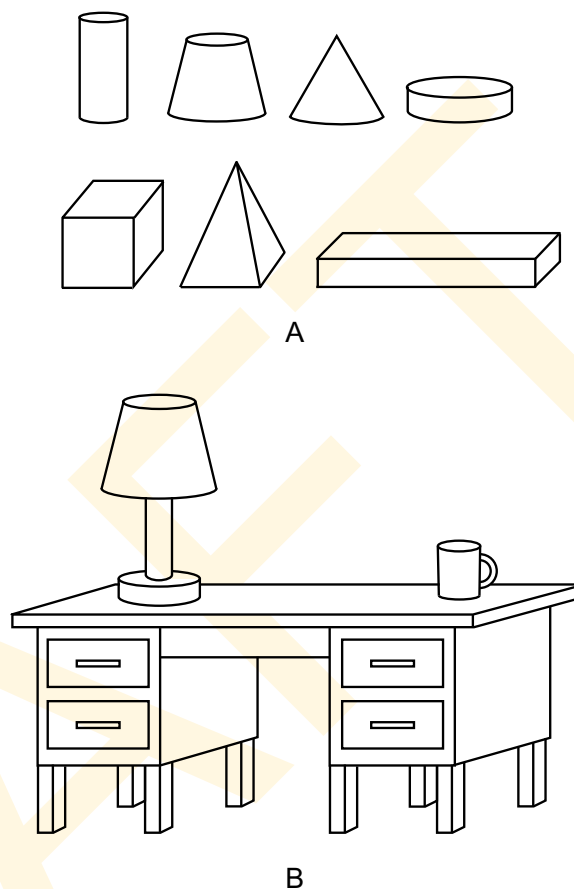
приоритетный будет распознан первым, а прочие окажутся «скрытыми». Правда, такие приоритеты до сих пор не выявлены — и сие, очевидно, весьма слабое звено в цепочке рассуждений.

Еще одним проблемным местом концепции формных примитивов является то, что мы воспринимаем объекты не только как состоящие из частей, но также и сами части как состоящие из подчастей: голова — часть тела, но и у самой головы есть глаза, нос, уши, рот и т.д. Более того, и сами подчасти обладают долевым структурой, например, глаза состоят из склеры, век, радужки, зрачка, ресниц и пр. То есть, теория долевого сегментации должна обладать способностью создавать потенциально сложную и продуманную *долевую иерархию* типа «часть — целое». Однако сие потенциальная проблема, поскольку лишь одно множество примитивов способно определить данный уровень долевого строения, но необходимо множество уровней.

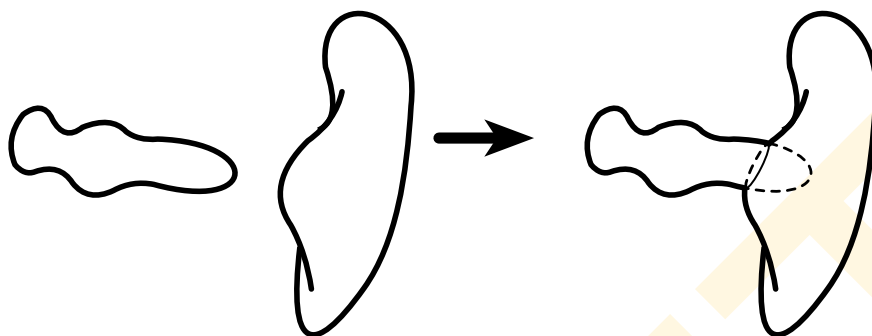
Описанную сложность можно разрешить двумя путями. Один из них — это предположить, что существуют некие *мультишкалы* детекции примитивов. Мы уже видели пример работы этой идеи в марровом дроблении частей тела цилиндрами разных пространственных разрешений (см. рис. 2.4.8): тело представлено цилиндром определенного размера, но рука — цилиндром большего пространственного разрешения, а каждый палец в отдельности — еще большего. Таким манером, различные примитивные части можно идентифицировать на различных уровнях пространственного разрешения, т.е. на различных шкалах, что в потенциале и дает иерархию этих частей.

Второй путь создания иерархии — это применение принципов группировки для объединения различных подмножеств примитивов в части более высокого уровня.

Однако еще одним скользким моментом в теории формных примитивов является то, что гипотетически должен иметь место некий четко определенный процесс их распознавания. Вопрос можно замять, когда речь идет о штриховом рисунке от руки (например, рис. 7.6.6), но от него никуда не деться при анализе ретинальных проекций реальных объектов, которые суть какофония яркостных уровней. Без набора процессов, способных найти исходные формные примитивы долевого сегментации по этим примитивам остается лишь изящной теоретической выдумкой.



**Рис. 7.6.6.** Типизированные цилиндры. Ряд текущих концепций построены на том, что повседневные объекты, в частности такие, что показаны в части В, можно разложить на формные примитивы, состоящие из типизированных цилиндров (примеры даны в А). Типизированные цилиндры варьируют по ряду параметров. (см. раздел 8.2.4).



**Рис. 7.6.7** Секущая регулярность. Когда одна поверхность проникает в другую, место их взаимодействия образует т.н. вогнутые прерывности. По идее Хоффмана — Ричардса прерывности эти играют ключевую роль в зрительном парсинге.

По материалам Hoffman & Richards, 1984.

Но самым сложным вопросом остается следующий: каким должен быть набор примитивов, чтобы описать все огромное разнообразие форм? Очевидно, не приходится рассчитывать на то, что простые геометрические фигуры, такие как цилиндры, конусы, эллипсоиды и призмы справятся с задачей описания гладких и тонких органических форм, например, лица человека. Правда, сие вовсе не значит, что примитивов, лежащих в основе восприятия частей объектов и самих объектов не существует, но только то, что отыскать их очень тяжело. Поэтому и остается лишь надеяться на то, что когда-нибудь они будут найдены.

### 7.6.2.2 Правило границ

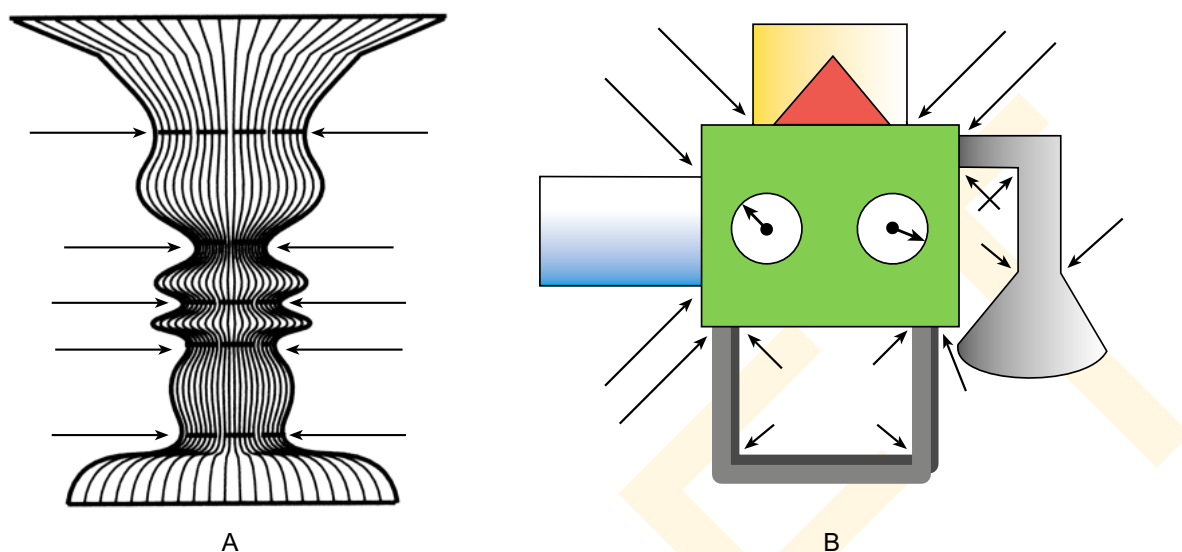
Вторым заслуживающим внимания концептом долевого сегментации является определение набора правил, описывающих границы между частями объекта. Сей подход не требует множества примитивных форм, поскольку отыскивает границы этих форм, и, следовательно, делит объект на составляющие, т.е. без предварительного описания характера частей. В логике формных примитивов части первичны, а границы между ними — плод идентификации этих частей. В правиле границ всё наоборот: границы первичны, а части — следствие детекции этих границ.

Самая широко известная концепция долевого сегментации на базе правила границ, принадлежит американцам Дональду Хоффману и Витману Ричардсу (1984). Ученые выстроили свой анализ на геометрических свойствах сложносоставных объектов, исходя из идеи проникновения одного в другой. Ключевой признак, названный ими *секущей регулярностью*, отнесли к тому факту, что когда один объект проникает в другой (рис. 7.6.7), то появляются т.н. *вогнутые прерывности* — участки, где поверхность объекта теряет гладкость и резко изгибается внутрь. Вогнутые прерывности образуются, если мы, например, сунем спичку в хлебный мякиш или свечку в шоколадную глазурь праздничного пирога.

Итак, в качестве первого шага к созданию концепта долевого сегментации Хоффман и Ричардс (1984) предложили правило вогнутой прерывности:

- ⇒ Зрительная система делит объекты на составляющие в тех местах, где имеются резкие изменения ориентации поверхности внутрь объекта, то есть — вогнутые прерывности этой поверхности.





**Рис. 7.6.8** Парсинг объекта по правилу границ. Хоффман и Ричардс (1984) предположили, что объекты можно делить на составляющие как минимум по криватурам гладких поверхностей (А), либо по вогнутым прерывностям (В).

Сие, кстати, объясняет то, почему две точки на неопределенной лестнице (рис. 7.6.3) выглядят расположенными на общей ступеньке при одном варианте интерпретации и на разных — при другом: лестница скорее поделится на части вдоль острых вогнутых углов между ступеньками, нежели вдоль выпуклых углов самих ступенек или плоских участков между. Если фигуру перевернуть вверх ногами, вогнутость/выпуклость каждого края обратится, и, следовательно, обратится долевая структура объекта.

Однако, строго говоря, секущая регулярность не применима к большинству случаев, проблема в том, что у многих объектов, в том числе у тела человека и разного рода неровных поверхностей переходы между частями весьма плавны (рис. 7.6.2). Хоффман и Ричардс тщательно проанализировали то, что происходит на этих переходах и переработали свой критерий секущей регулярности, придя, в итоге, к *правилу глубокой вогнутости*:

- ⇒ Поверхность перцептивно делится на составляющие в тех местах, где максимально искривлена внутрь (вогнута), даже если вогнутость эта (вопреки правилу вогнутой прерывности) плавна и непрерывна.

Пример приложения правила глубокой вогнутости к гладким поверхностям дан на рис. 7.6.8 А: судя по всему, парсинг этого объекта зрительной системой осуществляется тем же путем. Отметим, что правило глубокой вогнутости применимо и к волновой поверхности рис. 7.6.2: здесь оно работает по пунктирным линиям, т.к. они — суть места наибольших вогнутостей (если исходить из того, что наблюдатель смотрит снизу, то есть так, что «объект» лежит под поверхностью). И вновь теория прогнозирует то, что обращение вогнутости и выпуклости должно обращать долевую структуру объекта (что и происходит, когда фигуру переворачивают вверх ногами).

Как было показано, оба правила применимы скорее к фактическим поверхностям реальных объектов, нежели к двумерным ретинальным проекциям. Поскольку зрительная система не имеет прямо доступа к объектам внешнего мира, но только к их

проекциям, мы должны разобраться в том, как правила, касаемые вогнутостей *объектов*, можно транслировать в эвристику, работающую непосредственно с ретинальными изображениями. Хоффман и Ричардс предположили, что наиболее подходящим способом будет деление силуэта объекта на доли по максимальным вогнутостям, то есть, как скажут математики, — по локальным отрицательным минимумам кривизны.

Трудностью в практическом применении концепции Хоффмана — Ричардса является то, что она лишь находит точки, по которым идет деление объекта на перцептивные составляющие, но ничего не говорит о том, на какой паре точек это деление должно завершаться. Заметим, что примеры ученых были подобраны так, чтобы максимально облегчить собственно исследовательскую задачу, но нужно помнить при этом, что с реальными объектами работать намного сложнее. Пробел был заполнен сравнительно недавно через согласование правила глубокой вогнутости с множеством дополнительных правил, одно из которых — это т.н. разрез между ближайшими точками деления (Sing, Seyranian & Hoffman).

### 7.6.3 Общее и частное в нейронной обработке

Итак, ежели мы приняли то, что объекты реальных сцен воспринимаются нами как состоящие из частей, подчастей и т.д. возникает естественный вопрос: Какой из уровней приоритетен в перцептивной обработке? То есть, что первично — восприятие целого или частей? Несмотря на то, что гештальт-психология фактически никогда открыто не ставила этот вопрос, лежащая в ее основе доктрина холизма автоматически предполагает то, что восприятие целого берет верх над частностями. Сие, в свою очередь, значит что целое может обрабатываться раньше, благодаря чему и доминирует в результирующем перцепте. Однако почти все прочие подходы к восприятию оппонентны этой идее, и предполагают, что целое строится интеграцией мелких частей в крупные конгломераты. То есть, в логике таких воззрений части обрабатываются первыми. Какая же из точек зрения верна?

Как это часто бывает, ответ зависит от того, как задан вопрос. С точки зрения физиолога, безусловно, отдельные составляющие обрабатываются до целого. Как отмечалось в разделе 1.3, ретинальные рецепторы откликаются на чрезвычайно мелкие участки стимуляции, но если мы отследим ход информации по проводящим путям, то обнаружим, что рецептивные поля реагируют уже на более крупные и более сложные конфигурации стимулов. С такой (сугубо физиологической) точки зрения самоочевидно, что обработка зрительной информации идет от локальной стимуляции к общим структурам.

Однако если мы обратимся к своим зрительным ощущениям, то вряд ли сходу примем описанную выше логику. То есть, совершенно необязательно, что порядок, в котором зрительная обработка *начиналась*, определит результирующее зрительное ощущение, — скорее тот порядок, в котором она *завершилась*. Проще говоря, совершенно необязательно то, что нейронный процессинг идет непременно от частного к общему. Действительно, крайне маловероятно, что поток нейронной информации сугубо однонаправлен и течет исключительно от сенсорной поверхности глаза в высшие отделы мозга: данные гистофизиологии свидетельствуют, что в зрительной системе существует огромный массив обратных проекций из высших центров к низшим. Словом, очень высока вероятность существования обратной связи в зрительной системе, но никто толком до сих пор не знает, в какой форме реализована эта связь и за какую функцию она отвечает.

Еще одной сложностью в применении физиологических рассуждений к вопросу о

наших зрительных ощущений является спорность утверждения о том, что между порядком нейронного отклика и порядком возникновения зрительного ощущения можно поставить знак равенства. Можно ли быть уверенным в том, что факт первичной активности нейронов сетчатки в ответ на световую стимуляцию означает, что информация эта первой выводится и в ощущение? Более чем спорно. В 13 главе мы покажем, что ретинальная информация транслируется в перцепт далеко не напрямую. В отличие от любой формы зрительного ощущения, отклик фоторецепторов тотально и полностью дезорганизован и возникает *до* появления какой бы то ни было формы перцептивной константности.

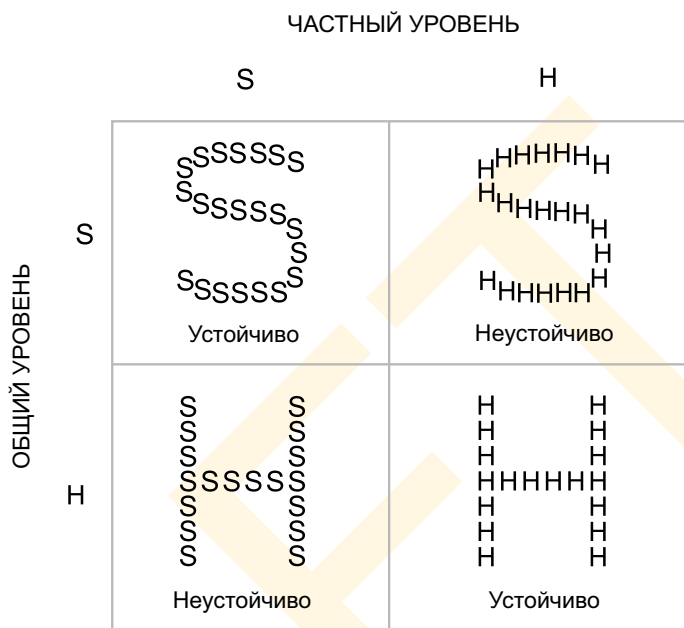
Вопрос о том, где же в недрах зрительной системы возникает осознанное ощущение (выраженное в нейронной логике), остается открытым, но, вполне вероятно, что образуется оно в разные периоды обработки, большая часть из которых, вероятнее всего — неосознанна. Осознанное восприятие может идти даже в направлении, обратном ходу нейронных процессов — от целого к частному, то есть, согласно холистической гештальт-доктрине.

Теперь рассмотрим некоторые данные, полученные из психологических опытов, убедительно свидетельствующих о том, что существует очень высокая вероятность того, что перцепт глобальных свойств объектов предшествует появлению перцептов их частей.

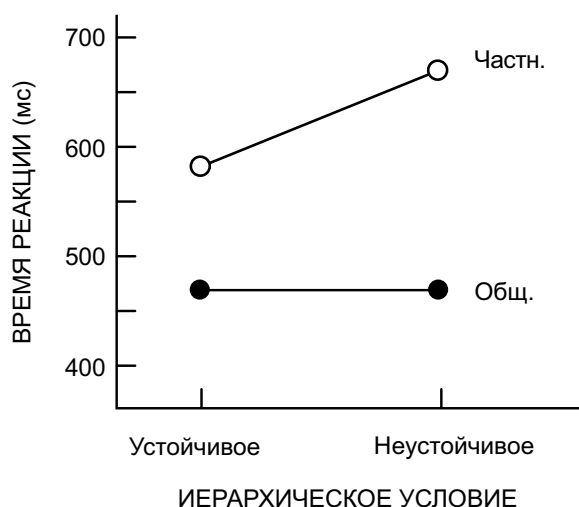
### 7.6.3.1. Приоритет целого

Взяв за основу рассуждений иерархичность долевого строения объектов, мы ставим вполне закономерный вопрос о перцептивной приоритетности целого над частным. Если стимул структурирован по двухуровневой иерархии, то нас интересует, что воспримется первым — объект в целом (общий уровень) или же его части (локальный уровень)? Психолог Дэвид Нэвон (университет Хайфы) исследовал этот вопрос, используя стимулы, в которых глобальный и локальный уровни несли в себе исключительно одну и ту же информацию: он попросту изобразил крупные литеры, состоящими из таких же мелких (рис. 7.6.9).

Нэвон (1977) предъявлял эти литеры своим испытуемым в двух конфигурациях — в *устойчивой* и *неустойчивой*. В устойчивой крупные и мелкие литеры были одни и те же, то есть, литера «Н» состояла из мелких «Н», а крупная «S» — из мелких «S». В неустойчивой конфигурации крупные и мелкие буквы конфликтовали: «Н» была построена из мелких «S» и наоборот. Наблюдателям было дано задание отловить идентич-



**Рис. 7.6.9** Иерархические стимулы. В ходе эксперимента испытуемые должны были выявить идентичность литер, представленных как на общем, так и на частном уровнях. По материалам Navon, 1977.



**Рис. 7.6.10.** Результаты опыта по перцепции иерархических стимулов. Идентификация общих литер (сплошные кружки) проходит и быстрее и не подвержена влиянию литер локальных. Идентификация локальных литер (полые кружки) идет медленнее, особенно когда идентичность глобальных литер конфликтует с таковой у литер локальных.

По материалам Navon, 1977.

ся первыми. Мы говорим при этом об *утрате локально-глобальной интерференции*.

Если же приоритетны признаки локального уровня, то мы прогнозируем обратные результаты: быстрейший отклик на локальном уровне (доминанта частного), замедление отклика при переходе на глобальный уровень в неустойчивых условиях (локально-глобальная интерференция) и отсутствие замедления откликов при переходе на локальный уровень в неустойчивых условиях (утрата глобально-локальной интерференции).

Результаты нэвонова опыта даны на рис. 7.6.10: мы видим, что по всем трем позициям они подтверждают идею приоритета общего: время отклика оказалось короче к общим литерам, а в неустойчивых условиях испытуемые «подтормаживались», только когда пытались перейти на локальный уровень.

Данные Нэвона убедительно свидетельствуют о том, что перцептивные процессы идут от общей обработки к частной, то есть ко все более и более тонкому анализу.

Дальнейшие исследования показали, что приоритет общего — это реальный и повторяемый феномен, но при этом не настолько мощный как казалось вначале. В определенных условиях стимуляции возникает иная картина. Р. Кинчла и Дж. Вольфи (Принстонский университет, 1979), например, обнаружили, что скорость именования локальных и общих форм зависит от их ретинальных габаритов: идентификация общих литер идет быстрее таковой у мелких, когда общий стимул оказывается меньше  $8 - 10^\circ$  зрительного угла, но при этом идентификация мелких срабатывает быстрее, когда общий стимул оказывается крупнее этого значения.

Кинчла и Вольфи показали, что зрительная система держится середины, начиная работать со структурами средних пространственных частот и продолжая анализ в обе стороны — в сторону низких и в сторону высоких. Сие очевидно согласуется с контрастно-сенситивной функцией человека (см. раздел 4.2), по которой мы четко видим,

ность литер на локальном и глобальном уровнях. Время отклика и точность измерялись в четырех условиях стимуляции (рис. 7.6.9).

Если объектные свойства общего уровня воспринимаются первыми, то:

1. Отклик на общем уровне должен возникать быстрее, то есть, мы говорим о *доминанте целого* на частным.

2. При переходе наблюдателя с общего уровня на частный отклик на неустойчивые общие литеры начинает тормозиться (поскольку признаки локального уровня воспринимаются только после признаков общего). Мы говорим при этом о *глобально-локальной интерференции*.

3. При переходе наблюдателя на общий уровень, неустойчивые локальные литеры не должны тормозить отклик, поскольку признаки общего уровня воспринимаются

что зрение наше наиболее чувствительно к стимулам средних пространственных частот и менее — к большим и малым. При этом глобальный приоритет фигур порядка  $8 - 10^\circ$  зрительного угла (т.е. размером примерно с кулак на вытянутую руку) — это весьма неожиданный результат, учитывая то, что рецептивные поля простых клеток V1 редко когда в размере своем превышают  $1^\circ$ .

Прочие опыты показали, что обработка на общих и локальных уровнях идет скорее параллельно, нежели последовательно. К примеру, когда испытуемые отслеживают целевую литеру на обоих уровнях (глобальном и локальном), они быстрее откликаются, если целевая литера представлена на обоих уровнях, но не на каком-то одном (Miller, 1981). Тщательный анализ времен откликов укрепил ученых во мнении, что обработка на обоих уровнях идет параллельно, и что результаты обоих процессов объединяются в единый отклик. Если сие верно, приоритет общего можно толковать как результат параллельной обработки по разным пространственно-частотным каналам. При этом по одним обработка идет чуть быстрее, чем по другим. Следовательно, тезис о том, что зрительная система последовательно отрабатывает уровни — неверен. Таким образом, приоритет общего — это открытие, подтверждающее примат глобальной информации в осознанной перцепции, но интерпретация этого факта не столь прямолинейна, как казалось на первый взгляд.

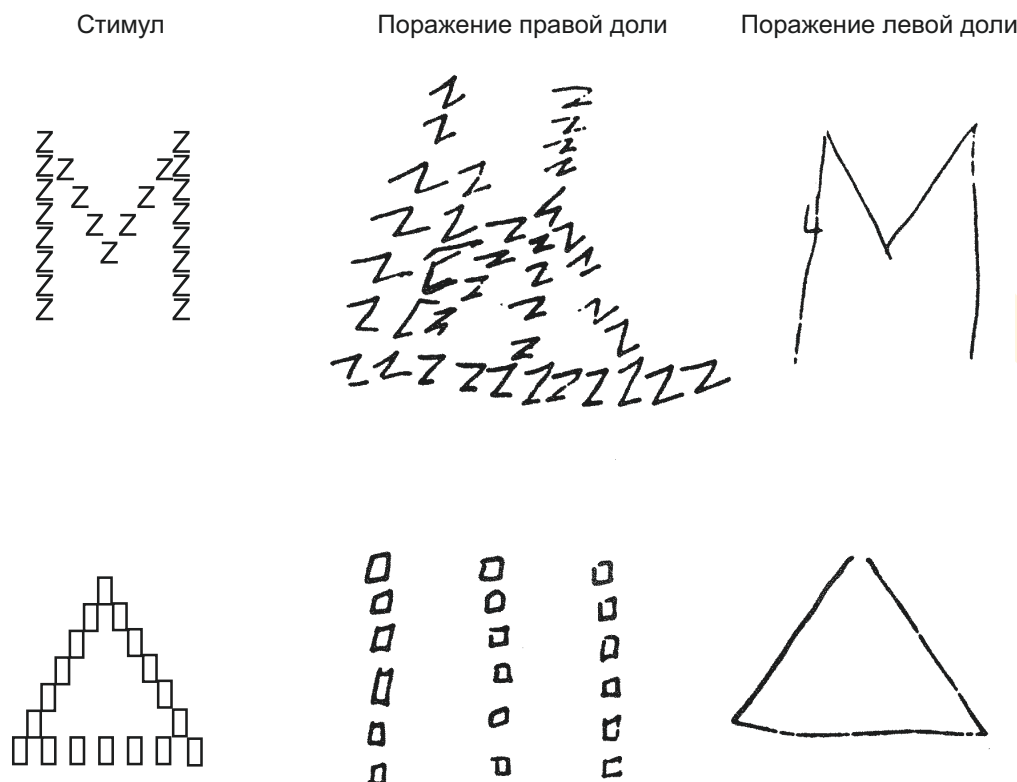
По данным Линны Робертсон, наблюдавшей за больными с травмами головного мозга, глобальная и локальная информации обрабатываются отдельно левым и правым полушариями. Было показано, что общие перцептивные свойства объектов обчисляются в правой доле, тогда как частные — в левой (Robertson, Lamb & Knight, 1988). Рис. 7.6.11 демонстрирует, как больные с поражением левого и правого темпорально-париетального участков коры перерисовывают один и тот же иерархически структурированный стимул (Delis, Robertson & Efron, 1986): больные с поражением правого полушария, то есть страдающие недостаточностью общей обработки, способны воспроизвести маленькие литеры (образуют большúю), но не могут организовать их в общую структуру (средняя колонка на рис. 7.6.11); пациенты с поражением левого полушария, т.е. страдающие недостаточностью локального процессинга, легко воспроизвели саму литеру, т.е. общую структуру стимула, но оказались не в силах нарисовать мелкие образующие ее литеры (правая колонка рис. 7.6.11).

Как легко догадаться, пациенты с поражением правой височно-теменной доли продемонстрировали устойчивую способность успешно решать нзвонову задачу идентификации, тогда как больные с поражением левой височно-теменной области легко решали задания общего плана (Robertson и колл., 1988). Исследования со здоровыми людьми, построенные по специальным методикам, позволяющим оценить степень полушарной асимметрии, привели исследователей к тем же выводам: информация общего плана эффективнее обрабатывается в правом полушарии, тогда как локальная — в левом (Sergent, 1982).

Неожиданным однако оказалось то, что ни в первой, ни во второй группах не наблюдалось перекрестных эффектов (Lamb, Robertson & Knight, 1989). И сие говорит о том, что в основе эффектов глобально-локальной конкуренции (например того, какой уровень идентифицируется быстрее), а также эффектов глобально-локальной и локально-глобальной интерференции (например того, какой из уровней подавляет другой уровень) лежат совершенно разные нейронные механизмы.

Дополнительные исследования, в частности, проведенные с участием пациентов с рассеченным мозолистым телом мозга, пролили некоторый свет на сие неожиданное открытие. В частности, стало понятно, что обработка информации, идущая в левой и в правой височно-теменной коре критична в действии эффектов глобально-локаль-





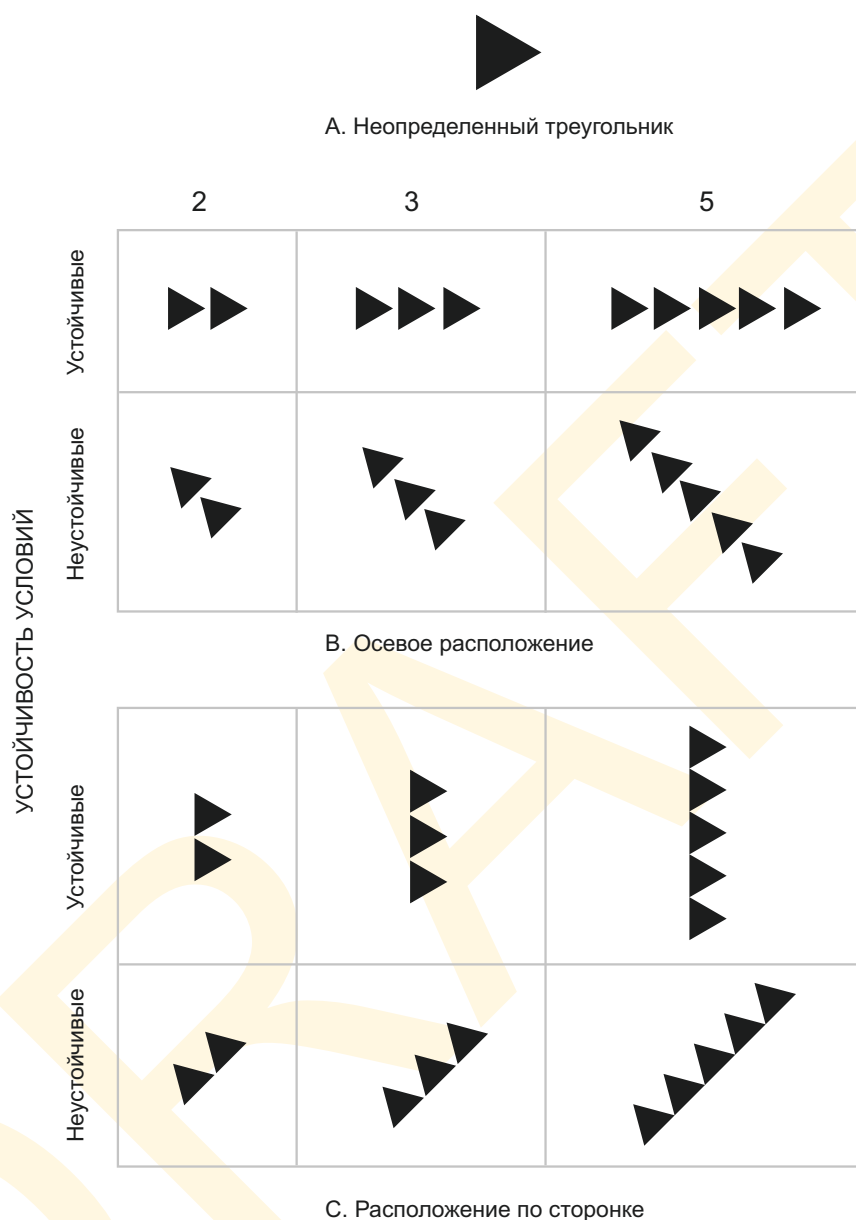
**Рис. 7.6.11** Перерисовка иерархически организованных фигур больными с травмой головного мозга. Больные с поражением правого полушария точно изображали локальные элементы, но не в силах были передать общую конфигурацию стимула (центральная колонка). Больные с поражением левого легко воспроизвели общую структуру стимула, но не могли изобразить его локальные элементы (правая колонка).

По материалам Delis, Robertson, & Efron.

ной конкуренции, тогда как передача информации между этими двумя корковыми областями через мозолистое тело — критична в действии эффектов интерференции (Robertson & Lamb, 1991). Очевидно при этом, что еще предстоит ответить на множество вопросов, прежде чем мы разберемся в механизмах, лежащих в основе комплексных эффектов. Однако уже сегодня у нас на руках ряд неожиданных и очень интересных нейрофизиологических данных, наподобие описанных выше.

### 7.6.3.2 Конфигурантно-ориентационные эффекты

Прочие данные, свидетельствующие о существенном влиянии общей структуры на восприятие частностей, получены в серийных экспериментах Пальмера и названы им *конфигурантно-ориентационными эффектами*. Оказалось, что на восприятие локальной пространственной ориентации влияет общая ориентация стимулирующей структуры (Palmer 1980, 1989). Вспомним, что треугольники Эттнива (1968) были неопределенными по направлению: мы их видели указывающими на одно из трех возможных направлений. К примеру, треугольник на рис. 7.6.12 А можно видеть как направленный на 3, на 7 или на 11 часов. Интересным фактом в контексте влияния общего на



**Рис. 7.6.12.** Восприятие направления. Сам по себе равносторонний треугольник (А) может восприниматься нами как указывающий на одно из трех возможных направлений. Расположение нескольких таких треугольников вдоль их осей симметрии (В) или же по одной из сторон (С) жестко влияет на перцепт направления этих треугольников. Подробнее в тексте.

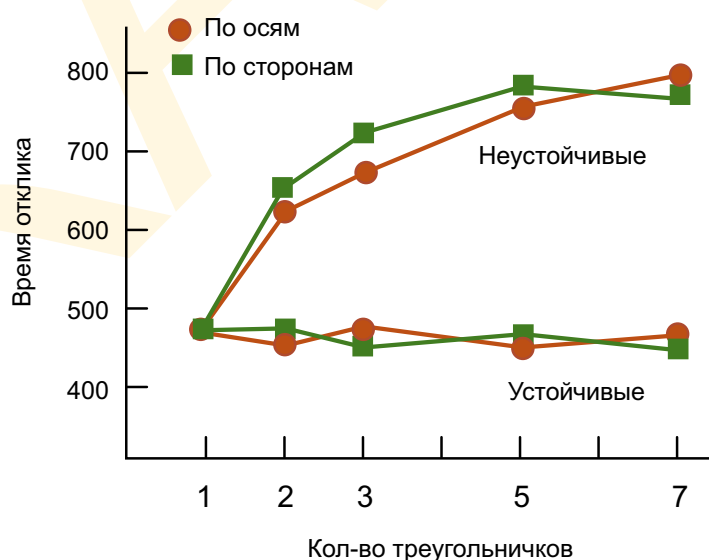
перцепцию частного является то, что ежели расположить треугольники определенным образом, конфигурация начинает жестко влиять на перцептивное направление каждого в отдельности: когда фигурки расположены вдоль оси симметрии (рис. 7.6.12 В), наблюдатели склонны видеть их указательную вершину по линии конфигурации; если же они расставлены вдоль одной из сторон (рис. 7.6.12 С) — направление указателя окажется перпендикулярным ей.

С. Пальмер и Н. Бучер (1981) исследовали то, насколько влияние общего зависимо от местоположения стимуляции, то есть, ученые захотели ответить на вопрос: влияет

ли конфигурация на перцепт ориентации по всей сетчатке или же ее действие ограничено определенной ретиальной локалью? Для ответа на него понадобилась методика измерения силы конфигурального эффекта. Поэтому исследователи придумали некое задание на скорость, в котором испытуемым по отдельности предъявлялись треугольники, каждый из которых указывал либо строго влево (на 9 часов), либо строго вправо (на 3 часа) — рис. 7.6.12. Наблюдателей попросили максимально быстро (но стараясь при этом не делать ошибок) нажимать на левую кнопку, если они видят левонаправленный треугольник и на правую — если правонаправленный. В одном из сеансов треугольники располагались так, чтобы облегчить восприятие направления (верхние ряды рис. 7.6.12 В и 7.6.12 С). Условия этого сеанса окрестили «устойчивыми», и время отклика ожидалось весьма коротким. В других сеансах были организованы помехи восприятию направления (нижние ряды рис. 7.6.12 В и 7.6.12 С) и эти условия называли «неустойчивыми». Ожидалось при этом, что время отклика удлинится. Отличия во времени отклика должны были выступить мерой того, насколько сильно ориентация стимула влияет на восприятие направления каждого треугольника в отдельности.

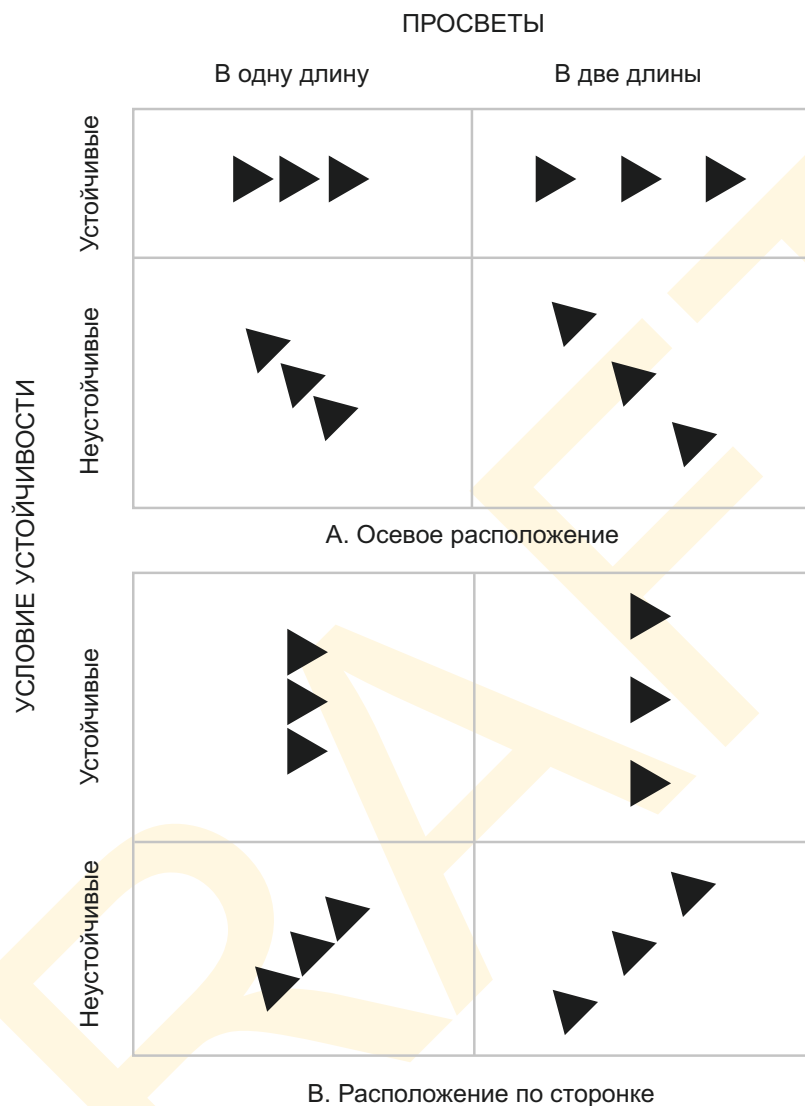
Ответ на вопрос о том, каковы эффекты общей ориентации группы фигур, ученые попытались получить в серии опытов. Во-первых, они взяли величину конфигурального влияния как функцию от числа треугольников в стимуле, используя при этом стимулирующие комплекты из одного, двух, трех, пяти и семи треугольничков (рис. 7.6.12 В и 7.6.12 С). Исследователи рассудили так, что ежели влияние конфигурации действительно глобально, добавка треугольников должна повышать степень эффекта. Если же эффект ограничен пространственно, то добавка треугольников за пределами критического участка не окажет влияния на восприятие. Результаты, показанные на рис. 7.6.13, говорят о том, что время отклика в неустойчивых условиях стимуляции быстро растет по мере прибавки треугольничков, но выходит на плато, начиная с пяти. Сие свидетельствует в пользу той точки зрения, что конфигуральный эффект возникает в результате ограниченного пространственного взаимодействия в рамках определенного участка.

Во втором опыте исследователи «играли» просветом между треугольниками в триплете (рис. 7.6.14), рассудив при этом так, что ежели существует ограниченный участок перцептивного взаимодействия между треугольничками, влияние конфигурации должно пропасть, когда расстояние между ними превысит некое кри-



**Рис. 7.6.13.** Результаты опыта по восприятию направления неопределенных треугольничков. Время, которое уходит на распознавание ориентации треугольника строго зависит от ориентации конфигурации, частью которой является этот треугольник. Разница между устойчивой и неустойчивой конфигурациями растет по мере роста числа треугольничков, но выходит на плато примерно при пяти.

По материалам Palmer & Bucher, 1981.



**Рис. 7.6.14** Эффект разрежения в восприятии направления треугольников. Эффект практически исчезает, когда расстояние между центрами фигур достигает их удвоенной длины.

тическое значение. Результаты показали, что эффект быстро уменьшается по мере роста просветов, сходя в ноль, когда расстояния между треугольниками выходят на их двойную длину.

Третий эксперимент показал, что сей эффект управляется скорее относительным расстоянием между фигурами, нежели абсолютным расстоянием между их проекциями на сетчатку.

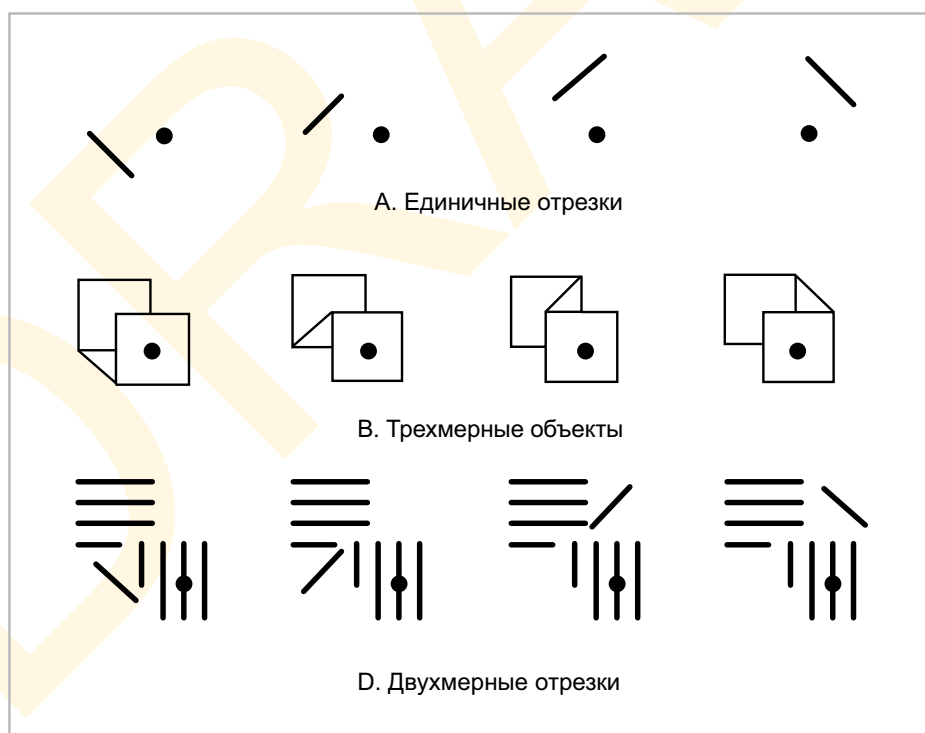
Данные трех опытов дружно свидетельствуют о том, что конфигуральные эффекты действительно оказывают сильное влияние на восприятие ориентации отдельных треугольников. При этом влияние это ограничено определенным участком взаимодействия, размер которого определяется относительным размером самих фигур, но не абсолютным ретинальным расстоянием.

### 7.6.3.3 Вербальная, объектная и конфигуральная доминанты

Прочие доказательства первичности общего в нашем восприятии мира получены из серии опытов, в которых выяснилось, что различение частей начинает доминировать, когда части эти образуют смысловое или четко структурированное целое. Реакция оказывается сильнее, чем в контрольных условиях, в которых те же части должны распознаться внутри бессмысленного или плохо структурированного контекста. Эффект оказался мощнее и в сравнении с попыткой распознавания этих же частей по отдельности.

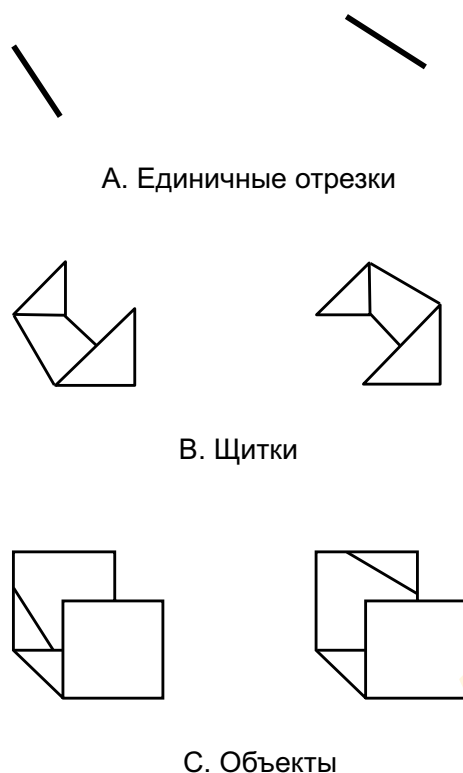
**Вербальная доминанта.** Первые данные об усилении перцепции частей внутри смыслового целого поступили из опытов Г. Рейчера (1969), исследовавшего т.н. *эффект вербальной доминанты*: оказалось, что наблюдатели лучше распознают единичные литеры, составляющие слово со смыслом, нежели когда они предъявлены сами по себе или в составе бессмысленной строчки. При этом наличие дополнительных литер в слове улучшает детекцию. Мы обсудим этот эффект (и то, как его можно истолковать теоретически) в гл. 9, когда будем говорить о моделях идентификации букв и слов.

**Объектная доминанта.** Н. Вайсшайн и К. Харрис (1974) доложили о сходном контекстуальном эффекте, названном ими *эффектом объектной доминанты*: испытуемые должны были определить, какая из фигур заключена внутри стимула, собранного из множества коротких отрезков. Некоторые из стимулов представляли собой простейшие трехмерные фигурки; прочие — набор линий (рис. 7.6.15). Испытуемые должны были найти четыре диагональных линии либо в свободном показе (рис. 7.6.15 А), либо внутри



**Рис. 7.6.15** Эффект объектной доминанты. Наблюдатели должны выделить четыре ориентации косых отрезков (А) в трехмерной конфигурации (В) и в двухмерной (С). В двухмерной конфигурации выявляемость оказалась на 20% ниже.





**Рис. 7.6.16.** Значимость структуры в эффекте объектной доминанты. Наблюдатели быстрее находят диагонали, когда те структурно значимы (B).



**Рис. 7.6.17.** Эффект конфигуральной доминанты. Наблюдатель лучше различает два диагональных отрезка внутри конфигурации, обладающей эмерджентными свойствами (B), чем когда они сами по себе (A).

возможного трехмерного объекта (рис. 7.6.15 B), либо внутри массива дезорганизованных отрезков (рис. 7.6.15 C). Вайсшайн и Харрис обнаружили, что наблюдатели лучше справляются с задачей в «объектных» условиях.

Последующие опыты показали, что эффект объектной доминанты касается не только собственно трехмерных стимулов, но может возникнуть даже если искомые линии принадлежат двумерной структуре, но *интерпретируемой* как трехмерная. Стимулы, подтверждающие этот вывод показаны на рис. 7.6.16 (McClelland & Miller, 1979). Испытуемые лучше находили линии только в тех условиях, в которых они меняли перцепт с прямостоящего щитка на опрокинутый (рис. 7.6.16 B). Прочие трехмерные условия, в которых они образовывали исключительно бессмысленные метки на другом объекте (рис. 7.6.16 C) не давали эффекта объектной доминанты.

**Конфигуральная доминанта.** Наконец, еще один пример доминанты в восприятии долей целого — это т.н. *эффект конфигуральной доминанты*, показанный на рис. 7.6.17 (Pomerantz, Sager & Stover, 1977). Испытуемых вновь просили отыскать простые элементы (часть A) в некоей конфигурации (часть B). Дабы убедиться в том, что экстраконфигуральные элементы сами по себе нельзя использовать при выполнении задачи различения, точно такие же элементы добавили к обеим целевым фигурам: на рис. 7.6.17 к целевой диагональной линии добавлен прямой угол, образующий стрелочный вертекс в первом случае и треугольник — во втором. Измерение времени отклика показало, что распознавание происходило быстрее внутри конфигурации, нежели изолированно.

Наиболее правдоподобным толкованием результатов опыта является следующее: диагональные отрезки объединяются с контекстуальными линиями, образуя *эмерджентные* (новоявленные) *признаки*, такие, в частности, как закрытые и открытые фигуры. Если эти признаки распознаются быстрее и легче, чем самостоятельные право- и леводиагональные линии, эффект конфигуральной доминанты можно объяснить. То есть, самоочевидно то, что как минимум некоторые глобальные эмер-

джентные свойства можно обработать быстрее, чем многие локальные. Однако же точный ответ на вопрос, почему эмерджентные признаки порой распознаются легче, чем их слагаемые — пока остается без ответа.

Эффекты словесной, объектной и конфигуральной доминант очевидно указывают на то, что решение перцептивных задач идет отнюдь не в локально-глобальном порядке, как представляется на базе одних лишь физиологических знаний. К примеру, маловероятно, что нейронная обработка таких конфигураций как треугольники и стрелки начинается с линий их составляющих (тем более, если учесть от факт, что размер рецептивных полей растет по мере движения сигнала от сетчатки к центральным отделам нервной системы). И уж совсем не стыкуется с нашими физиологическими представлениями то, что анализ перевернутости или неперевернутости трехмерных щитков из эксперимента Мак-Клелланда — Миллера (1979) проходит на самом деле до анализа ретинальной ориентации отрезков (последнее вероятно происходит где-то в V1, либо в нескольких синапсах за ее пределами).

Точная интерпретация описанных контекстуальных эффектов по сей день остается предметом бурного обсуждения. Возможно, что нейронная обработка идет-таки от частного к общему, но потом возвращается с холистического уровня назад на атомарный, что облегчает обработку локальных элементов. В 9-й главе мы рассмотрим модель, воплощающую эту идею в контексте эффекта вербальной доминанты.

Второй вариант объяснения звучит так, что хотя нейронная обработка идет от частного к общему, мы стремимся заполучить сознательный доступ к ее результатам, двигаясь в противоположном направлении, то есть, от общего к частному (Marcel, 1983 b).

Независимо от того, какой механизм в итоге будет найден, описанные выше эксперименты очевидно исключают всякую возможность того, что восприятие локальной структуры неизбежно предшествуют восприятию целого. Истина, по обыкновению своему, куда тоньше и интереснее.

**И**з всех воспринимаемых нами зрительно свойств объектов, форма, вероятно, главная, поскольку наиболее всего информативна: биологически форма объекта обладает наибольшей прогностической ценностью в отношении физических свойств этого объекта, нежели прочие его визуальные параметры. Вообразим, что нам известна лишь окраска некоего предмета. Какой нам прок от того, что это один из оттенков желтого? А теперь представим себе, что нам известна форма объекта, но более ничего. Распознав, к примеру, форму яблока, мы сможем очень многое понять об этом объекте: насколько он велик, какой вероятнее всего окраски, для чего сгодится, каков его вкус, где его можно добыть и т.д. Очевидно, что форма в данном случае выступает в роли наиважнейшего визуального признака объекта.

Однако на самом деле всё гораздо сложнее. О форме легко подумать как об унитарном свойстве<sup>1</sup> предметов, слагаемым из множества компонент, но очевидно должны существовать и иные унитарные свойства. К примеру, в гл. 3 мы разобрали то, что окраска поверхности оценивается нами по трем параметрам — цветовому тону, насыщенности и светлоте. И у такого анализа имеется ряд очевидных преимуществ перед унитарной точкой зрения. К примеру, сходство цветовых ощущений можно смоделировать сближенностью двух точек в трехмерном цветовом пространстве: точки, представляющие в этом пространстве подобные друг другу цвета, такие, например, как красный разной светлоты, расположатся рядом, тогда как точки красного и зеленого — вдали друг от друга.

Возможно ли провести сходный анализ формы по простейшим составляющим? Если да, то какие компоненты следует анализировать? Если нет, то можно ли форму как-то представить холистически? В любом случае, как определить меру сходства между двумя формами? Это лишь некоторые из вопросов, которые нам предстоит обсудить в текущей главе. Многие из них тесно связаны с темой перцептивной категоризации, которой мы коснемся в следующей главе (дело в том, что наиболее значимым свойством, определяющим категорию объекта, является его форма; следовательно, многие концепты объектной категоризации оказываются зависимыми от того, какое место в них уделено представлению формы).

Еще одной очевидностью в наших размышлениях о восприятии формы является тезис о том, что представление формы объекта можно идентифицировать по фасетному 2.5D-скетчу. Действительно, понятия, которые исходно применялись в литературе по машинному зрению в отношении 2.5D-скетчей, явно относятся и к текущему вопросу. Модули перехода от 2D-представления к представлению 2.5D часто называют «форма-из-яркости», «форма-из-движения» и т.п., как если бы 2.5D-скетч нес в себе представление формы. К сожалению, это не так. Хотя сие и справедливо в том смысле, что 2.5D-скетч *имплицитно* и *локально* несет в себе информацию о форме. Вспомним материал гл. 5: то, что фактически представлено в марровом 2.5D-скетче — это расстояние до локального фрагмента поверхности и его ориентация. Вспомним также, что

<sup>1</sup> Унитарный — объединённый, единый, составляющий одно целое, либо обозначающий, отражающий единство чего-либо. — *Прим. пер.*

2.5D-скетч организационно недифференцирован: он не представляет явном виде деление видимых поверхностей на отдельные объекты, еще меньше — части этих объектов и отношения между ними. И совсем не несет никакой информации о задних поверхностях объектов, тех поверхностях, что закрыты поверхностями передними. По всем этим причинам разумнее всего говорить, что в 2.5D-скетче отсутствует эксплицитное представление формы трехмерных объектов.

Как было сказано в конце 7-й главы, представление объектов часто задействует их деление на части — мы говорим, например, что лицо человека состоит из глаз, носа, ушей и т.д. Части, на которые перцептивно делится объект, пространственные взаимоотношения между этими частями, а также глобальные признаки их определяющие — все это важные аспекты их формы. И вновь мы скажем, что в явном виде информации об этих признаках в 2.5D-скетче нет. В зрительном восприятии формы объекта локальные фрагменты информации должны как-то организовываться в представления объектов, их признаков, частей, а также сложных взаимоотношений между ними внутри общей когерентной структуры. Понимание того, как всё это происходит — первейшая цель теории зрительного восприятия формы. В текущей главе мы разберем постулаты и перспективы этой теории.

## 8.1 ЭКВИВАЛЕНТНОСТЬ ФОРМ

Разговор о зрительном восприятии формы мы начнем с простейшей темы — вопроса *эквивалентности форм*. Вопрос эквивалентности форм включает в себя определение условий, в которых мы воспринимаем два разных объекта как обладающих одной и той же формой. Феномен восприятия эквивалентности форм теснейшим образом связан с классическим понятием константности формы. Вспомним (см. раздел 7.2), что:

- ⇒ Константность формы определена в логике того, что *один и тот же* объект обладает одной и той же перцептивной формой в самых разнообразных условиях просмотра.

Тогда как:

- ⇒ Эквивалентность форм касается вопроса о том, одинаковы ли формы двух разных объектов, несмотря на прочие пространственные отличия между ними?

К примеру, Boeing 747 в полный размер обладает той же формой, что и его масштабная модель на письменном столе, несмотря при этом на огромные отличия в абсолютном размере этих объектов; спортивная Cessna, бегущая по взлетной полосе, обладает совершенно той же перцептивной формой, что и перевернутая вверх брюхом при выполнении фигур высшего пилотажа.

### 8.1.1 Объективная форма

Разговор об эквивалентности форм мы начнем с разбора возможных трансформаций объекта, не затрагивающих при этом его *объективной формы* (Palmer, 1989). Приняв, что форма — это перцептивное свойство объектов, может показаться странным, что здесь мы говорим о форме как об «объективном» свойстве. Однако, разговор об объективной форме дает нам прочный базис рассуждений о ее зрительном восприятии, т.е. о *форме перцептивной*. Фактически, та мысль, что каждый объект обладает собствен-

ной объективной формой мало чем отлична от той очевидности, что каждый объект обладает объективными размером, расположением, ориентацией и т.д.

Рассмотрим, к примеру, простую геометрическую фигуру на рис. 8.1.1. Каждый из нас скажет, что фигурки в правой колонке имеют ту же форму, что и фигурки в левой, несмотря на отличия в их расположении, ориентации, размере и т.д. Сие подводит нас к определению объективной формы, которое прозвучит так:

⇒ Объективная форма — это пространственная структура объекта, которая остается неизменной, при выполнении в отношении этого объекта следующих трансформаций:<sup>1</sup>



А. Перенос



В. Поворот



С. Масштабир



Д. Отражение

**Рис. 8.1.1** Инвариантность формы при трансформациях подобия. У фигур, относимых к этим четырём видам трансформации или любой их комбинации — одна и та же объективная форма.

— *переноса*, при котором меняется только расположение объекта;  
— *поворота*, при котором меняется только ориентация объекта;  
— *масштабирования* (увеличение или уменьшение), при котором меняется только размер;  
— *отражения*, при котором меняется только т.н. хиральность (т.е. правшество или левшество) или направление.  
— любой комбинации из перечисленного.

В математике сей специфический набор трансформаций именуют *группой подобия*.<sup>2</sup> Мы уже сталкивались с группами подобия в школьном курсе геометрии, когда, например, через равенство углов доказывали, что два треугольника «подобны».

Несмотря на то, что пример формо-сохранного характера этих преобразований мы дали на базе двухмерных фигур, объективные формы трехмерных объектов зиждутся на трехмерной версии того же набора трансформаций: *переносе, повороте, масштабировании, отражении* или комбинации этих признаков. Они хранят форму в т.н. строго определенном смысле — сие значит, что о двух объектах, которые

<sup>1</sup> Форма также остается инвариантной, когда выполняются непространственные трансформации, такие как изменения окраски и текстуры. Но мы не станем рассматривать их просто потому, что они вне контекста разговора о восприятии формы. Мы поговорим о них несколько позже, когда речь пойдет о вопросах представления формы.

<sup>2</sup> «Подобие» — это только название данного математического набора преобразований, и его ни коим образом не следует относить к принципу перцептивного подобия фигур.



можно привести к строгому соответствию, применив определенную последовательность таких трансформаций, мы можем говорить как об имеющих одну и ту же объективную форму. Словом, преобразования эти есть инструмент объективной проверки того, обладают фигуры данной пары объективной формой или нет.

Если же для приведения пары объектов в строгое пространственное соответствие требуются иные трансформации, например, сплющивание, вытяжка или какая-то иная деформация, то, согласно данному выше определению, у этих объектов разные объективные формы, даже если *воспринимаются* они строго одинаково.

Тот факт, что мы видим фигурки рис. 8.1.1 обладающими одной и той же формой, подводит нас к самоочевидному предположению, что восприятие эквивалентности формы в целом веридикально, то есть — два объекта воспримутся имеющими одинаковую форму всегда, когда их объективная форма одна и та же. Но как мы узнаём вскорости, из правила сего имеют место серьезные исключения. Давайте возьмем определение эквивалентности формы и проверим, насколько ее восприятие зрением человека отвечает этому определению. Отметим при этом, что ситуация чем-то напоминает определение объективной ориентации прямых по отношению к гравитационной вертикали и оценке того, насколько человеческое зрение адекватно ей. Восприятие может давать различные иллюзии и дисторсии, но их легко учесть, если четко определить объективную ориентацию. То же должно касаться и восприятия формы.

Определение перцептивной эквивалентности форм, основанное на инвариантности (отсутствии изменений) при трансформации подобия — это количественный уровень теории эквивалентности форм, не претендующий на описание того, какие процессы задействованы в определении этой эквивалентности форм в смысле, определенном в гл. 2. На алгоритмическом уровне рассуждений, однако, встает вопрос о том, какая информация и какие процессы задействуются при идентификации равенства форм двух объектов? Существуют как минимум три сценария:

1. *Гипотеза инвариантных признаков.* Предполагает, что эквивалентность форм верифицируется сравнением некоего множества «формных признаков» двух объектов. Если эти два множества (или списка) признаков идентичны, у объектов будет одна и та же перцептивная форма. Для работы по этой схеме признаки должны оставаться инвариантными по всему набору дискретных трансформаций.

2. *Гипотеза трансформирующего расположения.* Предполагает, что эквивалентность форм проверяется тем, действительно ли два объекта могут прийти в точное соответствие через одну из возможных трансформаций, в частности преобразование подобия. Если сие возможно — формы воспримутся как идентичные; если нет — формы различны.

3. *Гипотеза объектно-центрированного репера.* Предполагает, что эквивалентность форм проверяется сравнением относительно некоего эталона — репера,<sup>1</sup> определенного действительными свойствами этого объекта. Если объекты одинаковы относительно своих собственных реперов, они воспринимаются одинаковыми по форме.

Понятно, что существуют промежуточные, комбинированные подходы к вопросу эквивалентности, но для простоты мы рассмотрим «чистые» версии этих гипотез.

<sup>1</sup> Реперы, реперные точки — точки, на которых основывается шкала измерений. К примеру, на реперных точках построена Международная практическая температурная шкала. Реперные точки на шкале Цельсия — температура замерзания (0°C) и кипения воды (100°C) на уровне моря. — *Прим. пер.*

### 8.1.2 Инвариантность признаков

Гипотеза инвариантных признаков предполагает, что восприятие формы объекта зависит от способа кодирования неизменных (инвариантных) свойств этого объекта, когда его трансформируют определенным образом. Рассмотрим, что же сие значит и как соотносится с восприятием эквивалентности форм.

Любая трансформация делит множество всех возможных свойств объекта на два взаимообратных подмножества: те, что остаются неизменными (и называются *инвариантными признаками*) и те, что меняются (и называются *признаками переменными*). В качестве примера рассмотрим эффект двухмерного переноса множества свойств плоской фигуры, показанной на рис. 8.1.2. Изменим расположение объекта так, чтобы изменились абсолютные позиции его компонент — линий, углов и т.д. Мы говорим при этом, что позиции эти являются *трансляционно-переменными признаками* (что отмечено серыми квадратами с минусом в первой колонке первого ряда). Сие не влияет на *относительные* позиции этих компонент и не влияет на прочие свойства, такие как число линий и углов в фигуре или же ориентацию и длину этих линий, а также величину углов. Следовательно, мы говорим, что позиции относительные полагают *трансляционно-инвариантными признаками*, что отмечено белыми квадратами с плюсом в остальных колонках этого ряда.

Теперь рассмотрим, что происходит, когда меняется ориентация объекта. К примеру, поворот меняет не только абсолютные и относительные позиции компонент, но также их абсолютные ориентации, которые, следовательно, являются *ротационно-переменными признаками*. Поворот не меняет относительных ориентаций компонент объекта (то есть, угловых отличий в ориентации между ними), не оказывает влияния

	Абсолютное расположение	Относительное расположение	Абсолютная ориентация	Относительная ориентация	Абсолютный размер	Относительный размер	Количество	Сближенность	Связность
Перенос	—	+	+	+	+	+	+	+	+
Поворот	—	—	—	+	+	+	+	+	+
Масштабирование	—	+	+	+	—	+	+	+	+
Отражение	—	+	—	+	+	+	+	+	+
Все комбинации	—	—	—	+	—	+	+	+	+

**Рис. 8.1.2** Концепция инвариантных признаков в контексте эквивалентности форм. Трансформация каждого из типов оставляет одни признаки инвариантными, а другие меняет. В таблице показано то, какие признаки меняются (—), а какие остаются инвариантными (+) при трансформациях подобия.

на такие признаки как число или длина линий и величина углов (которые, соответственно, являются *признаками ротационно-инвариантными*).

В связке с определением эквивалентности объективных форм, озвученным, напомним, в логике трансформаций подобия, гипотеза инвариантных признаков указывает на то, что в рамках группы подобия форму можно представить множеством инвариантов: переносов, поворотов, отражений, масштабирований и их различных комбинаций.

Сие был базовый посыл «чистой» гипотезы инвариантных признаков.

Нижний ряд рис. 8.1.2 указывает на то, что, согласно данной гипотезе, признаки-кандидаты — это число линий, величина углов и различные прочие компоненты, их относительные ориентации, размеры, плюс их завершенность и связность: ни один из этих признаков не меняется при трансформациях подобия.

Отметим также, что список включает в себя малое число абсолютных метрических свойств, но множество свойств относительных. Сие согласуется с ранними наблюдениями гештальт-психологов, исповедовавших идею *детерминанты относительности*, то есть того, что восприятие зиждется на *конфигурации отношений* между признаками и частями объекта, нежели на их абсолютных свойствах.

Яркой особенностью инвариантных свойств является то, что свойства эти, не меняясь при трансформациях, гарантируют то, что любые две фигуры, которые можно совместить через трансформации подобия, будут иметь идентичные представления формы. Следовательно, эквивалентность *перцептивных* форм — это гарантия эквивалентности форм объективных, то есть, исходя из определения трансформаций подобия, представление формы в логике признаков инвариантных по подобию — это решение перцептивного вопроса эквивалентности форм.

Со времен классической работы американских нейрофизиологов Вальтера Питтса и Уоррен Мак-Каллока (1947) и вплоть до самого недавнего времени гипотеза инвариантных признаков в той или иной своей форме оставалась доминирующим концептом. Скрыто ли, явно ли догадки Питтса — Мак-Каллока лежат в основе гештальт-теорий восприятия формы, гибсоновской идеи константности формы и классических концепций распознавания паттернов, в частности таких как т.н. «теория преисподней» (Selfridge, 1957; Selfridge & Neisser, 1960).

Теории, построенные в парадигме инвариантных признаков, подкупают своей простотой: форму можно представить в виде множества атрибутов, а свойства признаков внутри такого множества — дифференцированно взвесить согласно их перцептивным особенностям или значимости. Основная же идея такова, что достаточно иметь на руках совокупность инвариантных признаков, чтобы объяснить, почему некоторые пары отличных друг от друга объектов воспринимаются одинаковыми по форме, а другие — нет.

Однако, к сожалению, существуют убедительные доказательства того, что концепция инвариантных признаков как теория перцептивной эквивалентности форм — весьма слаба. Сложности проистекают из серьезных отличий между перцептивной и объективной эквивалентностями форм. Наиболее показательным здесь является феномен строго определенной связи между перцептивной ориентацией и перцептивной формой, чего мы слегка коснулись в гл. 7. Одна из наиболее ранних, простейших и, в тоже время, изящных демонстраций этой связи — наблюдение Маха (1914/1959): ежели квадрат повернуть на  $45^\circ$ , мы станем воспринимать его как прямостоящий даймонд,<sup>1</sup> но не как повернутый квадрат (рис. 8.1.3). Разумеется, мы *можем*, приложить усилие и

<sup>1</sup> Даймонд или диамант — прямоугольный ромб. — Прим. пер.

увидеть эту фигуру как повернутый квадрат, если мысленно положим его плоскую сторону, расположенную под  $45^\circ$  к вертикали, как верхнюю. Но если верхний вертекс воспринимается как вершина, форма фигуры окажется ромбовидной и весьма отличной от прямостоящего квадрата.

Иными словами, махов квадратодаймонд выявил одно весьма слабое место в концепции инвариантных признаков: если восприятие формы определено исключительно признаками, инвариантными при повороте, то любые две фигуры, которые являются лишь повернутыми версиями друг друга (в частности маховы квадрат и даймонд) должны восприниматься обладающими одной и той же формой. Но сие не так, как видим (что всерьез подтачивает основу гипотезы инвариантных признаков как концепции восприятия эквивалентности форм алгоритмического уровня).

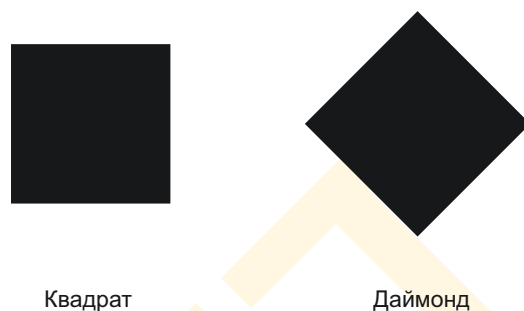
Может показаться, однако, что проблема решается простым исключением поворота из списка трансформаций, сохраняющих перцептивную форму. Такое решение, однако, вместе с водой выплескивает и младенца, поскольку довольно случаев, когда две фигуры, объединенные общим поворотом спонтанно воспринимаются имеющими одинаковую форму. Одна такая пара показана на рис. 8.1.1 В. Примеры, подобные сему, эффективно блокируют любую попытку «причесать» концепцию инвариантных признаков простым исключением поворотов. Приходится искать другие решения.

### 8.1.3 Трансформирующая привodka

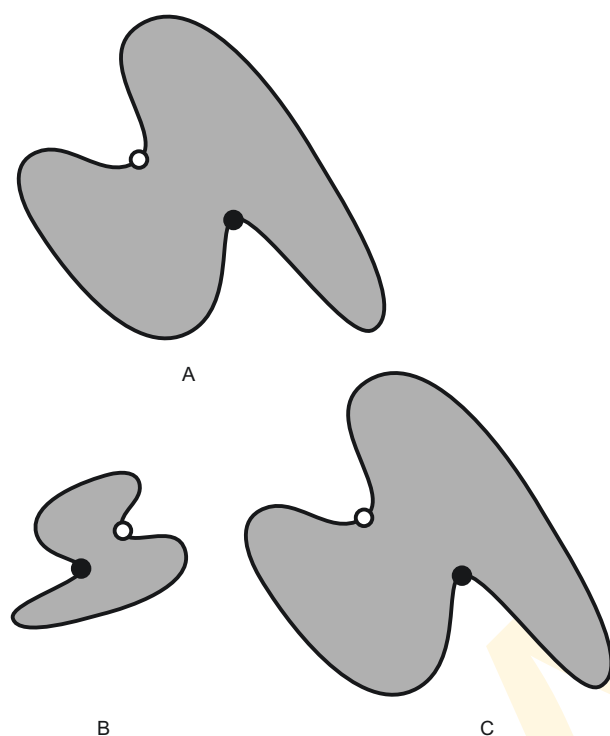
Второй путь ответить на вопрос, являются два объекта эквивалентными по форме или нет — это подобрать такую пространственную трансформацию, которая расположит один объект точно так же, как другой. Если такая трансформация существует и ее можно взять из дозволенного множества (то есть, из группы подобия) — объекты будут иметь одну и ту же объективную форму. В противном случае, формы окажутся разными.

Сей подход подкупает не только тем, что весьма схож с определением эквивалентности объективных форм, данным выше, но еще и своей тесной связью с некоторыми иными весьма важными зрительными феноменами. Когда, к примеру, в гл. 10 мы поведем разговор о восприятии движения, то увидим, что зрительная система склонна воспринимать формосохраняющее движение, нежели формодеструктивное. А в гл. 12 мы поговорим о значимости изображающих трансформаций при сравнении форм двух похожих объектов. Объектные трансформации и процессы расположения также занимают важное место в некоторых теориях детекции объектов, о чем речь пойдет в гл. 9.

Итак, чтобы проиллюстрировать действие трансформирующей привodka, вообразим два идентичных друг другу двухмерных объекта, разнящихся расположением, ориентацией и размером — рис. 8.1.4 А и 8.1.4 В. У обоих объектов имеется по две точки излома в одних и тех же местах (относительно объекта). Оба объекта (А и В) мож-



**Рис. 8.1.3** Дуализм квадрата/даймонда. Одна и та же фигура воспринимается квадратом, когда ее стороны вертикальны и горизонтальны, но при этом как даймонд, когда они диагональны. Перцептивную форму даймонда легко изменить, если вообразить, что его повернули на  $45^\circ$ . То же возможно в отношении квадрата, но существенно сложнее.



**Рис. 8.1.4** Определение эквивалентности форм через трансформирующую приводку. Тот факт, что форма объектов А и В одна и та же можно доказать, поставив в каждой фигуре по паре точек, найдя трансформацию, которая выровняет линии, соединяющие эти точки в каждой фигуре, а затем применив масштабирование. Результат применения такой трансформации к объекту В (то есть, теперь объект С) можно затем напрямую совместить с объектом А.

ственной приводки фигур на фронтальной плоскости, но трехмерным объектам необходимы как минимум три неколлинеарных точки. Поскольку опорные точки естественных фигур специально никто не окрашивает, то по началу и неочевидно, какая точка на одном объекте с какой согласуется на другом. Может помочь, однако, разметка этих точек по характеру (выпуклость, вогнутость, центровка и пр.), поскольку выпуклость на одной фигуре может сесть только на выпуклость другой. Но даже в этом случае пример на рис. 8.1.4 остается неопределенным, поскольку обе точки расположены на вогнутостях. Следовательно, процедура должна испытать оба соответствия и выбрать правильное.

Итак, несмотря на все плюсы, у гипотезы трансформирующей приводки очевидны недостатки, один из которых в том, что, ежели фигуры сложны и опорных точек множество, придется искать принципы исключения лишних вариантов, или же, альтернативно — перепробовать все возможные. Без специальных меток (например, окраски)  $n$  опорных точек в каждом объекте дадут  $n!$  [ $= n \times (n - 1) \times (n - 2) \times \dots \times 1$ ] возможных соответствий. Сие огромное число. К примеру, при пяти недифференцированных

но сравнить затем по критерию эквивалентности форм:

1. Найти соответствие между точками — черная к черной; белая к белой.

2. Определить перенос, поворот, отражение и масштабирование, необходимые для точного позиционирования точек фигуры В на точках фигуры А.

3. Применить ту же операцию ко всей фигуре В, включая все точки (а не только черную и белую).

4. Определить, действительно ли трансформированная версия фигуры В (фигура С) идентична фигуре А. Если идентична — у фигур А и В одинаковая форма; если не идентична — формы разнятся. Вывод сей гарантирован, потому что выбранная нами трансформация взята из группы подобия.

С реальными фигурами процедура приводки проходит намного сложнее по ряду причин, и одна из главных в том, что объекты редко когда оказываются помеченными черной и белой точками. Скорее то небольшое число точек излома (иногда называемых *опорными*) приходится идентифицировать из структуры самой фигуры. Точки максимальной вогнутости по контуру (на рис. 8.1.4 черная и белая точки) — это удачные опорные точки, равно как и точки максимальной выпуклости в центре фигуры (Huttenlocher & Ulman, 1987). Две опорных точки достаточны для простран-



опорных точках имеются 120 возможных соответствий, а при десяти — 3628800. То есть, множество возможных соответствий следует искать до тех пор, пока не отыщется приемлемое совпадение по расположению, либо же, пока все варианты не будут испробованы.

Вторая проблема заключена в том, что у обеих фигур должно быть видимо одно и то же множество опорных точек. С планарными фигурами (рис. 8.1.4) или трехмерными объектами того же класса (проволочные фигурки, у которых нет самоокклюзии поверхностей) проблем не возникает, но сие крайне затруднительно в отношении большинства трехмерных *непрозрачных* объектов, поскольку точки хорошо видимые у одного, могут оказаться невидимыми у другого из-за естественной самоокклюзии.

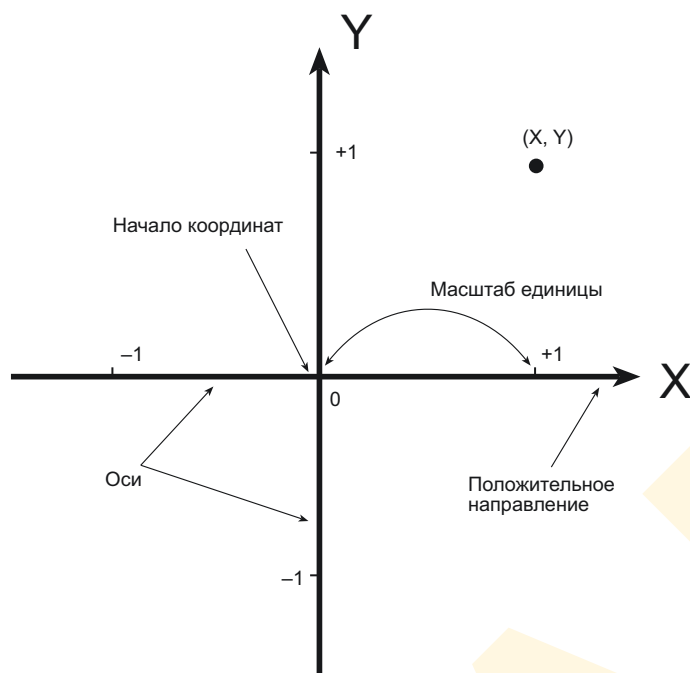
Третья проблема в том, что существуют объекты, которые видятся нам разными по форме, но при этом приводятся в строгое соответствие трансформацией подобия. Махов квадратодаймонд — один из примеров (см. рис. 8.1.3), поскольку даймонд — это квадрат, повернутый на  $45^\circ$ . Гипотеза трансформирующей приводки предполагает, что такие фигуры должны всегда выглядеть одинаковыми по форме, но, как видим, сие не обязательно.

Для учета данной нестыковки концепция трансформирующей приводки должна как-то объяснять то, почему такая приводка не срабатывает в отношении махового квадратодаймонда, тогда как поворот, скажем литеры «R» на  $45^\circ$  никак не приводит к изменению ее перцептивной формы. В первом случае зрительная система либо ошибочно выбирает разные опорные точки у квадрата и даймонда, либо по-разному согласует одни и те же точки, но при этом корректно согласует их в литерях «R». Виды опорных точек, описанные выше — центры масс и максимальные вогнутости и выпуклости — поворотно-инварианты и малопригодны к толкованию феномена квадратодаймонда. Прочие варианты опорных точек и и/или ориентационно-смещенных схем установки соответствий между опорными точками — это уже некая махинация. К примеру, если расположение определено осями симметрии и если зрительная система склонна отсчитывать симметрию от вертикальных осей — маховы квадрат и даймонд будут ориентационно располагаться так, как выглядят на рис. 8.1.3, и, следовательно, будут иметь различные перцептивные формы.

#### 8.1.4 Объектно-центрированный репер

Третий вариант решения вопроса эквивалентности форм — это определение формы относительно объектно-центрированного репера. Во многом эта идея сходна с процессом трансформирующей приводки, но при этом отлична тем, что требует отдельного описания.

Мысль о том, что восприятие формы может задействовать объектно-центрированные реперы, навеяна тем наблюдением, что махов даймонд можно увидеть повернутым квадратом, если наклоненную сторону перцептивно истолковать верхней стороной фигуры. Что сие значит? Сие значит, что махову фигуру можно воспринимать как обладающую одной формой (даймонд) относительно гравитационной вертикали и иной (квадрат) относительно диагональной ориентации, переопределяющей перцептивные верхнюю, нижнюю и боковые стороны. Отметим при этом, что сии альтернативы взаимоисключающи, то есть фигуры не могут восприниматься обладающими двумя формами одновременно. Ирвин Рок (1973) был одним из первых, кто показал, что связь между формой и ориентацией имеет место благодаря тому, что перцептивная форма — это описание относительно перцептивного репера (хотя мысль о нем звучала и ранее из уст гештальтистов в их толкованиях прочих феноменов восприятия ориентации, например,



**Рис. 8.1.5** Эвклидова координатная система. геометрические точки в двухмерном пространстве можно описать парой чисел, если перед этим установить оси каждой мерности, а также позицию, доступную обеим осям (начало координат — «0»), направление по оси («+» или «-») и масштаб единицы («1»).

как сие осуществимо, обратимся к аналитической геометрии, в которой координатные системы играют роль такого репера в описании объектов.

#### 8.1.4.1 Геометрические координатные системы

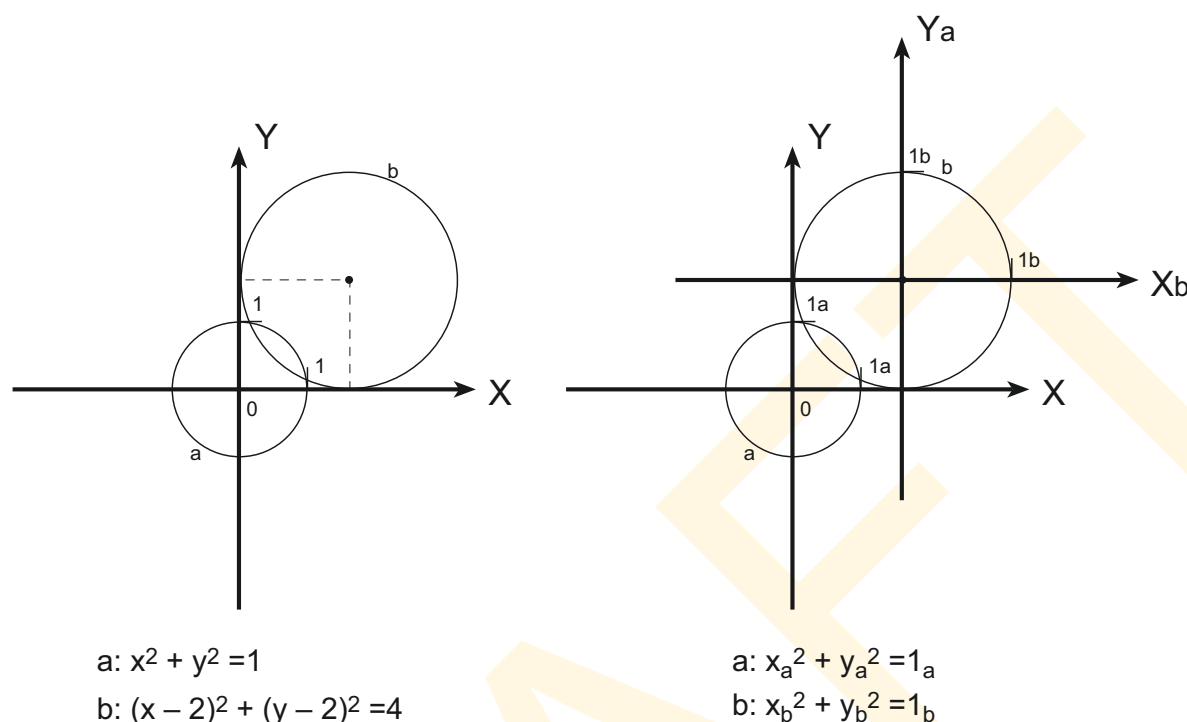
Аналитическая геометрия способна давать символическое описание различных объектов (линий, кривых, кругов и пр.), т.е. описание в виде уравнений. Такую возможность предоставляет т.н. *координатная система* — формальная структура, придуманная французским философом и математиком Рене Декартом. Координатная система позволяет каждую точку  $n$ -мерного пространства представить в виде упорядоченного множества  $n$ -значений, именуемых *координатами*. Набор точек, входящих в описание геометрического объекта (линии, круга и пр.) можно описать затем уравнением, выражающим отношения между числовыми координатами этих точек.

Рассмотрим обычную двумерную координатную систему, работающую через установку структурированного множества пространственных реперов в координатах, обозначаемых обычно « $x$ » и « $y$ ». Каждая размерность координатной системы включает в себя:

- репер *расположения*, определяющий *начало координат*;
- репер *ориентации*, определяющий *ось*;
- репер *расстояния*, определяющий *размер единицы величины*;
- *смысловой репер*, определяющий *положительное направление по оси*.

в объяснении прутикооконого эффекта Эшем и Виткиным / Asch & Witkin, 1948a, 1948 b). Рок тоже касался этой идеи в контексте ориентации, но прочие исследователи (например Marr & Nishihara, 1978; Palmer, 1975b, 1989) в дальнейшем и по аналогии с геометрическим репером (см. ниже) разнесли ее на прочие свойства объектов, в частности такие как расположение и размер.

В общем виде гипотеза объектно-центрированного репера дает альтернативное двум другим толкование механизмов перцептивной эквивалентности форм, базируясь на той идее, что эффекты трансформаций подобия в восприятии формы можно устранить, позиционировав некий внутренний репер, эффективно отгораживающий представление формы от воздействия пространственных трансформаций. Чтобы понять,



**Рис. 8.1.6** Объектно-центрированные координатные системы. Слева показано, как два круга можно описать двумя уравнениями внутри одной координатной системы — у каждого круга свое описание. Справа показано, что если каждому кругу дать свою собственную координатную систему, описание останется неизменным относительно собственного объектно-центрированного репера.

Palmer, 1983.

Понятно, что при работе на плоскости, т.е. в двухмерном пространстве, достаточно присвоить пару значений  $(x, y)$  каждой точке на этой плоскости (см. рис. 8.1.5).

В качестве конкретного примера того, как координатная система дает символическое описание, рассмотрим круг, отмеченный «а» в левой половине рис. 8.1.6. В декартовой координатной системе он описывается уравнением  $x^2 + y^2 = 1$ .

Для данного конкретного уравнения центр координатной системы должен располагаться в центре круга, а цена деления должна быть равна радиусу круга.<sup>1</sup> Итак, сей репер позволяет нам символически описать данный конкретный круг с помощью данного конкретного уравнения.

Можно ли использовать это уравнение для представления формы всех кругов? Чтобы ответить на этот вопрос рассмотрим то, как другие круги, в частности тот, что отмечен литерой «b» в левой половине рис. 8.1.6, описываются внутри координатной системы. В целом уравнение будет похоже, но кое в чем и отличным:  $(x - 2)^2 + (y - 2)^2 = 4$ .

Разница возникает из-за того, что расположение и размер круга *b* отлично от таковых у круга *a*, поэтому, когда для описания используется одна и та же координатная система потребуются немного разные уравнения.

<sup>1</sup> Ориентация и направление вдоль по осям не играют роли для круга, поскольку круги инвариантны, то есть — симметричны по всем вариантам поворота и отражения.

Ключевым моментом нашего понимания роли объектно-центрированных реперов в зрительном восприятии является то, что для описания обоих кругов вместо применения одной и той же координатной системы мы можем использовать разные. Если центр координатной системы круга  $b$  совпадает с его центром, и если шаг этой системы совпадает с радиусом  $b$ , то уравнение для круга  $b$  во второй координатной системе окажется точно таким же, что и уравнение для круга  $a$  внутри первой. Сие продемонстрировано в правой половине рис. 8.1.6, где даны те же два круга, но описанные в разных координатных системах. Важным выводом здесь явится то, что относительно их собственных внутренне определенных координатных систем уравнения обоих кругов окажутся одинаковыми. То есть, выбрав «правильную» координатную систему, переменные положения круга и его размера можно исключить из уравнения. Фактически эти данные поглощаются координатными системами и поэтому уравнения инвариантны для всех возможных кругов.

#### 8.1.4.2 Перцептивные реперы

Идея объектно-центрированного репера в восприятии аналогична координатной системе в аналитической геометрии и может применяться для символического перцептивного описания объекта. Один из способов достижения сего — это принятие т.н. *зрительно-центрированного репера*: простой координатной системы, чьи стандарты выбраны относительно наблюдателя с центром в точке фиксации, положительным направлением ориентации по вертикали и правосторонним направлением сетчатки, а также шагом, определенным неким произвольным зрительным углом. Относительно такого репера (который условно выглядит так, как если бы на сетчатку нанесли декартову прямоугольную сетку) можно дать описание любого объекта. Сие, наверное, наиболее простой путь применения перцептивного репера, но, разумеется, далеко не единственный.

Альтернативный вариант — присвоение каждому перцептивному объекту его собственного объектно-центрированного репера — перцептивной системы отсчета, выбранной на базе действительных свойств описываемого объекта. Гипотеза объектно-центрированного репера предполагает, что координатная система, использованная в описании каждого объекта, в какой-то мере «сделана под заказ» для данного объекта. Действительно, в дальнейшем мы покажем, что различные координатные системы можно даже использовать для описания различных *частей* одного и того же объекта.

Грубо говоря, объектно-центрированный репер в целом дает идентичное описание объектов эквивалентных форм, поскольку большинство (но не все) формы достаточно четко структурированы, чтобы индуцировать их описание зрительной системой в рамках одного и того же объектно-центрированного репера. К примеру, если ориентации двух во всем остальном идентичных объектов различны (скажем, нормально стоящий стул и перевернутый), ориентация их объектно-центрированных реперов задается так, что оба будут иметь одно и то же формное описание относительно своих объектно-центрированных реперов. Сходно, если размеры двух во всем остальном идентичных объектов различны (скажем, как у полноразмерного авиалайнера и его масштабной модели), размеры их объектно-центрированных реперов будут различаться соответственно. До тех пор пока отличия в реперах компенсируют трансформационные отличия между объектами, эквивалентность форм будет веридикально восприниматься во всей группе трансформаций подобия. Однако, если по каким-то причинам выбраны разные реперы — как, например, разные оси симметрии в маховых квадрате и даймонде — объекты с одной объективной формой могут восприниматься с разными формами перцеп-

тивными. Сие ключевой момент, благодаря которому обсуждаемая гипотеза выглядит столь привлекательно.

Отметим также, что «объектно-центрированными» реперы названы потому, что стандарты перцептивных реперов представления каждого объекта выбраны с целью заполнения его собственной структуры, также как центры и шаг координатных систем в правой половине рис. 8.1.6 выбраны так, чтобы совпасть с центрами кругов. В целом эквивалентность перцептивных форм имеет место всякий раз, когда двум объектам дается одно и то же символическое описание (что возможно) относительно их «персональных» объектно-центрированных реперов. Перцептивные реперы способны компенсировать отличия в ориентации двух объектов тем же путем, что и отличия в расположении и размере. Если ориентации двух (в остальном идентичных) объектов различны, то объекты эти до тех пор будут восприниматься обладающими одной и той же формой, пока ориентации их объектно-центрированных реперов остаются неизменными относительно объекта. Иными словами, если во всех ситуациях в отношении одной и той же формы использовать один и тот же репер, эквивалентность форм будет идеальной при любых трансформациях в рамках группы подобия.

#### 8.1.4.2 Анализ отказов

Сила, и одновременно слабость идеи реперов заключены в следующей фразе: «Если во всех ситуациях для одной и той же формы используется один и тот же внутренний репер». О слабости мы говорим, потому что в принципе невозможно удовлетворить этому условию. Насколько репер можно привязать к свойствам объекта, если выбирать этот репер всегда? Мы знаем, что порой выбираются «неправильные реперы», и, время от времени, в перцептивном определении эквивалентности форм возникают серьезные погрешности. Погрешности эти могут стать очень сильны, как, например, в случае махова квадратодаймонда.

Каким образом перцептивные погрешности толкуются с позиций концепции объектно-центрированного репера? Пальмер (1985) предложил три принципиальных допущения при различиях в ориентации объекта:

1. *Относительное описание.* Форма воспринимается относительно репера, в котором специфическая ориентация берется как описательный стандарт.

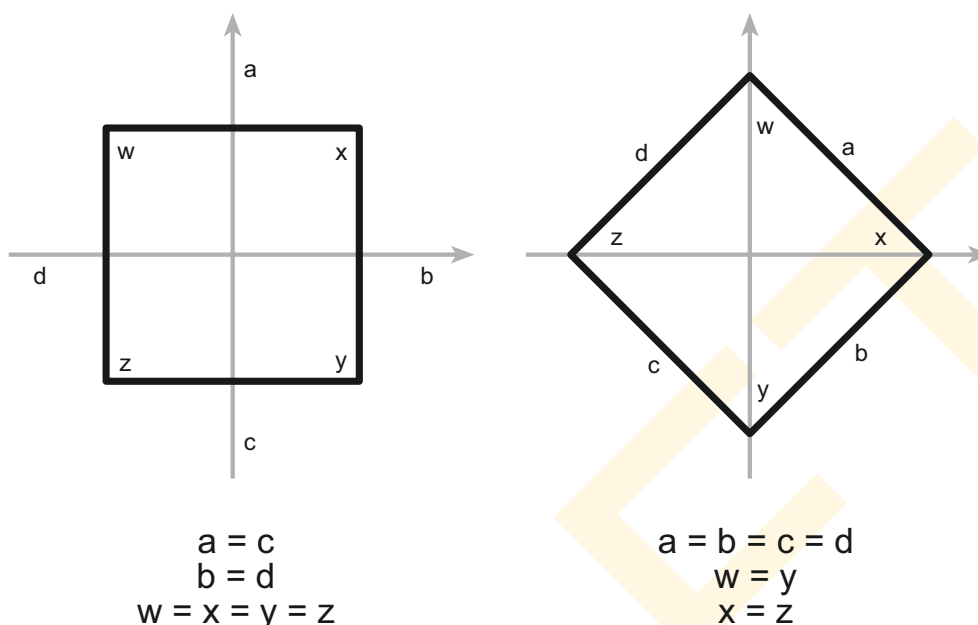
Сие допущение предполагает, что эквивалентность форм определена сравнением скорее символических представлений объектов, нежели непосредственным сравнением самих объектов как при трансформирующем расположении.

2. *Внутренние отклонения.* Для задания объектно-центрированного репера наше восприятие задействует эвристику, основанную на собственной структуре самого объекта.

Мы разберем то, каким образом эвристика эта может стать наикратчайшей. Но при этом всегда существует некая возможность того, что реперную ориентацию можно установить по оси симметрии или по длиннику. Эвристика дает веридикальный перцепт эквивалентности форм в большинстве случаев, поскольку одни и те же алгоритмы, приложенные к одному и тому же объекту, в целом дают один и тот же выход. Однако коль скоро любая эвристика несовершенна, имеет место ряд обстоятельств — в частности такие как множественность осей симметрии у объекта (например, как у приснопоминанного квадратодаймонда) — при которых возможны различные ориентации. Сие дает иллюзии в восприятии эквивалентности форм.

3. *Внешние отклонения.* Вдобавок к отклонениям объектного характера в выборе ориентации существуют смещения в сторону наиболее заметных ориентаций по отно-





**Рис. 8.1.7** Объяснение неопределенности формы квадрата/даймонда в логике реперов. Очевидно, что уравнения для перцептов квадрата и даймонда окажутся различными. В перцепте квадрата паритет противоположных сторон и соседних углов доминируют. В восприятии даймонда доминирует паритет соседних сторон и противоположных углов.

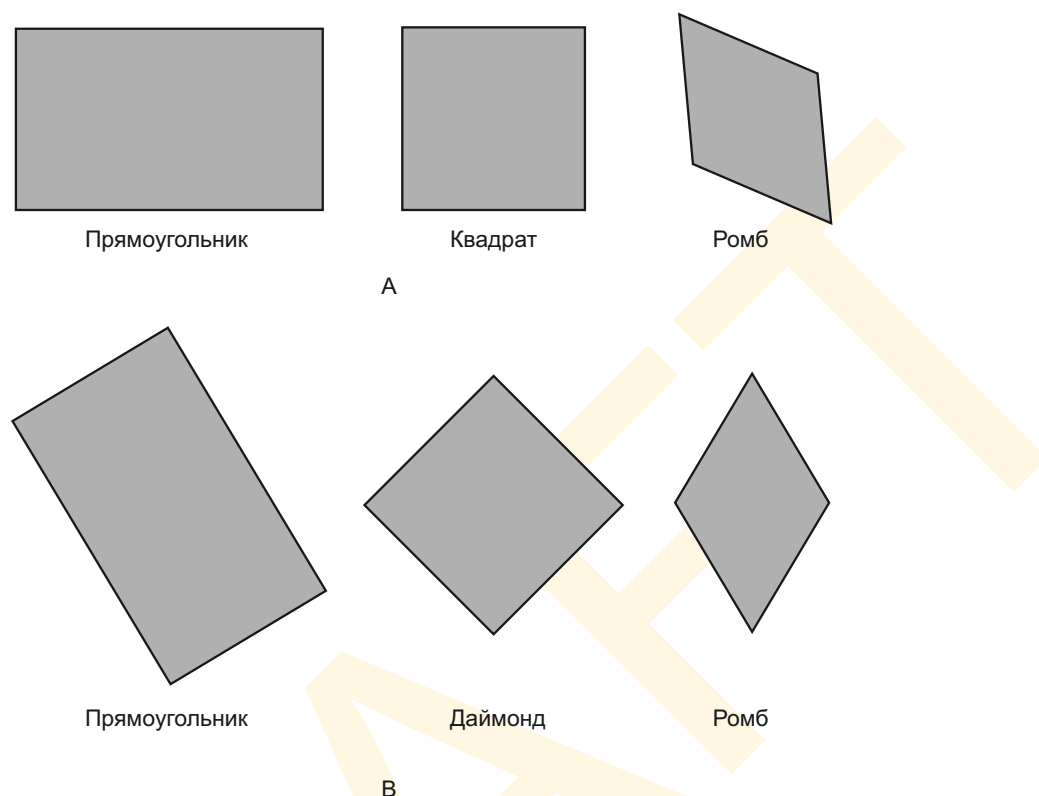
Palmer, 1983.

шению к другим реперным объектам и направлениям — гравитационной вертикали, ориентации корпуса наблюдателя, или оси «верх-низ» сетчатки.

В зрительной системе прямостоящего наблюдателя всё перечисленное работает сочетанно, но может начать конфликтовать, когда его голова и/или корпус отклоняются от гравитационной вертикали. Этих отклонений порой достаточно для того, чтобы вызвать серьезные погрешности в перцепции эквивалентности форм.

Три озвученных выше допущения вместе предполагают, что перцептивная реперная ориентация данного объекта является совокупной функцией его внутренней структуры, ориентации относительно наблюдателя и ориентации экологической (т.е. относительно других объектов реального мира). К примеру, если у фигуры имеются две и более внутренних оси для объектно-основанной реперной ориентации и одна из них оказывается расположенной вдоль гравитационной вертикали, то зрение обычно предпочитает вертикальную ось, что, в итоге, дает некое специфическое описание формы. Если иная, идентичная первой, фигура рассматривается в иной ориентации и при этом так, что иная внутренняя ось располагается по вертикали, эта новая ось выбирается в качестве репера ориентации. Следовательно, один и тот же объект в различных объективных ориентациях может дать разные описания формы.

Сказанное фактически, и есть объяснение махова феномена (то есть, феномена недостаточности эквивалентности форм) силами гипотезы объектно-центрированных реперов. Сие могло бы дать два различных описания формы, в зависимости от чего оси симметрии расположились бы по объективным горизонталям и вертикалям (как говорят, своим *биссектрисам сторон* или *угловым биссектрисам*). Когда биссектрисы сторон располагаются по горизонталям и вертикалям, мы воспринимаем фигуру в рамках описания, соответствующего перцепту «квадрат»; когда угловая биссектриса располагается по гравитационной вертикали, мы воспринимаем фигуру по сценарию, со-



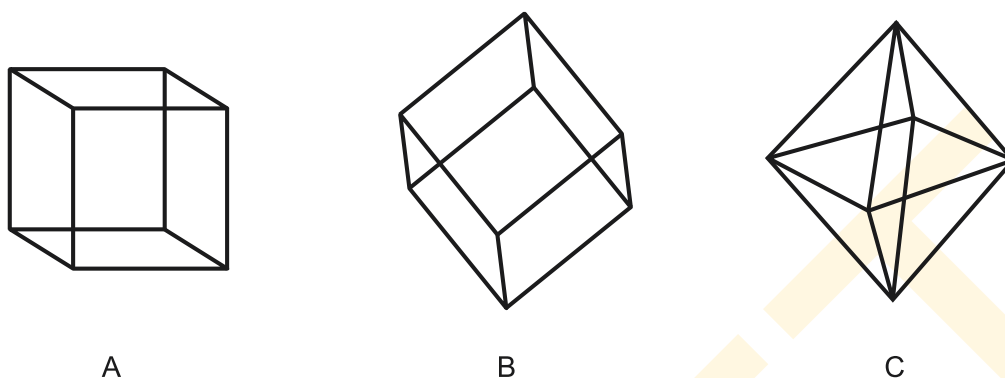
**Рис. 8.1.8** Отличия в подобии форм квадратов и даймондов. Квадрат воспринимается подобным скорее прямоугольнику, нежели ромбу, тогда как даймонд — наоборот. Сии факты согласуются с равенствами на рис. 8.1.7.

ответствующему перцепту «даймонд». Таким образом, сбой в восприятии эквивалентности объективных форм может стать результатом работы одних и тех же реперов, которые в большинстве случаев дают веридикальный перцепт.

Из-за чего же различные реперы дают различные описания формы одного и того же объекта? Пальмер (1983) говорит о том, что они задают различные соотношения свойств, доступных зрительному восприятию (см. рис. 8.1.7 А). Сие и далее предполагает, что квадрат должен по форме более походить на прямоугольник (у которого та же структура), нежели даймонд (чьи стороны наклонены относительно реперных осей и чьи соседние углы не равны). Сравнение, заимствованное из иллюстрации Голдмайера (1936/1972), дано на рис. 8.1.8 А.

Напротив, перцепт даймонда подчеркивает наклон сторон относительно осей и то, что противоположные углы равны по размеру, а соседние стороны — по длине (рис. 8.1.7 В). Сие предполагает, что даймонд должен больше походить именно на прямоугольный *ромб*, но не на прямоугольник, обращая при этом сходные отношения, данные в примере с квадратом (рис. 8.1.8 В). Сие верно, даже когда все три фигуры — это лишь повернутые на  $45^\circ$  объекты, показанные рис. 8.1.8 А.

Канадский исследователь Джеффри Хинтон (1979, 1981), применяя разнообразные реперы сделал ряд ценных наблюдений за утратой эквивалентности форм путем. Исследователь изучал вопросы воображения: сперва он просил своих испытуемых внимательно рассмотреть некий проволочный куб, стоящий перед ними на столе (рис. 8.1.9 А), а затем вообразить его повернутым так, чтобы две противоположных



**Рис. 8.1.9** Реперные эффекты в воображении. Когда Хинтон (1979) просил испытуемых про-  
сил вообразить поворот стандартного куба (часть А) так, чтобы он встал на одну из вершин,  
то большая часть из них не смогли правильно расположить остальные вершины (как в части  
В), а получалось в основном то, что показано в части С.

Пальмер, 1983.

друг другу вершины расположились по вертикали (рис. 8.1.9 В). Как только испыту-  
емый произносил «Да, готово!», исследователь просил указать воображаемое распо-  
ложение остальных вершин фигуры. Почти все при этом указывали на четыре точки в  
квадратной конфигурации, лежащие в горизонтальной плоскости, которая делила по-  
полам линию между вертикально расположенными вершинами. Фактически же сие не  
определяло собственно куб, но двойную пирамиду (рис. 8.1.9 С). Правильный же от-  
вет, как видимо, куда сложнее: существуют фактически шесть вершин, лежащие по-  
переменно на двух параллельных плоскостях и соединенных краями так, что можно  
ходить вперед-назад между ними (рис. 8.1.9 В).

Отметим, что выбор перцептивного репера дает наблюдателю различные множества  
геометрических отношений, заключенных в структуре фигуры. Поскольку из фигуры  
невозможно извлечь все возможные отношения (их неопределенно много), поскольку  
многие из них сверхизбыточны, зрительная система выбирает из доступных наиболее  
стабильное и удобное множество. Это множество потенциально перцептивных отноше-  
ний, о которых Пальмер (1983) говорил как о составляющих структуру перцептивно-  
го репера. Остается теперь точно определить, какими эти отношения должны быть, и  
действительно ли это наилучший путь характеристики перцептивных реперов.

#### 8.1.4.3 Ориентация и форма

Разбор махового феномена, который выглядит лишь любопытной зрительной ано-  
малией, может показаться совершенно бессмысленным. Однако Ирвин Рок (1973) убе-  
дительно показал, что феномен сей куда более значим, чем может показаться нам на  
основе повседневного опыта. Вывод ученого таков, что объектно-центрированные ре-  
перы постоянно участвуют в восприятии эквивалентности форм объектов. При этом  
сбои системы порой говорят нам много больше о механизмах этой системы, нежели ее  
нормальная работа. Плюс к тому, мы не должны упускать из виду тот факт, что верное  
восприятие эквивалентности форм — это правило, а сбой — исключение.

Рок (1973) продемонстрировал также, что при некоторых условиях восприятие экви-  
валентности форм при поворотах в плоскости изображения проходит с очень большим  
трудом. Он предъявлял своим испытуемым последовательность из нескольких аморф-

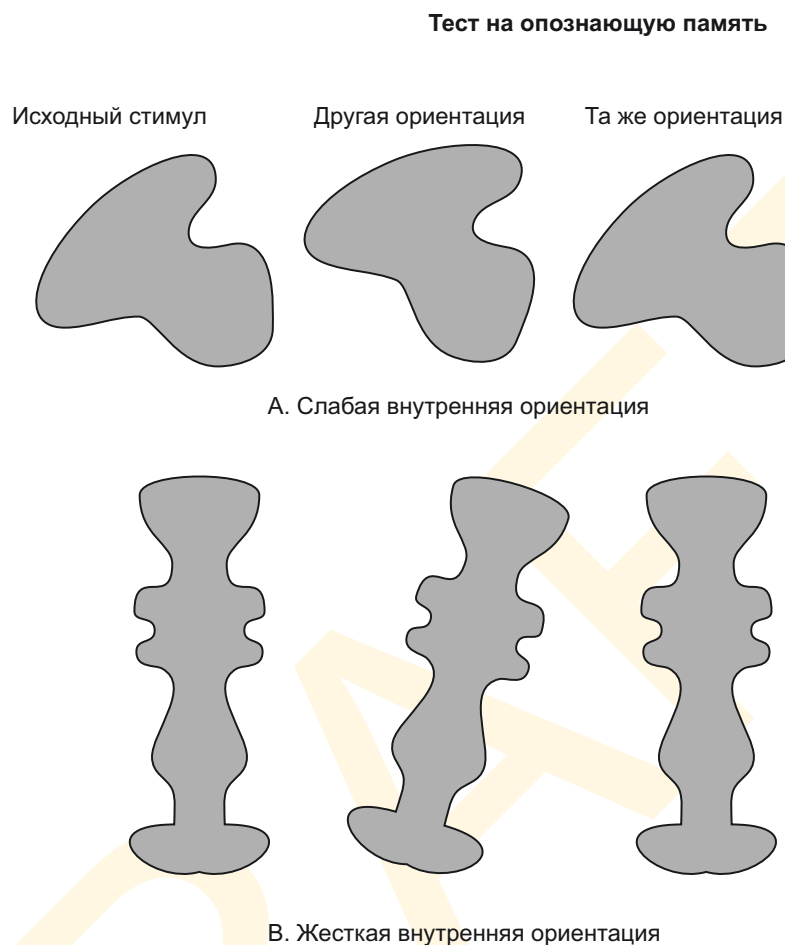
ных и ранее не виденных ими форм в определенной ориентации, а затем проверял их опознающую память по тем же фигурам, но уже в иной ориентации (рис. 8.1.10 А). Результаты показали, что мы намного хуже узнаем объект в новой его ориентации, нежели в старой. То есть, слабый уровень опознавания говорит об отсутствии перцептивной эквивалентности форм исходных и тестовых объектов.

В серии дальнейших опытов Рок показал, что первичными факторами, определяющими реперную ориентацию сих слабоструктурированных фигур, являются факторы не ретинальные, но экологические и/или гравитационные. К примеру, когда наблюдатель после тестовой стимуляции, наклонив голову на  $90^\circ$ , смотрит на стимул тестовый (без изменения фактической ориентации фигур) распознавание оказывается намного лучшим, чем когда на  $90^\circ$  меняется ориентация самих фигур, но при этом без наклона головы. Эти и ряд других фактов ученый использовал в качестве доказательства того тезиса, что форма воспринимается относительно экологического, т.е. *внешнего репера*, в котором сила тяжести задает реперную ориентацию как минимум в отсутствие собственных осей у объекта. Если экологическая (фактическая) ориентация фигуры меняется от исходной к тестовой, описание тестовой фигуры разойдется с описанием, сохранившимся в памяти наблюдателя, и, следовательно, перцепта эквивалентности форм может не возникнуть.

Классическое роково открытие сбоя в перцепции эквивалентности форм относится, в первую очередь, к аморфным формам. Но что происходит, когда используется фигура с четко выраженной собственной осью? Ведет ли это к выбору соответствующего объектно-центрированного репера? Бельгийка Марианна Вайзер (1981), детально занимавшаяся этим вопросом, обнаружила, что фигуры с четко выраженными собственными осями одинаково распознаются и когда исходная фигура с тестовой находятся в разных ориентациях — и когда в одинаковых (рис. 8.1.10 В). Взяв аморфные фигуры, наподобие роковых, Вайзер повторила его результаты (рис. 8.1.10 А), и полученные данные целиком «сели» в рокову теоретическую схему (1973), хотя исследователь больше интересовалась сбоями в восприятии эквивалентности форм, нежели нормой.

В ходе дальнейших опытов Вайзер (1981) показала, что если ось четко структурированной фигуры исходно отклонена от гравитационной вертикали, дальнейшее распознавание ускоряется, если фигуру тестовую поставить *вертикально*. Исследователь интерпретировала результат опыта как указующий на то, что форма фигуры хранится в памяти такой, как если бы предмет располагался вертикально относительно собственного объектно-центрированного репера. Сей результат особо важен потому что опровергает расхожий миф, будто распознавание формы всегда лучше, когда исходная и тестовая фигуры всегда представлены в одной ориентации. Нам просто нужно допустить, что объектно-центрированный репер фигуры задается собственной внутренней структурой, когда структура эта четко выражена. Следовательно, вайзеровы результаты предполагают, что процессы распознавания работают рациональнее и эффективнее, когда объектно-центрированный репер фигуры совмещен с гравитационным репером сцены.

Отметим, что данные Вайзер по фигурам с четко выраженными внутренними осями демонстрируют высокую устойчивость нашего восприятия эквивалентности форм. Объект отлично распознается в любой ориентации, поскольку воспринимается имеющей ту же форму. Однако сие происходит только тогда, когда у фигуры четко выражена внутренняя ось, оба раза приводящая реперную ориентацию к этой оси. Ежели у фигуры более одной оси (как, например, у махового квадрата-даймонда) или же четких осей нет (как у роковых стимулов) перцепт эквивалентности форм не возникает, когда разные оси садятся на очевидно внешние ориентации у исходной и тестовой фи-



**Рис. 8.1.10** Влияние ориентации на восприятие формы. Когда наблюдателям в некоей определенной ориентации предъявлялись фигуры со слабо выраженными внутренними осями (А), а затем выполнялся тест на опознающую память по фигурам той же и другой ориентаций, они лучше справлялись с заданием, когда ориентация была той же. Фигуры с четко выраженными внутренними осями (В) распознавались почти одинаково при исходной ориентации и при ориентации иной.

гур. Словом, результаты опытов Рока и Вайзер укрепляют позиции идеи объектно-центрированных реперов.

#### *8.1.4.4 Эвристика при выборе реперов*

Если расчет перцептивной эквивалентности форм в логике объектно-центрированных реперов оказался верен, то на следующем этапе нужно оценить факторы, управляющие выбором этих реперов. По некоторым свойствам объекта ответ самоочевиден, по некоторым — необычайно труден. Нулевая позиция репера (соответствует центру геометрической системы координат) может находиться в центре масс объекта. Сие легко «вычисляемое» свойство этого объекта, свойство это вполне можно использовать для задания центра объектно-центрированного репера. Реперный шаг (соответствует цене деления осей в геометрической системе координат) можно идентифицировать по длиннику объекта. Сие тоже легко определяемое свойство, которое можно использовать для



оценки размера объектно-центрированного репера. Таким образом, абсолютная позиция и абсолютный размер — это два параметра, которые можно смело факторизовать (т.е. вынести за скобки) с минимальной неопределенностью или даже вообще без нее.<sup>1</sup>

Нетрудно догадаться, что выбор ориентации внутреннего репера объекта намного сложнее. Очевидно значимыми факторами здесь являются:

1. *Гравитационная ориентация.* Поскольку сила тяжести — это наиважнейший параметр, определяющий ориентацию объектов сцены, он важен и в определении объектно-центрированных реперов. Гравитационная вертикаль выступает в роли *значения по умолчанию* для ориентации, которая будет выбрана для объектно-центрированного репера в отсутствие четко ориентированной структуры самого объекта (Rock, 1983). Отметим при этом, что горизонталь — это еще одна ярко выраженная ориентация, влияющая на этот выбор.

2. *Оси двусторонней симметрии.* Объектная структура первого типа, пригодная для определения ориентации внутреннего репера — это ось двусторонней (билатеральной) симметрии (Palmer, 1983, 1985). Когда в наличии лишь одна такая ось, то ею можно однозначно определить ориентацию репера (рис. 8.1.10 В), но ежели осей больше (как, например, в маховом квадратодаймонде — рис. 8.1.3) — ориентация репера потенциально неопределенна.

3. *Длинники.* Параметр второго типа, вполне пригодный для определения ориентации репера — это длинник фигуры (Marr & Nishihara, 1978; Palmer, 1975 b; Rock, 1973). Опять же, когда в наличии один очевидный длинник, его можно использовать для спецификации однозначной реперной ориентации (рис. 8.1.10 В). Но ежели длинников более одного — исходом явится неопределенность.

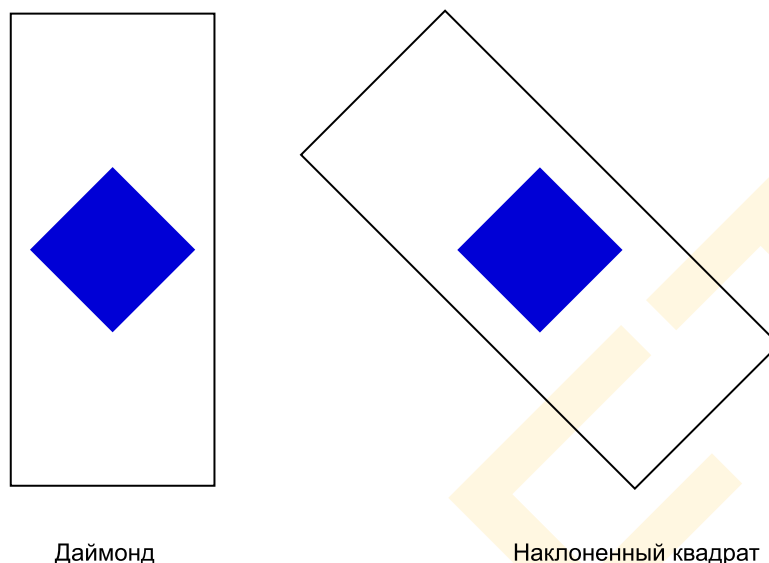
4. *Ориентация контура.* Еще один фактор, определяющий ориентацию объектно-центрированного репера — это ориентация контура объекта, особенно если объект прямой. При этом исключительно важен нижний край, поскольку задает гравитационную стабильность — экологически крайне важное свойство объекта. Контурные и иных ориентаций тоже способны дать явный реперный эффект, особенно если они достаточно длинные и/или идут параллельной парой.

5. *Ориентация текстуры.* Ориентация текстурных элементов внутри фигур может стать очень важной в определении ориентации репера, особенно если в сравнении с самой фигурой элементы относительно крупны в длину и в ширину (Palmer & Bucher, 1982).

6. *Ориентация контекста.* Ориентация близлежащих объектов может также повлиять на ориентацию репера, особенно если они сами по себе строго ориентированы и состоят в тесной структурной связи с интересующей нас фигурой. В частности, ярко выраженное контекстное влияние оказывает окружающий прямоугольник. К примеру, на рис. 8.1.11 показано, что прямоугольник, наклоненный на 45° может прямостоящий махов даймонд заставить восприниматься наклоненным квадратом (Korferman, 1930; Palmer, Simone & Kube, 1988). Ориентация фоновой текстуры, окружающей фигуру, также может повлиять на ориентацию перцептивного репера (Palmer & Bucher, 1982).

7. *Движение.* Направление движения объекта оказывает сильное влияние на перцепт его ориентации и формы (Bucher & Palmer, 1985). Предположительно сие благодаря тому, что движение выделяет ориентацию, сходную с направлением движения всего остального. Та часть объекта, что обращена по направлению движения, обычно

<sup>1</sup> Сказанное справедливо только в отношении двухмерных объектов во фронтальной плоскости. Когда задействуется трехмерное восприятие, то и длинник, и центр масс могут измениться в зависимости от перспективы.



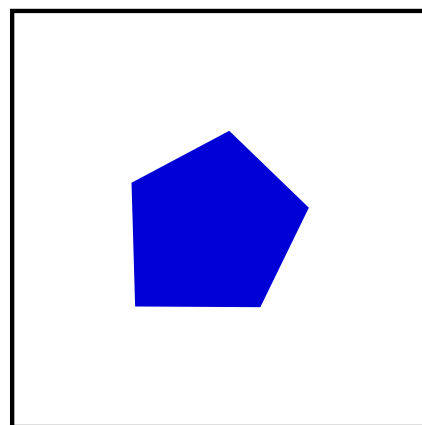
**Рис. 8.1.1** Влияние ориентации репера на перцепт формы. Когда стороны неопределенного квадрато-даймонда диагонально ориентированы внутри прямоугольного репера, квадрат этот воспринимается как даймонд, когда репер стоит вертикально, но как наклоненный квадрат, если репер тоже наклонен.

Корферман, 1930.

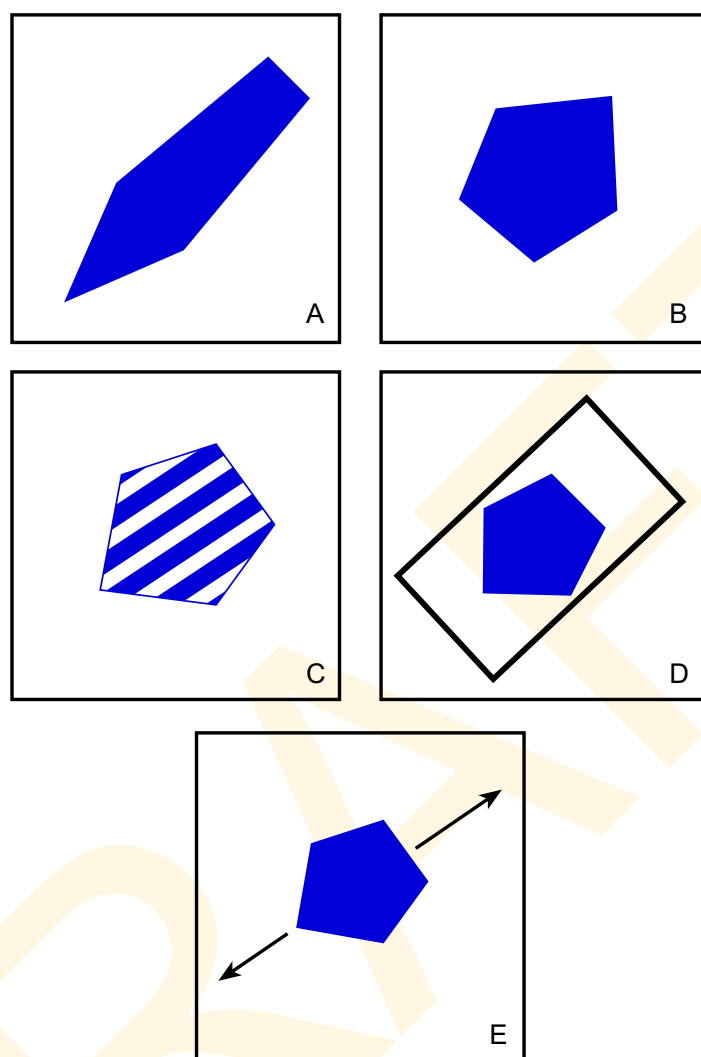
берется как передняя или верхняя у объекта, репер которого расположен по направлению движения.

Итак, описанные принципы выбора репера тем схожи с гештальт-принципами группировки, что работают в режиме *ceteris paribus*. То есть, они являются потенциально независимыми факторами, которые можно использовать для прогнозирования выбора репера только в «чистых» случаях или же в случаях, когда отдельные факторы конвергируют совокупно влияя на ориентацию объекта. Во многих ситуациях однако, многие факторы напротив — конфликтуют (см. рис. 8.1.12): большинство из нас воспримет эту фигуру прямостоящим асимметричным полигранником. Таким он и будет оставаться до тех пор, пока мы не обнаружим в нем наклоненную ось симметрии. Наиболее подходящим объяснением сего, пожалуй, является то, что объект этот смотрится прямостоящим благодаря строго горизонтальной ориентации нижнего контура. То есть, результирующий перцепт гравитационной стабильности усиливает естественное смещение в сторону вертикальной ориентации фигуры как ориентации реперной.

На рис. 8.1.13 А дан тот же самый объект, но сильно вытянутый по оси симметрии. Теперь намного легче увидеть его наклоненным, сим-



**Рис. 8.1.12** Неопределенная фигура. Обычно эта фигура выглядит асимметричным многогранником, лежащим на одной из своих сторон. Однако ее вполне можно увидеть симметричной, если мысленно поставить на правую верхнюю сторону, то есть, повернуть на  $45^\circ$  по вертикали.



**Рис. 8.1.13** Примеры влияния эталонных реперов на восприятие формы. На перцептивную ориентацию пятигранника может повлиять вытянутость (А), гравитационная стабильность В), ориентация текстуры (С), окружающего окна (D) и движение (Е).

метричным и стреловидным, указывающим вниз и влево. На рис. 8.1.13 В дана та же фигура, что и на рис. 8.1.12, но повернутая на  $180^\circ$ . В такой ориентации она скорее воспримется повернутой симметричной фигурой, указывающей вверх и вправо, поскольку гравитационно нестабильна. То есть, хочется повернуть ее против часовой стрелки так, чтобы основание встало строго горизонтально, а ось симметрии — вертикально. На рис. 8.1.13 С тот же объект с толстыми текстурными полосками внутри смещает наше восприятие в сторону наклоненного репера. На рис. 8.1.13 D вновь та же фигурка, но уже внутри прямоугольника, сторонами своими жестко задающего ее ориентацию. На рис. 8.1.13 E показано влияние движения по оси симметрии, что задает очень жесткий сдвиг восприятия в сторону наклоненной симметричной формы.

К сожалению, до сих пор не существует универсальной теории выбора репера, описывающей то, как множественные факторы объединить воедино, дабы прогнозировать результирующую ориентацию фигуры.

Стоит отметить, что теории эквивалентности форм, основанные на идее объектно-

центрированных реперов плотно связаны с таковыми, построенными на базе трансформирующей приводки. Поскольку в обоих случаях мы используем множество пространственных трансформаций, принимая, факторизуя (вынося за скобки) отличия между двумя объектами или вообще отказываясь от них. Но в концепциях пространственной приводки один из объектов трансформируется во имя точного соответствия другому объекту, тогда как в теории реперов сам репер трансформируется с целью соответствия структуре объекта. Трансформирующая приводка основана на согласовании двух изображений, тогда как сравнение реперов — на символическом описании формы внутри репера. Плюс к тому, теория приводки часто прибегает к опорным точкам (к примеру, Huttenlocher & Ullman, 1987), тогда как концепция реперов обычно базируется на поиске собственной ориентации, что реализуется через поиск разного рода симметрий, длинников, текстур и т.д. (Palmer, 1983, 1985). Отметим наконец, что до сих пор так и остается неясным, какой из подходов ближе к фактической работе зрительной системы человека.

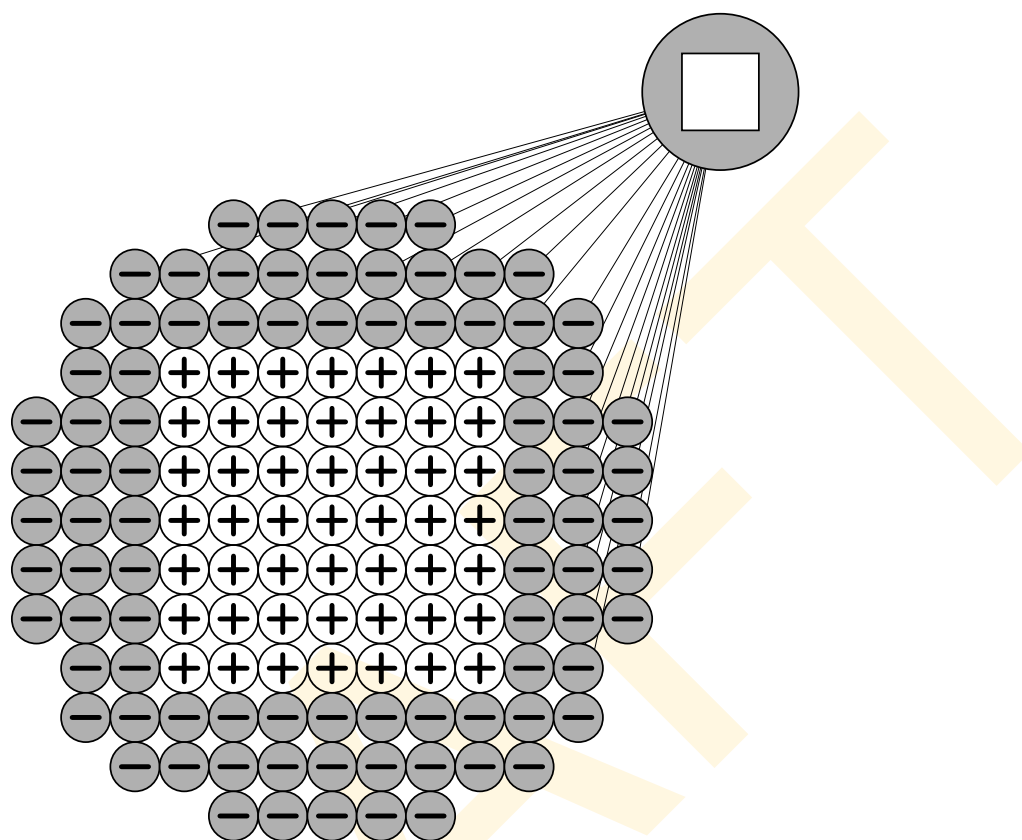
## 8.2 ТЕОРИИ ПРЕДСТАВЛЕНИЯ ФОРМЫ

Эквивалентность форм — это лишь один из аспектов зрительного восприятия формы объектов, касаемый вопроса их *идентичности*, но никак не затрагивающий ключевой проблемы *подобия* форм объектов, неидентичных по форме. Подобие форм — это куда более сложный вопрос, поскольку требует наличия теории, описывающей то, как следует представить форму, чтобы заполучить достоверную шкалу перцептивного подобия. Отметим, что в нашем разговоре об эквивалентности форм мы ни слова не сказали о представлении формы, поскольку первое сугубо бинарно и требует лишь ответа на вопрос: совпадают объекты по форме или нет? Сие упрощение позволяет нам, не имея на руках полноценной теории восприятия формы, определить условия, в которых два объекта будут обладать одной перцептивной формой. В данном разделе мы поговорим о различных потугах, претендующих на объяснение того, как форму объектов и их частей можно фактически представить в зрительной системе человека и как оба эти представления можно сравнить по критерию подобия.

Легко догадаться, что все теории, которые нам предстоит разобрать неадекватны той изумительной мощи, универсальности и утонченности перцепции формы зрительной системой человека. Вопрос о том, за счет чего мы безошибочно воспринимаем форму предметов — это один из списка наиболее сложных вопросов зрительного восприятия, настолько сложных, что до сих пор не прозвучало хоть сколько-нибудь внятного ответа на него. Но все же имеется некоторый прогресс, и многое уже понято. Однако, поскольку безусловного решения пока нет, нам придется разобрать как сильные, так и слабые стороны наиболее популярных гипотез.

### 8.2.1 Шаблоны

Идея *шаблонов*, представляющих форму как форму, весьма близка рассуждениям первых структуралистов. В шаблонных представлениях форма описывается через объединение рецепторов, на которых оптически сформировалось проекционное изображение данного объекта. Шаблон квадрата может появиться, к примеру, за счет объединения всех простимулированных рецепторов, плюс окружающих непростимулированных. Самый простой путь реализации этой идеи — это объявить о существовании «клетки-детектора квадратов», чье рецептивное поле по форме своей повторяет форму нас интересующую (в данном случае, квадрата на некоем фоне). Как показано



**Рис. 8.2.1** Шаблонное представление квадрата. Клетка — детектор квадрата (справа сверху) возбуждается сигналом рецепторов центрального участка (белые кружки с «+») и тормозится рецепторами окружения (серые кружки с «-»). Такая клетка максимально возбудится при стимуляции белым квадратом на темном фоне и максимально затормозится при стимуляции темным квадратом на фоне светлом.

на рис. 8.2.1, сие значит, что центральный — квадратный участок рецептивного поля будет оказывать возбуждающее влияние на клетку квадрат-детектор, а окружающий регион получит ингибирующее соединение с ней (что нужно для профилактики отклонений от квадратности, тем более, что, как известно, у корковых рецептивных полей в основном ингибиторное окружение). Так, к примеру, белый квадрат на черном фоне максимально возбудит клетку квадрат-детектор, поскольку его пространственная структура оптимально согласуется с таковой у рецептивного поля этой клетки. Мы назовем такое представление *стандартным шаблоном*, дабы отличать его от более сложных альтернативных представлений, о которых нам предстоит поговорить позднее.

Атомарные элементы, из которых собран шаблон, обычно имеют сугубо бинарную характеристику, соответствующую тому, возбужден данный конкретный рецептор или нет. Сие значит, что каждый элемент в шаблоне либо активен (т.е. отвечает положительно на световую стимуляцию) или заторможен (т.е. отвечает отрицательно на воздействие светового потока). Однако сия бинарная схема не единственно возможная: мы можем довольно легко соорудить «полутонные» шаблоны, конвертировав представление каждого пространственного элемента из бинарного вида (черный/белый по каждому из рецепторов) в непрерывную серую шкалу, плавно меняя мощность связей от +1 (возбуждающая) до -1 (ингибиторная). Таким манером, мы можем построить множество шаблонов для локальной синусоидальной решетки, то есть, организовать базис



локальной пространственно-частотной концепции восприятия формы. Более того, мы можем даже соорудить шаблон «детектора бабушки», способный в специфических условиях стимуляции дать максимальный отклик на изображение лица нашей бабушки (при определенных угле просмотра и расстоянии до источника света в определенном положении). Однако, как мы покажем далее, тот факт, что полутонные шаблоны возможны, вовсе не означает, что вопрос восприятия форм решен.

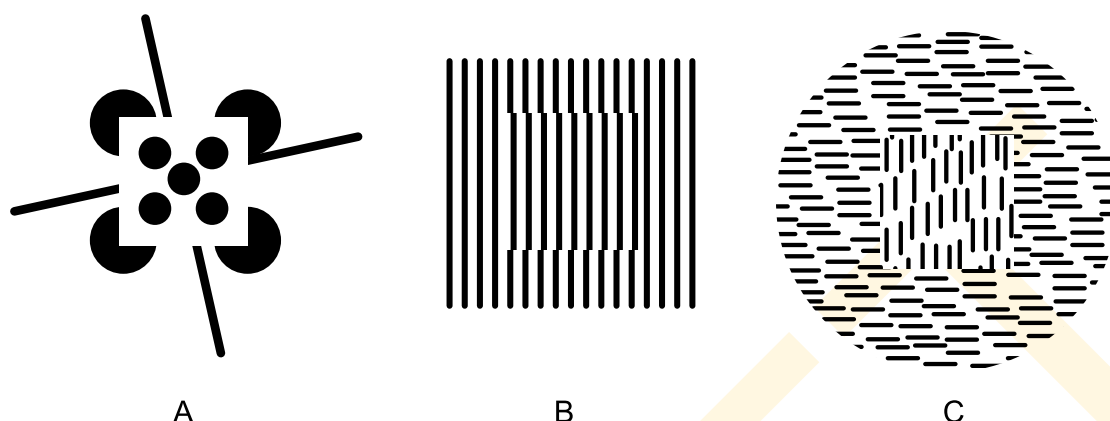
Чтобы установить подобие между двумя формами, нам понадобится методика сравнения представлений. Два шаблона можно сравнить простым расчетом степени их точечного соответствия или степени заполнения. Одна из простейших схем — вычисление корреляции между входным изображением и шаблоном. Степень эта ранжируется в диапазоне от  $+1.0$  по максимальному соответствию (два идентичных изображения) до  $-1.0$  по максимальному несоответствию (изображение и его черно-белая обращенная версия). Ожидаемое значение между двумя разными изображениями равно нулю. Сию меру сравнения можно использовать как с бинарными, так и с полутонными шаблонами.

### 8.2.1.1 Плюсы

На страницах научных монографий и статей идею шаблонов обычно высмеивают как совершенно неадекватную представлению формы в зрительной системе человека. Однако у нас есть «железобетонный» аргумент в ее пользу: шаблоны *обязаны* использоваться на определенном этапе нейронной обработки, а именно — при конверсии пространственной структуры ретинальных изображений в символичные представления. Шаблоны всегда служат отправной точкой рассуждений о представлении формы по той простой причине, что являются наиболее очевидным механизмом этой конверсии. Теория детекции линий и краев простыми клетками V1 — это теория шаблонного представления, хотя и очень примитивная. Согласно ей, каждая клетка, к примеру, клетка — детектор линий, дает максимальный отклик на линию определенного расположения, ориентации и контраста (светло-темная или темно-светлая). Частота пульсации падает, если один из этих параметров отклоняется от оптимального значения, давая тем самым критерий подобия локальной структуры ретинального изображения структуре рецептивного поля. Минимум пульсации наблюдаем при обратном распределении света внутри рецептивного поля.

То же справедливо и при Фурье-подходе к функции клеток V1 (см. раздел 4.2.2): каждая клетка эффективно рассчитывает степень подобия локальной пространственной структуры входному ретинальному изображению, но при кодировании используется габорова функция. При этом происходит конверсия ретинального изображения в локальные пространственно-частотные компоненты, но не в линии и края. Однако базовое представление все равно можно назвать одним из видов шаблонного соответствия. Отличие же будет в том, что локальные пространственно-частотные компоненты не рассматриваются как самодостаточный аспект формы объекта, тогда как линии и края полагаются таковыми.

Наконец, мы почти наверняка знаем то, что зрительная система использует шаблоны для представления таких простейших «форм», как линии и края или локальные фрагменты синусоидальных решеток. Главный вопрос в том, действительно ли сей подход можно применить к кодированию сложных форм реальных объектов, таких как квадрат, авиалайнер, пароход или лик нашей драгоценной бабушки.



**Рис. 8.2.2** Зрительные каналы в шаблоне квадрата. Мы воспринимаем квадраты, заданные мнимыми контурами (А и В) и текстурными краями (С) даже тогда, когда эти паттерны не вызывают высокоуровневой активации в яркостном шаблоне квадрата с рис. 8.2.1.

### 8.2.1.2 Минусы

Даже приняв то, что шаблоны — это неизбежно первый шаг в создании символического представления формы, мы должны понимать, что как общая теория такого представления, эта концепция не выдерживает критики. Рассмотрим наиболее очевидные ее про-  
ре-  
хе-  
и будем использовать их в качестве опорных позиций в анализе прочих теорий.

**1. Множественность сенсорных каналов.** Большая часть зрительных факторов оказывает весьма слабое влияние на перцептивную форму объекта, но при этом жесткое на корреляцию между шаблоном и ретинальной проекцией этого объекта. Возможно, наиболее очевидным фактором является отличие в окраске. Зеленый квадрат на желтом фоне по форме выглядит таким же как и синий на красном (или в какой угодно иной комбинации окрасок). Сие позиционирует проблему стандартизации шаблонных представлений, поскольку представления эти определены в логике специфического рецепторного отклика. Шаблонное представление зеленого квадрата желтом фоне не будет соответствовать во всем остальном идентичному шаблону для синего квадрата на красном фоне, поскольку лежащие в его основе хроматические элементы различны. Общая проблема в том, что ретинальные шаблонные представления не инвариантны при смене окраски объекта, тогда как перцепт формы — инвариантен.

Итак, очевидно, что ситуация с окраской — это лишь макушка большого айсберга. В структуре зрительных каналов множество различий, из-за которых представление квадрата может оказать радикальное влияние на любое стандартное шаблонное представление, но при этом минимальный эффект на восприятие самой квадратности. Вспомним, что квадрат можно увидеть в т. ч. через иллюзорные контуры, отрезки линий, различные текстуры и т.д. (см. рис. 8.2.2), но перечисленным способом представления квадрата не исчерпываются. Например, общий ток может вызвать перцептивную группировку в квадрат точек, движущихся на некоем статичном фоне или же фоне движущемся в противоположном направлении. Бинокулярная диспаратность способна оквадратенную стохастической стереограммой группу точек перцептивно отодвинуть в глубину от фона (см. рис. 5.3.8). Контуры квадрата могут также проступить со временем через тонкую апертуру (см. раздел 10.2.4) и т.д. Основная сложность в том, что существует множество различных зрительных каналов, через которые можно представить контуры квадрата и способных использовать принципиально разные режимы сенсор-

ной стимуляции. Следовательно, тот тезис, что шаблоны формы можно организовать, собрав вместе множество периферических сигналов, представляется весьма спорным.

Существует два пути ухода от описанной сложности. Первый путь — это создание отдельного шаблона по каждому сенсорному каналу. Однако тогда вместо одного мы получим множество шаблонов квадрата. Второй путь — определить шаблоны представления формы на некоем абстрактном уровне работы зрения причем после того как все контур-экстрагирующие каналы (окраска, текстура, диспаратность, движение, мнимые контур и пр.) выполняют свою работу. В этом случае всё, что потребуется — это шаблон, который каким-то образом представляет отличия между (квадратом) и фоном безотносительно к тому, какой сенсорный канал доставил информацию о контурах. Такой шаблон должен явиться общим для всех каналов.

**2. Пространственные трансформации.** В первом разделе текущей главы мы обсуждали значимость того факта, что перцептивная форма инвариантна при трансформациях подобия — переносах, поворотах, масштабировании, отражениях и их всевозможных комбинациях. Одной из основных сложностей, возникающих при использовании стандартных шаблонов, является то, что сравнение шаблонных представлений двух объектов одной формы в целом не дает точной приводки, если они отличаются только одной из таких трансформаций.

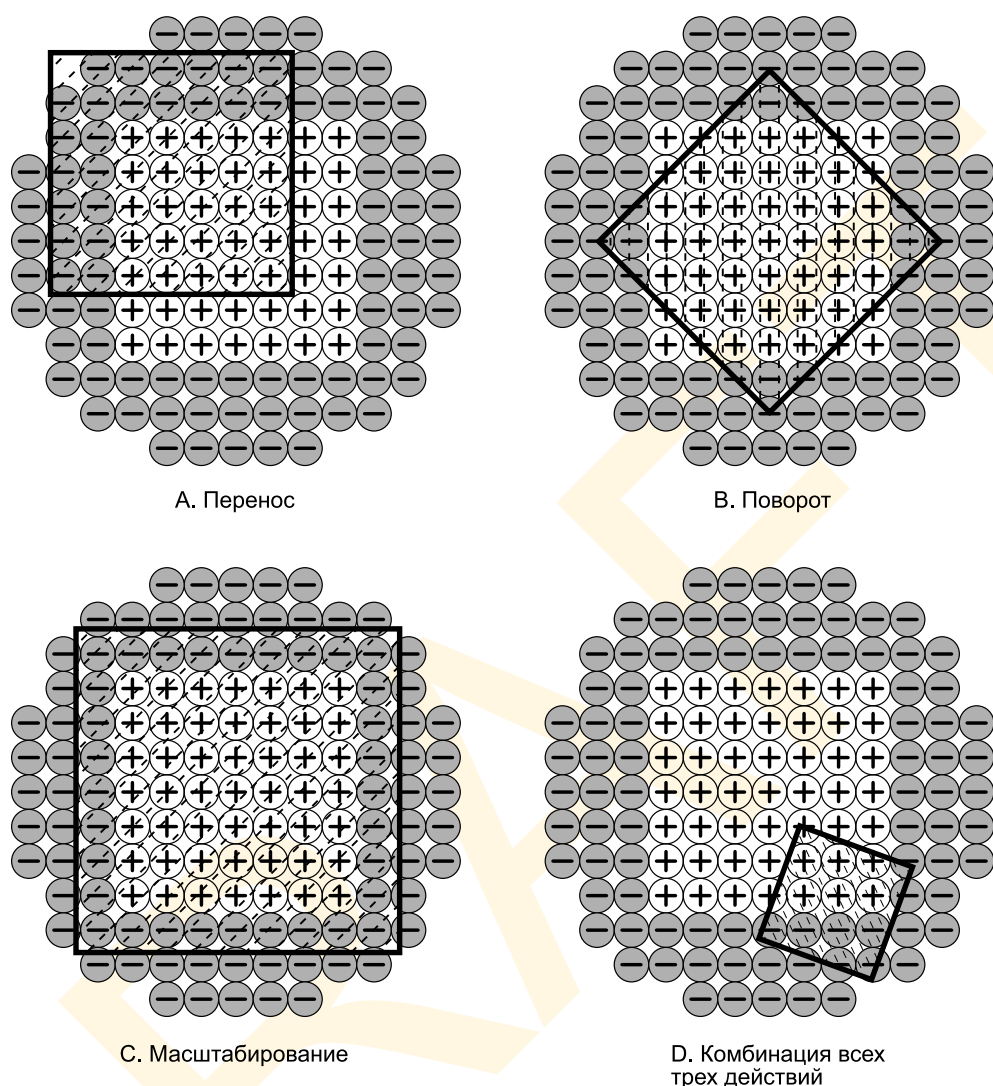
На рис. 8.2.3, к примеру, показано, что, ежели сравнить второй квадрат с первым, изменив расположение, ориентацию и размер, то влияние на корреляцию будет огромным. Поскольку у этих объектов одна объективная форма (см. раздел 8.1.1), простой расчет корреляции между двумя шаблонами не даст удовлетворительной меры подобия форм.

Однако существуют как минимум два пути решения этой проблемы: *репликация* и *нормировка*.

**Репликация** — это тактика построения различных шаблонов для каждой формы в каждом возможном ее расположении, ориентации, размере и отражении. После чего шаблон, наиболее подходящий к целевой фигуре, явится более или менее приемлемой аппроксимацией их «истинного» подобия в рамках принятой схемы подобия форм. Судя по всему, зрительная система пользуется такой подход на уровне V1 в отношении линий и краев (и/или локальных пространственно-частотных компонент), повторяя ту же структуру рецептивных полей по всем возможным ретинальным позициям, ориентациям, размерам и отражениям.

Как легко догадаться, сей подход действенен только в том случае, если множество шаблонов формы малочисленно. И линии, и края (или габоровы функции в косинусной или синусной фазах) — это действительно малые множества, однако если все распознаваемые нашим зрением формы повторить по каждой из позиций, каждой из ориентаций, размеров и отражений, шаблоны расплодятся так, что исходом явится т.н. *комбинаторный взрыв*, поражающий воображение. Следовательно, репликацию шаблонов вряд ли можно назвать подходящим решением вопроса представления формы, хотя оно может стать удовлетворительным при минимальном наборе очень простых шаблонов, таких как линии и края.

**Нормировка** — это принципиально иной подход к решению вопроса пространственных трансформаций, весьма тесно связанный с трансформирующей приводкой и внутренними реперами. Предположение таково, что внятную теорию шаблонного подобия форм можно построить без упоминания о репликации шаблонов, если ретинальное входное изображение вначале трансформировать или «нормировать» к набору реперных стандартов. Например, расположение фигуры можно нормировать относительно ее центра масс, ориентацию — относительно длинника и т.д., т.е. почти так, как об-

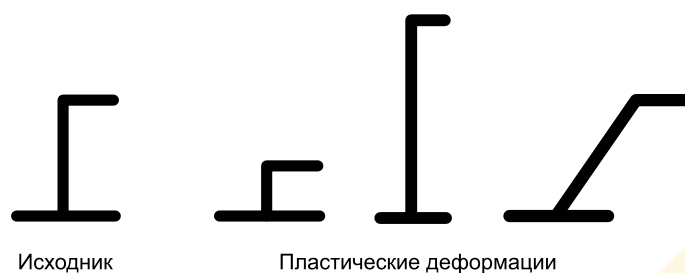


**Рис. 8.2.3** Пространственные трансформации в контексте шаблона квадрата. Мы вполне успешно воспринимаем заштрихованные участки А, С и D как квадрат независимо от того, что пространственно они слабо согласуются с шаблоном. Более того, мы не воспринимаем даймонд в части В квадратом, даже несмотря на то, что он неплохо ложится на шаблон.

суждалось в разделе 8.1.4, когда мы говорили об объектно-центрированных реперах. Поместив сей нормирующий процесс в систему до начала процесса сравнения шаблонов на подобие, мы сможем закрыть вопрос пространственных трансформаций. Но, скорее всего, сие ничего не даст в решении вопроса подобия форм, поскольку в процесс внедряются прочие трансформации.

Трансформации подобия очевидно достаточны, чтобы определить эквивалентность объективных форм, но, ежели мы обращаемся к более общей проблеме схожести (*подобия*) форм, к делу следует подключить дополнительные пространственные трансформации.

Рассмотрим пластические деформации, такие как растяжка, расплющивание и срез — рис. 8.2.4. Такие трансформации, могут изменить «количественную» форму фигуры, сохранив при этом многие из ее «качественных» аспектов. Однако простое срав-



**Рис. 8.2.4** Пластические деформации и подобие формы. Формы трех фигур справа очень схожи с формой исходной фигуры, несмотря на то, что весьма слабо «сажаются» на ее шаблон.

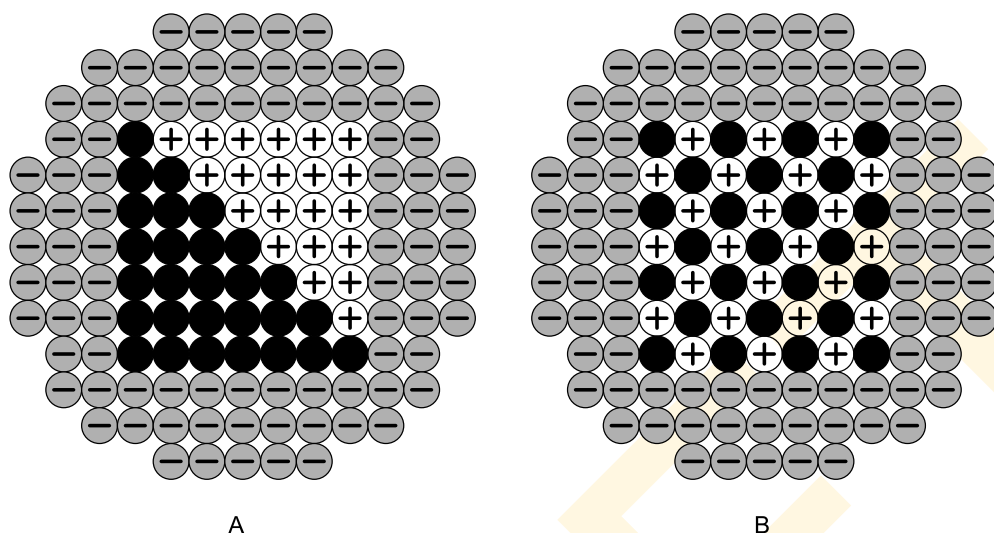
нение трансформированных фигур с использованием стандартных шаблонных представлений в целом не даст достаточного перехлеста, позволяющего говорить о высоком перцептивном подобию таких фигур. Устойчивые соответствия можно получить, применяя шаблоны, нормированные по отношению к этим дальнейшим трансформациям. Но выполнение сего рождает новые и неожиданные проблемы.

В качестве примера: расширенные нормировки ведут к тому, что определенные пластические трансформации, такие, например, как растяжка, могут изменить важные свойства фигуры, в частности ее длинник. Сие повлечет за собой изменение в ориентации, и, как следствие, изменение в результатах ее ориентирующей нормировки. К примеру, если длинник фигуры на рис. 8.2.4 использовался для нормировки ориентации, первый (сплюснутый) вариант не будет соответствовать стандарту, потому что длинник этой фигуры теперь стал скорее горизонтальным, нежели вертикальным. Мы можем уйти от этой беды, перепробовав все возможные трансформации, но такое решение позволяет избежать «размножения» шаблонов только только через невероятное усложнение обработки. Фактически, до сих пор никто не смог предложить удовлетворительной схемы нормировки шаблонных представлений, дающей без комбинаторного взрыва устойчивый прогноз перцепта подобия форм. Было предложено множество сложных систем, основанных на преобразовании изображений с целью их пространственной приводки друг к другу (см. раздел 8.1.3), но процессы согласования в них при этом основываются на очевидных признаках, таких как контурные выпуклости и вогнутости, но не на недифференцированных шаблонах. Мы рассмотрим такие системы подробнее в гл. 9, когда поведем разговор о распознавании объектов (например, Huttenlocher & Ullman, 1987; Lowe, 1985; Ullman, 1996).

**3. Дольчатость структуры.** В конце 7-й главы было показано, что большинство объектов мы воспринимаем состоящими из иерархически структурированных частей. Возникает закономерный вопрос: где в такой структуре место шаблонному представлению? Ответ предельно прост: таких шаблонов там нет. Существуют лишь два уровня шаблонного представления — цельный шаблон и атомарные элементы (рецепторы и пиксели, если речь об искусственных зрительных системах), ассоциируемые в шаблон. И ничего более. Следовательно, на базе шаблонов весьма сложно прогнозировать результат решения пальмеровой задачи распознавания (1977), в которой дольчатость структуры является ключевым фактором (см. раздел 7.6.1).

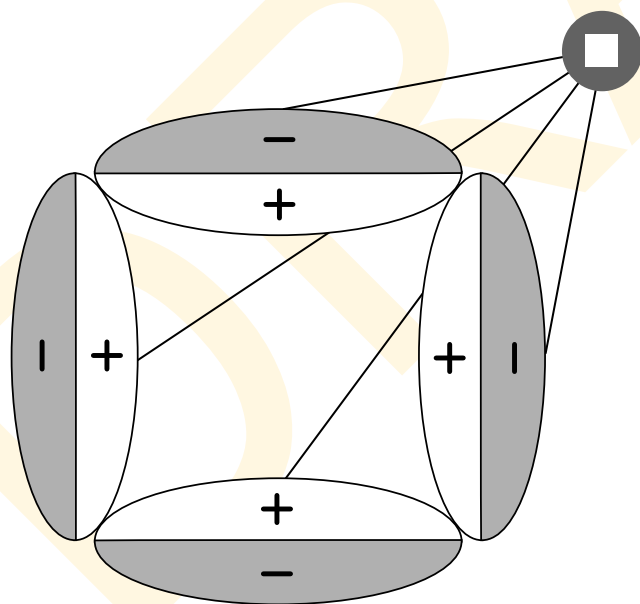
Утрата стандартными шаблонами дольчатости также вырастает в проблему в контексте разговора о характере долевого соответствия. Давайте в качестве примера рассмотрим два возможных пути сравнения некоторых элементов квадратного шаблона с таковыми у другого шаблона. Треугольник на рис. 8.2.5 А покрывает примерно половину элементов квадратного шаблона. Столько же — «точечный квадрат» на рис. 8.2.5 В. Но не найдется никого, кто скажет, что эти две фигуры подобны по форме квадра-





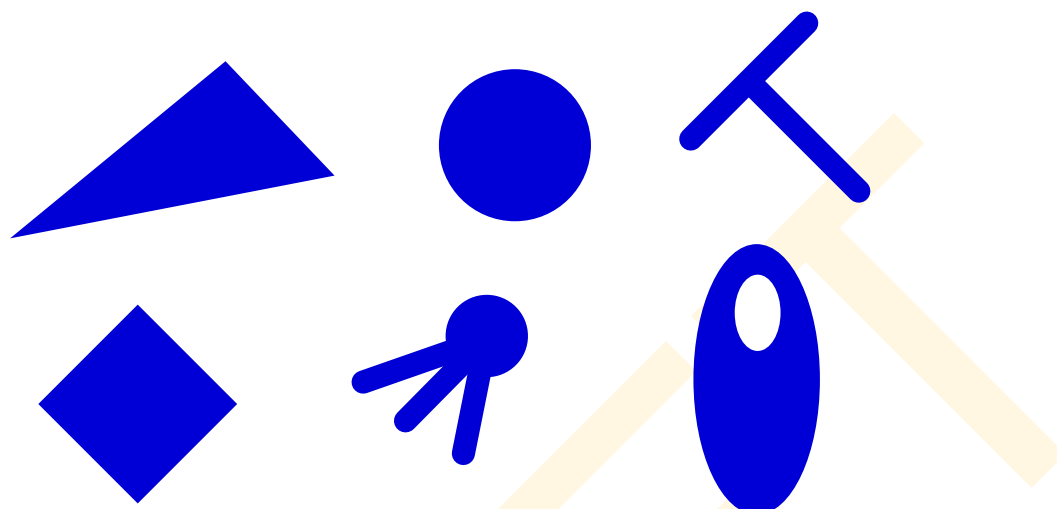
**Рис. 8.2.5** Проблема частичных соответствий. В обоих случаях (А и В) черные кружки заполняют половину кружков шаблона квадрата, но в А фигура воспринимается очевидно треугольником, тогда как в В — рыхлым квадратом. Концепция шаблонов не может оперировать такого рода различиями.

ту, представленному оригинальным шаблоном: треугольник сильно отличен, а точечный квадрат подобен лишь условно. Проблема проистекает из отсутствия в структуре шаблона квадрата дольчатости. Единственное, что совпадает в обоих случаях — это число элементов.



**Рис. 8.2.6** Иерархический шаблон квадрата. Клетка — детектор квадрата (справа вверху) получает сигнал от других шаблонов (в данном примере детекторы краев определенной ориентации и определенной полярности), но не от отдельных рецепторов.

Итак, что нужно для формирования дольчатости шаблонных представлений? Вопрос решается через построение т.н. *иерархических шаблонов*: комплексных шаблонов, организуемых сцепкой простейших шаблонов (но не элементарных единиц). Например, квадрат можно представить, объединив выходной сигнал четырех отдельных краевых шаблонов соответствующей пространственной организации (см. рис. 8.2.6). Приняв на веру сегодняшние представления о строении и функции зрительной системы, иерархические шаблоны начинают смотреться единственной жизнеспособной версией идеи шаблонов как таковой. Судя по всему, в V1 представлен лишь весьма ограниченный список шаблонов, и поэтому для обработки их выходного сигнала очевидно потребуется множество иных процессов. Как было показано в гл. 4 структура рецеп-



**Рис. 8.2.7** Определенные нейроны T1 откликаются на некоторые основные формы стимула, что позволяет как из блоков строить более сложные формы.

Tanaka et al., 1991

тивных полей клеток V1 — это, в целом, либо шаблоны краев и линий, либо пространственно-частотные шаблоны. Ежели так, то на высших корковых уровнях из этих примитивов должны строиться много более сложные анализаторы формы. Скажем, в области V4 обнаружались клетки высокочувствительные к паттернам стимуляции более сложным, чем простые линии, края и пространственно-частотные решетки, то есть, например, к фигурам, представляющим собой концентрические, радиальные, спиральные и гиперболические решетки (см. рис. 4.4.3). Однако по сей день неясно, как такие рецептивные поля формируются из выходных связей клеток V1 (Gallant, Braun & Van Essen, 1993).

Вспомним однако, что рецептивные поля клеток IT-области коры макака весьма крупны, составляя базис интеграции информации по большим областям поля зрения и анализа сложных паттернов стимуляции (Gross, 1973; Fuster & Jervey, 1981; Rolls & Tovee, 1995). Некоторые формы стимулов, к которым чувствительны нейроны IT, даны на рис. 8.2.7. Если клетки стриарной коры лучше откликаются на штрихи, края и решетки, нейроны IT дают наибольший отклик на сравнительно сложные, хотя и базовые, формы (т.е. круги, квадраты и овалы), что вполне может лечь в основу конструкции блоков, из которых затем строятся более сложные формы (Gallant et al., 1993).

Вероятно, та область коры обезьян, что получила название «IT», эквивалентна т.н. *латерально-окципитальному комплексу* (LOC) коры человека.<sup>1</sup> Методом функциональной МРТ было показано, что клетки LOC в коре человека лучше откликаются на отдельные объекты, но не их смесь или их фрагменты (Grill-Spector, 2001). При этом похоже, что они откликаются на все объекты подряд, но не на объекты какого-то определенного типа.

Сравним сие с более специализированными свойствами двух соседних корковых

<sup>1</sup> Несмотря на то, что кора человека и кора макаки анатомически весьма сходны, имеют место кое-какие функциональные отличия, побуждающие исследователей идентифицировать корковые области в головном мозге у представителей одного биологического вида (т.е. IT у макаки) как функциональный эквивалент иной области в коре представителей другого биологического вида (т.е. LOC у человека).

областей — веретенообразной и парагиппокампальной (Tong et al., 1998): клетки первой дают наибольший отклик на лица (и не на какие иные объекты) и вполне могут участвовать в детекции человеческих лиц и/или в их различении (Kanwisher, 1997); нейроны парагиппокампальной области дают отклик на объекты (например, дом) и их позицию, но не на лица, что указывает на их возможную роль в восприятии сцен (Epstein et al., 1999).

Отметим также, что у иерархии шаблонов имеется ряд любопытных свойств.

Во-первых, шаблоны физиологичны: сложные рецептивные поля строятся из простых, расположенных на нижнем иерархическом уровне (древние зрительные системы строго иерархичны). Стандартные же шаблоны не согласуются с физиологией зрительной системы, поскольку предполагают, что все формы можно представить непосредственно объединением атомарных ретинальных элементов.

Во-вторых, если требуется лишь небольшое число простых частей (например, прямых линий, краев и оконцовок — как в марровом опорном скетче), то иерархические шаблоны, судя по всему, потенциально способны разрешить проблему трансформации подобия, разрешив репликацию нужных частей во всех возможных позициях, ориентациях, размерах и отражениях. Как уже было сказано, судя по всему это то, что делает зрительная система в V1, повторяя структуру одного и того же рецептивного поля по всему полю зрения.

В-третьих, иерархические шаблоны можно применить к учету долевого строения сложных объектов, просто добавив в иерархию дополнительные уровни: углы можно собрать из линий и краев, простые формы наподобие квадратов и треугольников — из линий и углов, более сложные формы — из квадратов и треугольников, и т.д., и так до бесконечности. По этой и ряду других причин физиологи часто говорят о том, что нечто подобное иерархическому шаблону используется нашим зрением при кодировании формы сложных объектов.

Важно отметить, однако, что все эти выгоды даются ценой заметного усложнения представления формы. Мы более не рассуждаем в логике простого кодирования «формы как формы», но в логике кусочковой версии такого кодирования. Здесь важно следующее. К примеру, мы говорили, что такой простой критерий подобия форм как корреляция, нам более не подходит. Безусловно, она может использоваться на самом нижнем уровне, на котором располагаются стандартные шаблоны, но соответствие на более высоких уровнях требует несколько более сложной интеграции входных компонент, получаемых с нижних уровней. Плюс к тому понадобятся несколько базовых способов спецификации «подходящей организации» низкоуровневых шаблонов. Если базовые краевые и линейные шаблоны непосредственно соединить с единицами более высоких уровней, то возникнет старая проблема генерализации формы при трансформациях подобия, и для каждой такой трансформации понадобится отдельно взятый шаблон по каждой форме в каждой позиции, в каждой ориентации, каждом размере и т.д. Чтобы уйти от этой сложности, придется искать пути более общего описания организации базовых шаблонов. Как вариант — локализовать шаблоны низкого уровня по их координатам внутри объектно-центрированного репера. Второй способ — описать расположение частей относительно друг друга. Однако наиболее простая схема решения выходит за рамки темы шаблонных представлений, и мы опишем ее позже, когда будем говорить о классах представления формы, известных как «структурные описания».

**4. Трехмерность.** Если применение шаблонов в представлении формы простейших фигур (квадратов и треугольников), так скажем, проблематично, то с представлением формы трехмерных объектов — человеческого тела, собак, столов, стульев и пр. —

дело обстоит еще хуже. Причина в трехмерности. Шаблоны — это простой и примитивный способ представления формы как формы (и тем симпатичный), но идея резко сдает свои позиции, когда мерность представления меньше, чем мерность формы, которую оно представляет.

Существует лишь два пути создания размерности шаблонов, соответствующих трехмерным объектам, но ни один из них при этом нельзя назвать удовлетворительным. Первый — это создание трехмерных шаблонов. Сие уравнивает число размерностей у шаблона и у объекта, но, к сожалению, уравнивание сие опосредуется сетчаткой, которая лишь двумерна. Следовательно, трехмерные шаблоны должны строиться каким-то очень сложным образом, интегрирующим различные двумерные проекции в трехмерное представление. Второй путь — сделать внутреннее представление трехмерного объекта двумерным через двумерное представление его формы. Очевидной трудностью здесь является то, что понадобится большое число разных шаблонов. В крайнем выражении — по одному на каждую возможную перспективу. Ясно, что сие приведет к лавинной пролиферации (умножению) числа шаблонов, требуя, возможно сотен, если не тысяч для одного мало-мальски сложного объекта.

Однако специалистам машинного зрения удалось показать, что вполне можно ограничиться лишь небольшим числом шаблонов, если организованы процессы, высчитывающие промежуточные представления из представлений соседних (например, Poggio & Edelman, 1990; Ullman & Basri, 1991; см. раздел 9.3.3). К сожалению, сие осуществимо лишь ценой усложнения обработки и никак не решает других проблем шаблонopodobных представлений, в частности таких, как утрата дольчатости в структуре (см. Hummel) или зависимость от различных входных режимов сенсорных каналов. Мы вернемся к вопросу трехмерности еще неоднократно, поскольку сие камень преткновения многих теорий представления формы и не только в контексте шаблонов.

### 8.2.2 Спектры Фурье

Вторая принципиальная схема представления формы в зрительной системе опирается на пространственно-частотное содержимое ретинальных изображений или, как еще говорят, на их Фурье-спектр (например, Ginsburg, 1971, 1986; Kabrisky, 1966). Идея, лежащая в основе этого подхода заключена в том, что, если на начальных этапах зрительной обработки выполнить фурьеподобный анализ ретинального изображения по слагающим его пространственным частотам (см. раздел 4.2), то сие даст прочную основу представления формы, а также прочих высокоуровневых свойств зрительного восприятия, описываемых в логике пространственно-частотного концепта.

Итак, общий Фурье-анализ ретинального изображения складывается из двух спектров: *спектра мощности (амплитудного спектра)* и *фазового спектра*. Спектр мощности описывает амплитуду (светло-темный контраст) синусоидальной решетки на каждой из пространственных частот и ориентаций (см. раздел 4.2). Фазовый спектр характеризует фазу (пространственную позицию) компонент решетки на каждой из частот и ориентаций.

Сразу отметим, что еще самые первые психофизические опыты показали, что человек совершенно слеп к фазовой информации, когда решетки представлены на яркостно-контрастном пороге различимости (например Graham & Nachmias, 1971). Потому большинство Фурье-подходов к восприятию формы сосредоточены на спектрах мощности.

Одним из очевидно положительных свойств амплитудных спектров как базиса восприятия формы является то, что автоматически решается проблема эквивалентности форм при смене пространственного положения объекта: коль скоро информация

о пространственном расположении полностью зашита в фазовый спектр, спектр мощности оказывается инвариантен при смене локали стимуляции. В остальном, однако, спектры мощности — не более чем стандартные шаблоны представления формы, и без существенной доработки они не способны управлять прочими трансформационными аспектами эквивалентности форм, в частности такими как изменения в размере, ориентации и отражении. Например, спектр мощности большого квадрата и квадрата маленького окажутся отличными от спектров мощности двух совершенно идентичных квадратов, но расположенных в разных позициях.

В этом отношении стандартные спектры Фурье неудобны тем, что представляют *абсолютные* амплитуды синусоидальных компонент на *абсолютных* пространственных частотах и в *абсолютных* ориентациях (т.е. определенных относительно сетчатки). Но, как мы уже говорили ранее, форма более зависима от *относительных* свойств фигуры. Это значит, что предпочтительно кодирование *относительных* амплитуд, *относительных* пространственных частот и *относительных* ориентаций в представлении. Давайте назовем такое представление *объектно-относительным спектром мощности*.

Отметим, что объектно-относительное кодирование реализуемо по принципу, лежащему в основе объектно-центрированных реперов: простое кодирование амплитуд, пространственных частот и ориентаций относительно некоего репера определяется внутренними свойствами фигуры. К примеру, работа ведется относительно компонент с нижней (или основной) пространственной частотой, представленной в паттерне стимуляции. Мы уже сталкивались с этой схемой в гл. 4, когда описывали квадратную волну как сумму бесконечного числа синусоид:

$$F + 1/3(3F) + 1/5(5F) + 1/7(7F) + \dots,$$

где  $F$  представляет амплитуду базовой пространственной частоты квадратной волны,  $1/3(3F)$  — компонент утроенной базовой частоты при амплитуде в  $1/3$  от амплитуды основной волны и т.д.

Отметим, что все частоты и амплитуды описываются в объектно-относительных показателях: частота кодируется как произведение базовой частоты, а амплитуда — как доля от амплитуды базовой частоты. При такой схеме представления, все квадратные волны обладают одним и тем же объектно-относительным спектром мощности.

Аналогичная тактика применима к остальным размерностям. Например, информацию об ориентации объекта можно закодировать по отношению к ориентации, на которую приходится максимум энергии (Jañez, 1983). То есть, если начать кодировать спектры изображений в относительных показателях, то станет понятным, что вполне практичный критерий подобия форм можно рассчитать из соответствия относительных спектров мощности двух паттернов.

### 8.2.2.1 Плюсы

У Фурье-подхода есть ряд очевидных преимуществ. Во-первых, он не просто согласуется с теорией низкоуровневого зрения, но проистекает из нее. В разделе 4.2 мы уже озвучили кое-какие аргументы в пользу концепции пространственных частот: вспомним, что существует множество физиологических фактов, активно свидетельствующих о том, что некоторые виды локального пространственно-частотного анализа происходят в V1 зрительной коры. Словом, Фурье-подход к восприятию формы вполне применим в контексте низкоуровневой обработки ретинального изображения.

Еще одно весьма ценное свойство Фурье-анализа в свете разговора о восприятии



Еще одно весьма ценное свойство Фурье-анализа в свете разговора о восприятии формы — это его формально-математический статус. С тех пор как месье Фурье в 1822 г. озвучил свою знаменитую теорему, анализ стал весьма популярен, благодаря его развитой аналитической технике, применимой к обработке ретинального изображения. Математическая строгость метода Фурье контрастирует с прочими умозрительными теориями представления формы, чьи формальные «положения» в этой сфере мы снисходительно назовем «неточными». Локальный Фурье-подход, основанный на габоровой и вейвлет-функциях (см. гл. 4) хоть и не очень ясно нами понимаем, но зато четко и полностью формализован. Сие дает возможность приложить мощь математики к решению проблемы перцептивного подобия форм, причем, в отличие от прочих теорий представления формы — наилучшим образом.

Наиболее очевидное преимущество Фурье-анализа в данном контексте состоит в том, что относительные спектры мощности эффективно решают вопрос эквивалентности форм при трансформациях подобия. Иными словами, двумерные фигуры одной объективной формы в одних и тех же условиях просмотра обязательно дадут идентичные объектно-относительные спектры мощности. Относительные спектры мощности инвариантны также при отличиях в фотометрических (яркостных) контрастах, поскольку все амплитуды понижены или повышены пропорционально. Даже текстурные отличия можно потенциально исключить, отфильтровав высокие пространственные частоты, поскольку основная информация о форме передается на низких, а информация текстурная — на высоких пространственных частотах.

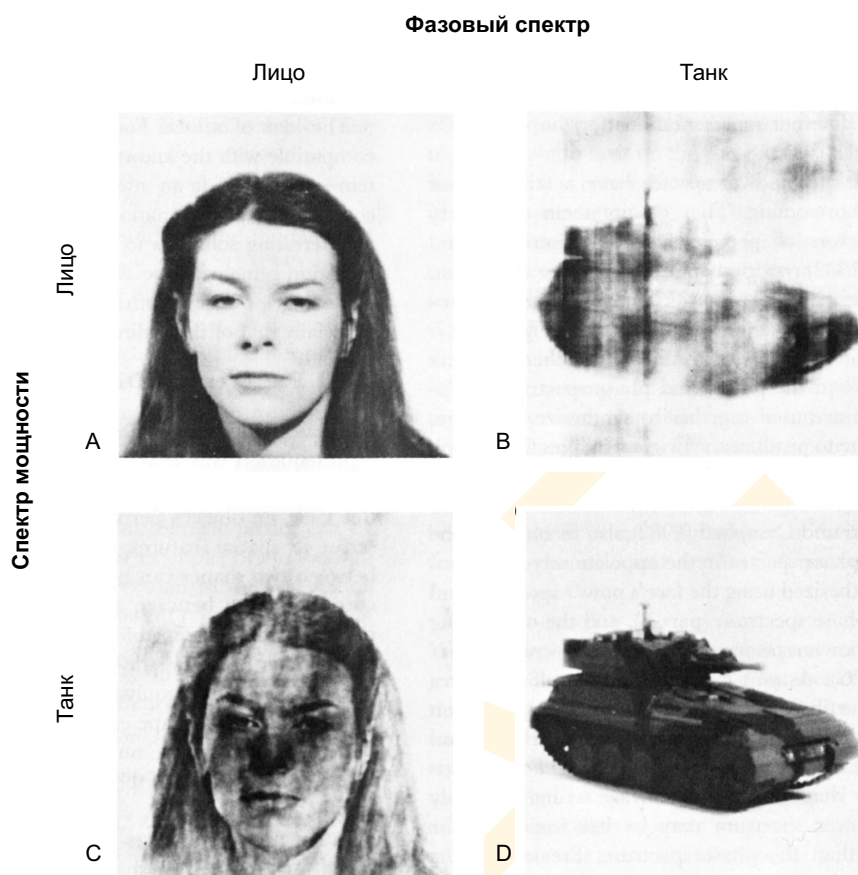
#### 8.2.2.2 Минусы

Несмотря на колоссальные преимущества перед остальными, идея представления формы объектно-относительными спектрами мощности имеет ряд недочетов, всерьез затрудняющих ее применение. Одной из основных проблем является то, что Фурье-анализ представляет цельное, неинтерпретированное ретинальное изображение, но не отдельные объекты. К примеру, относительный спектр мощности сцены не в силах отделить представление фигуры от представления фона. Сие значит, что представление формы фигуры окажется прочно сцеплено с этим фоном. И сие критически слабое место данного концепта, потому что один и тот же объект на разных фонах будет иметь принципиально разные спектры мощности.

Объектно-относительные спектры мощности также не в силах решить (и даже пролить хоть какой-то свет) на вопрос дольчатости структуры объекта и его трехмерности. Поскольку спектры мощности описывают неинтерпретированное ретинальное изображение холистически, они не включают в себя какого-либо эксплицитного представления естественно воспринимаемых частей. Действительно, способ, которым пространственная информация представлена в спектрах Фурье делает проблему практически неразрешимой, поскольку в отличие от шаблонов, не существует подмножества спектров мощности, образующих эксплицитное представление частей объекта.

Спектры Фурье не в силах решить и вопроса трехмерности, поскольку, подобно шаблонам, представляют двумерные ретинальные изображения, а не трехмерные объекты. Поэтому при различных углах зрения на один и тот же объект, кодирование будет идти разными спектрами, что приведет к огромному числу различных представлений формы одного сложного трехмерного объекта.

Наконец, у относительных спектров мощности имеется весьма серьезный эмпири-



**Рис. 8.2.8** Спектры мощности и фазовые спектры в восприятии формы. Каждый квадрант матрицы представляет комбинацию спектра мощности и фазового спектра лица девушки и танка. Когда объединяются собственные спектры, получается исходное изображение (слева сверху и справа внизу). Когда спектры не соответствуют друг другу (справа сверху и слева внизу), фазовая информация очевидно доминирует в восприятии.

Piotrowski & Campbell, 1982.

ческий недостаток: они вряд ли способны выступить добротным предиктором<sup>1</sup> перцепта формы.

Леон Пиотровский и Фергус Кемпбелл (Беркли, 1982) сообщили о весьма любопытной демонстрации: они провели общий Фурье-анализ двух изображений, расположенных вдоль положительной диагонали рис. 8.2.8: одно — лицо женщины (часть A), второе — танк (часть D). Объединяясь, спектры мощности и фазовый спектр одного и того же изображения дают изображение исходное (см. левый верхний и правый нижний квадранты рис. 8.2.7).

Пиотровский и Кемпбелл (1982) рекомбинировали спектры мощности и фазовые спектры исходных изображений: одно изображение попытались получить, объединив спектр мощности лица девушки с фазовым спектром танка (часть B), а другое наоборот — соединив спектр мощности танка с фазовым спектром лица девушки (часть C).

<sup>1</sup> Предиктор — величина или символ, прогнозирующие что-либо, например: снижение атмосферного давления — это устойчивый предиктор непогоды. — *Прим. пер.*

Хорошо видно, что фазовый спектр здесь явно доминирует над спектром мощности. Однако сия демонстрация не говорит о том, что Фурье-подход к представлению формы неверен, но лишь то, что спектр мощности менее значим для восприятия, нежели спектр фазовый. Напомним, что предыдущие исследования показали, что зрение человека нечувствительно к фазовой информации на пороге различимости (Graham & Nachmias, 1971), но что очевидно не распространяется на надпороговые уровни контраста.

Любопытно, что новорожденные в этом отношении принципиально отличаются от взрослых: экспериментально подтверждено, что они дольше удерживают взгляд на гибридах со спектром мощности, взятом с лица. Так продолжается вплоть до двух месяцев, пока малыши не начнут отдавать предпочтение изображению с фазовым спектром лица человека (Kleiner & Banks, 1987).

Отметим, наконец, что фазовый спектр кодирует информацию об *относительном* расположении компонент синусоидальной решетки на различных пространственных частотах и в различных ориентациях, особенно, когда компоненты пространственно выровнены друг с другом. Тот факт, что фазовый спектр доминирует в перцепции формы танка/лица, указывает на то, что сторонникам Фурье-подхода к представлению формы следует чем-то заменить исходный посыл, будто фазовая информация к делу не относится. Значимость ее становится нам очевидной, хотя до сих пор толком неясно, как с этим знанием обходиться.

Идея Фурье-анализа — это очень серьезная альтернатива сегодняшним представлениям о путях кодирования данных о пространственном расположении стимула. Однако нам предстоит еще управиться с другими сложностями, описанными выше и особенно — с трехмерностью.

### 8.2.3 Признаки и размерности

В последнее десятилетие в среде исследователей зрения наиболее популярным классом представления формы стал т.н. *список признаков* — символическое описание, являющее собой простое множество зрительных атрибутов объекта. Согласно сему подходу перцептивная форма объекта определяется множеством пространственных признаков этого объекта, а степень подобия двух форм поверяется мерой соответствия между двумя множествами признаков. Мы сталкивались уже с таким подходом в начале текущей главы, когда рассуждали об инвариантных признаках в контексте эквивалентности форм. Мы говорили тогда о том, что эквивалентность форм можно зафиксировать через такие признаки, как число линий и углов у объекта, их относительные длины и ориентации.

В целом, признаки, которые обычно полагают представляющими форму, делят на два класса: *общие свойства* (такие как симметрия, завершенность и связность) и *локальные* (прямые линии, кривые и острые углы). Оба типа свойств чаще всего представлены бинарно: симметрия либо есть, либо ее нет; фигура либо завершенная, либо нет; горизонтали либо представлены, либо нет. Тем не менее, концепцию можно развить, наделив ее способностью менять *степень* представленности тех или иных признаков. В этом случае объект уже не обязан быть либо симметричным, либо нет, но может иметь некую *степень* симметрии. Аналогично объект не обязан либо нести в себе вертикали, либо нет, но может иметь линии *примерно* вертикальные.

Представления, состоящие из нескольких размерностей, формально относятся к *представлениям многомерным*, но поскольку их сильные и слабые стороны в контексте представления формы плотно связаны с таковыми у стандартных дискретных пред-

ставлений признаков, мы будем говорить о всех сразу. Один из весьма значимых вариантов концепции непрерывных признаков можно сформулировать в логике *теории нечетких множеств* (например, Massaro & Hary, 1986; Oden & Massaro, 1978). Вспомним, что в гл. 3 мы говорили о том, что модель нечеткой логики цветовых наименований выступает продолжением теории множеств классических (Zadeh, 1965). Вспомним также, что главное достоинство концепта нечетких признаков — это его способность представлять эти признаки в количественных размерностях и более или менее сходно с тем, как мог бы работать наш мозг, если представление признаков действительно кодируется частотой пульсации нейронов. Частота пульсации клеток — детекторов краев и линий меняется скорее градуально, нежели бинарно, т.е. по схеме «всё или ничего». Нечеткие признаки позволяют представлять эти градации с разной степенью выраженности.

Итак, один из важнейших вопросов мультипризнакового и мультиразмерного подхода к представлению формы таков: каким путем мы можем определить виды признаков, которые должны быть включены в данное представление формы? Возможно, один из наиболее очевидных ответов звучит так: включить в список те признаки, что инвариантны при формосохраняющих трансформациях. Ранее мы убедились в том, что сей подход способен справиться с трансформациями подобия, хотя в данном отношении толку от него довольно мало (см. раздел 8.1). Также, возможно, он способен решить и прочие проблемы восприятия формы. Трудность же в том, что алгоритм не может тем же манером идентифицировать множество трансформаций, лежащих в основе подобия форм, каков применим при решении вопроса эквивалентности.

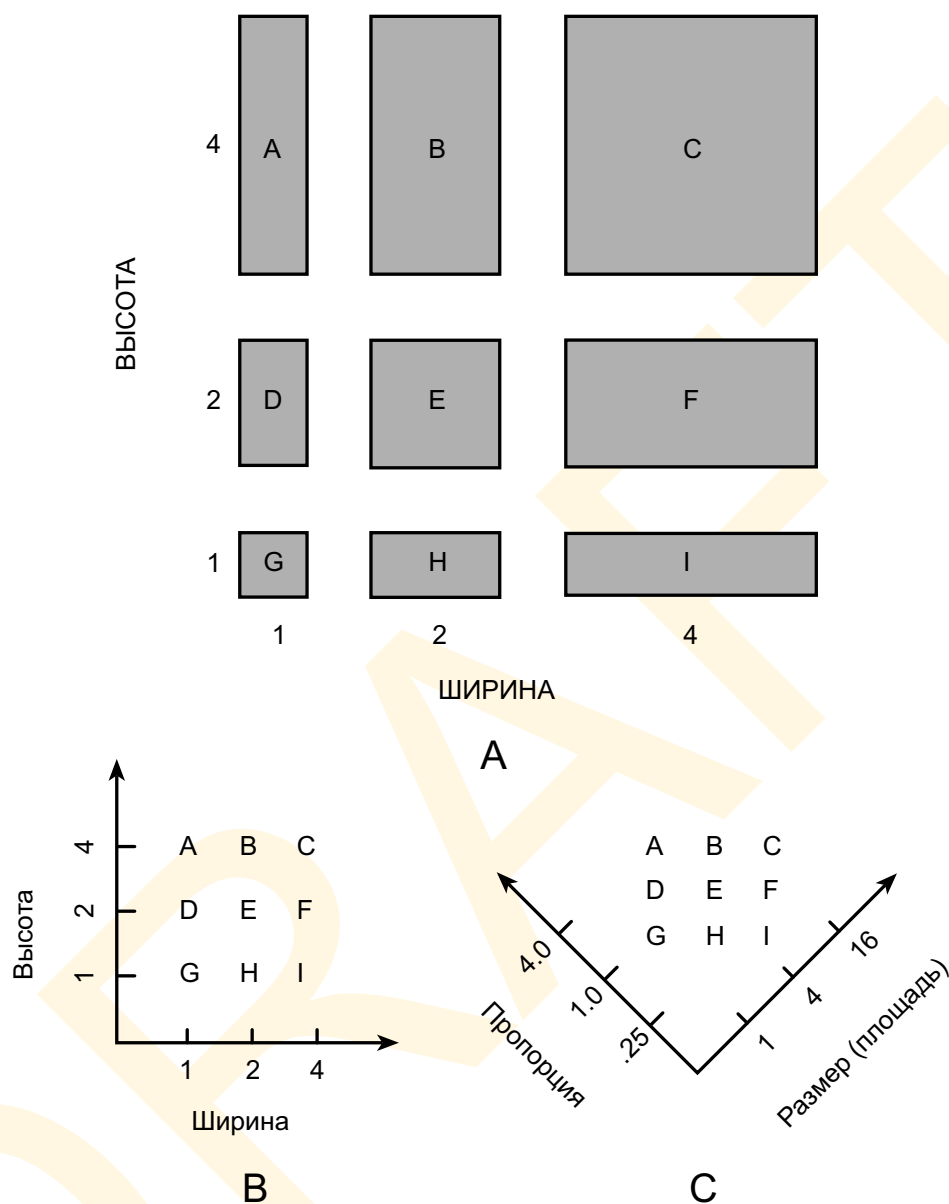
В отсутствие внятной методики, были предложены как минимум два весьма хитроумных способа создания представлений признаков из эмпирических критериев подобия: первый для построения стандартного бинарного представления признаков и второй — для построения многомерного (т.е., непрерывного). Однако же вначале мы рассмотрим непрерывное представление.

### 8.2.3.1 Многомерные представления

Естественный путь представления множества двумерных объектов — это определить их в виде множества точек в многомерном пространстве (от двух размерностей и более). Размерность признака задает оси пространства, и каждый объект будет точкой внутри этого пространства. Пример многомерного представления мы уже обсуждали в гл. 3, когда говорили о цветовых ощущениях как о точках в пространстве цветовом. Вспомним, что каждый цвет можно представить как точку в трехмерном пространстве, определенную позицией по осям цветового тона, насыщенности и светлоты (см. рис. 3.1.5). Идея, лежащая в основе многомерных представлений формы аналогична: модель того же типа можно использовать для внутреннего представления перцептивной формы сложных объектов.

Простейший пример — представление множества прямоугольников внутри двумерного пространства (рис. 8.2.9 А). Каждый прямоугольник можно определить через его высоту и ширину, то есть оба эти показателя можно взять для фиксации координат каждого из прямоугольников в двумерном пространстве (рис. 8.2.9 В), где координаты литер указывают высоту и длину объектов. Однако существует и альтернативная структура для того же множества стимулов — иной мерности, состоящая из общего размера прямоугольников и их пропорции (отношение высоты к ширине) — рис. 8.2.9 С. Отметим при этом, что конфигурация точек в этом пространстве будет той же, что и на рис. 8.2.8 В, но оси перешкалированы и повернуты на 45°.

Многомерные представления вряд ли станут интересными и практически полез-



**Рис. 8.2.9** Представление формы прямоугольников в многомерных пространствах. Множество из девяти прямоугольников определено тремя величинами высоты и тремя ширины (А). Следовательно, эти прямоугольники можно представить в многомерном пространстве, заданном ортогональными осями высоты и ширины (В) или осями пропорции и размера (С).

ными, если их применять только к тем фигурам, которые можно легко определить очевидными физическими размерностями (например, высотой и шириной). Однако, существуют машинные алгоритмы, именуемые *многомерными шкалирующими программами (MDS)*, способные из соответствующих данных внутри многомерного пространства автоматически строить представления, касаемые подобия или неподобия двух объектов. Такие программы исходят из того, что исследователь в состоянии выделить психологически релевантные размерности по данному набору объектов, но при этом без углубленного знания их. Проще говоря, MDS-программы — это потенциально весьма мощ-



	Фиолет	Синий	Зеленый	Желтый	Красный
Фиолет (434 нм)	99	85	40	25	70
Синий (472 нм)	85	99	70	25	25
Зеленый (504 нм)	40	70	99	55	10
Желтый (584 нм)	25	25	55	99	55
Красный (674 нм)	70	25	10	55	99

**Рис. 8.2.10** Матрица близости для пяти цветовых тонов. Каждый элемент представляет усредненную оценку подобия в паре стимулов, определяемую цветовым тоном в ряду и колонке по шкале от 0 до 99 (где 99 означает идентичность). Матрица близости была использована в качестве входного массива данных для работы программы многомерного шкалирования (MDS).

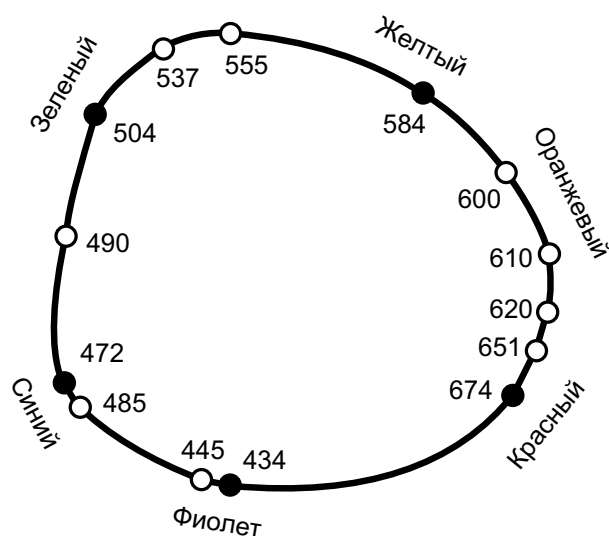
ный инструмент определения того, какие признаки могут лежать в основе внутренних представлений в зрительной системе человека при восприятии ею формы объектов.

Самые известные MDS-алгоритмы выведены из метода исходно сформулированного психологом Роджером Шепардом, долгое время работавшим в Bell Telephone Laboratories, а ныне в Стэнфордском университете. Основаны они на той общей идее, что перцептивное подобие двух объектов моделируемо расстоянием между соответствующими точками в многомерном пространстве. Чтобы построить многомерное представление  $n$ -объектов, MDS-программа берет в качестве входных данных т.н. *матрицу близости*  $n \times n$ : таблицу, состоящую из  $n$  рядов и  $n$  колонок, в которой каждый элемент — это число, представляющее подобие (или неподобие) ячейки в ряду — ячейке в колонке. На рис. 8.2.10 дан пример матрицы цветовых образцов.

Элементы матрицы близости обычно получают, предъявляя наблюдателям пару объектов, и прося затем по некоей числовой шкале оценить степень субъективного подобия (или неподобия) объектов в этой паре. В данном случае от 0 (максимальное различие) до 99 (идентичность). Средний рейтинг по множеству испытуемых — это тот или иной элемент по ряду и по колонке. Повторяя данную процедуру по каждой паре, строится полная матрица близости, поставляющая входные данные шкалирующей программе.

MDS-программа работает, отыскивая множество локалей внутри многомерного множества, что дает наилучшую стыковку двух показателей: рейтинга подобия объектов в парах матрицы близости и расстояния между соответствующими парами точек в многомерном пространстве. Программа начинает работу, разместив объекты в случайных позициях, а затем рассчитывает величину (обычно именуемую *напряжением*) того, насколько расстояния между парами точек в пространственной конфигурации стыкуются/не стыкуются с рейтингом подобия соответствующих объектов в матрице близости. Если конфигурация точек хорошо стыкуется с данными подобия, высокий рейтинг по парам объектов должен соответствовать малым межточковым расстояниям, а низкий рейтинг — большим. Когда сие истинно для данной конфигурации точек, напряжение будет низким; когда ложно — высоким.

Смысловым ядром MDS-программы является метод управления локалями объектов внутри пространства по уравнению, систематически понижающему напряжение при каждой последующей итерации. После серии итераций программа дает на выходе оп-



**Рис. 8.2.11** Реконструкция круга цветковых тонов с применением многомерного шкалирования. Используя матрицу близости  $14 \times 14$ , подобную той, что показана на рис. 8.2.10, шепардов оригинальный MDS-алгоритм восстановил более или менее циркулярную конфигурацию. Черные точки соответствуют подмножеству из пяти цветковых тонов с рис. 8.2.10.

Shepard, 1962b

сала при этом сами оси размерностей: дело в том, что в большинстве MDS-приложений оси могут находиться в любой ориентации и, следовательно, должны быть определены исследователем, который обычно в начале опыта проверяет конфигурацию точек в MDS-решении и находит множество ортогональных осей, в которых стимулы располагаются правильно. В данном случае обе оси очевидно искривлены (не прямые) и открыты (не закрыты), о чем свидетельствует регулярный порядок стимулов в этих размерностях.

Теоретически, с помощью MDS любое множество зрительных стимулов можно представить в многомерном пространстве. Когда на входе имеется матрица близости, то процедура направляется на поиск решения оптимальной подгонки в рамках многомерной схемы представления с заданным количеством размерностей.

Стоит сказать, что, несмотря на популярность MDS-представлений как моделей структуры сложных стимулов, у моделей этих имеются два существенных ограничения:

— во-первых, сфера их применения резко ограничена тремя и менее размерностями. Дело в том, что, дабы заставить систему работать, исследователь должен оценить конфигурацию объектов в пространстве. Выполнить сие с физическими моделями возможно максимум по трем размерностям. Правда, существуют методы восстановления структу-

туральную подгонку, которую мы полагаем «решением» в рамках той или иной размерности.<sup>1</sup>

Пример решения, предлагаемого MDS, дан на рис. 8.2.11. Речь идет об оценке подобия высоконасыщенных цветковых образцов (сплошные кружки — множество стимулов, чьи рейтинги подобия представлены в матрице близости на рис. 8.2.10). В данном случае нам уже известно, что представление должно являть собой двухмерную конфигурацию высоконасыщенных тонов, лежащих по краю цветкового круга, то есть исключительно совокупность точек, восстановленных алгоритмом MDS. Мы видим, что программа вполне корректно реконструировала двухмерное представление круга цветковых тонов.

Второй пример работы MDS дан на рис. 8.2.11 (Shepard, Kilpatrick & Cunningham, 1975). Здесь размерности перцептивного подобия заранее не были известны, но были найдены MDS-программой, которая восстановила конфигурацию точек, но не опи-

<sup>1</sup> Шепардов оригинальный метод (1962а, 1962b), позднее был переработан Крускалом (1964), что позволило выполнять т.н. *процедуру понижения градиента* (см. Приложение В), дабы гарантировать то, что напряжение понижено до минимума по каждой локале.

ры высокоразмерных решений (Wish, Deutsch & Biener, 170), но они весьма сложны и применение их затруднительно. И, все-таки, MDS-представления — это вполне достойный кандидат на теорию представления формы. Единственная проблема в том, что высокоразмерные решения, выполненные MDS, очень сложны в интерпретации;

— во-вторых, встает серьезный вопрос о том, насколько посылы, лежащие в основе MDS-алгоритма, схожи с тем, что происходит в зрительной системе человека при оценке им подобия объектов друг другу. К примеру, израильский когнитивный психолог Амос Тверски (1977) показал в свое время, что близость двух точек в пространстве может выступить добротной моделью тогда, когда оценка подобия симметрична — А должен быть подобием В, а В подобием А. Сие верно по определению расстояния в многомерном пространстве, но верно ли при перцептивной оценке подобий?

Вопреки очевидности такого рассуждения, в некоторых случаях оно оказывается неверным. К примеру, мы регулярно оцениваем периферийный красный как более схожий с фокальным красным, чем фокальный красный с точно таким же красным на периферии поля зрения (Rosh, 1975a). Амос Тверски и Итамар Гати нашли множество ошибочных посылов, лежащих в основе пространственных моделей, хотя большинство из этих моделей были куда более абстрактны и упрощенны, чем реальное восприятие. В отличие от первой проблемы, сие больше вредит жизнеспособности MDS-моделей. Некоторые исследователи пытаются найти некие ассиметричные позиции между точками (Krumhansl, 1978), но как правило толку от таких усилий очень мало.

### 8.2.3.2 Многопризнаковые представления

Второй подход к представлению перцептивно сложных объектов, лишенный недостатков своего предшественника, предполагает, что объекты можно анализировать по дискретным признакам, а не в непрерывных размерностях. Представление объекта в рамках многопризнакового подхода обычно понимается как некая совокупность (неупорядоченный список) его признаков. Квадрат, например, можно представить следующим комплектом признаков: завершенность, четыре стороны, четыре прямых угла, вертикальная симметрия, горизонтальная симметрия и т.д.

Если многомерные представления мы рассматриваем как мэппинг объекта на некую точку в  $n$ -мерном пространстве, многопризнаковые представления — это мэппинг объектов на узелки некоего структурного дерева. На рис. 8.2.13 дан пример *иерархического древа*, представляющего признаки нескольких прописных букв алфавита. В вершине этого древа имеется т.н. *корневой узелок*, представляющий множество всех объектов. Затем начинается разветвление по связям, представляющим наличие/отсутствие

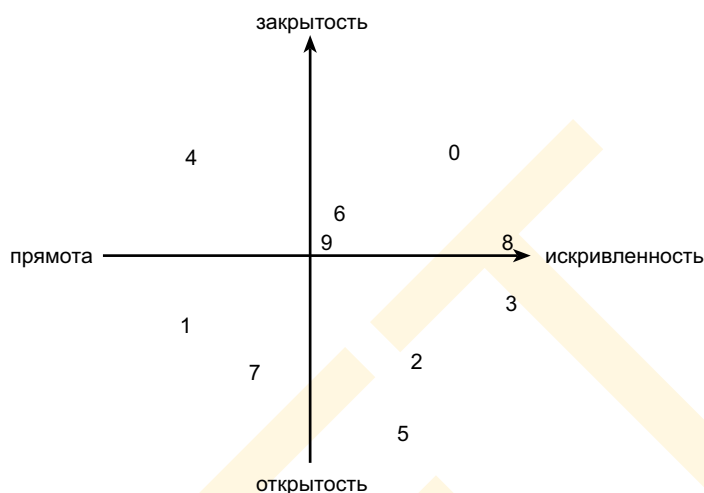
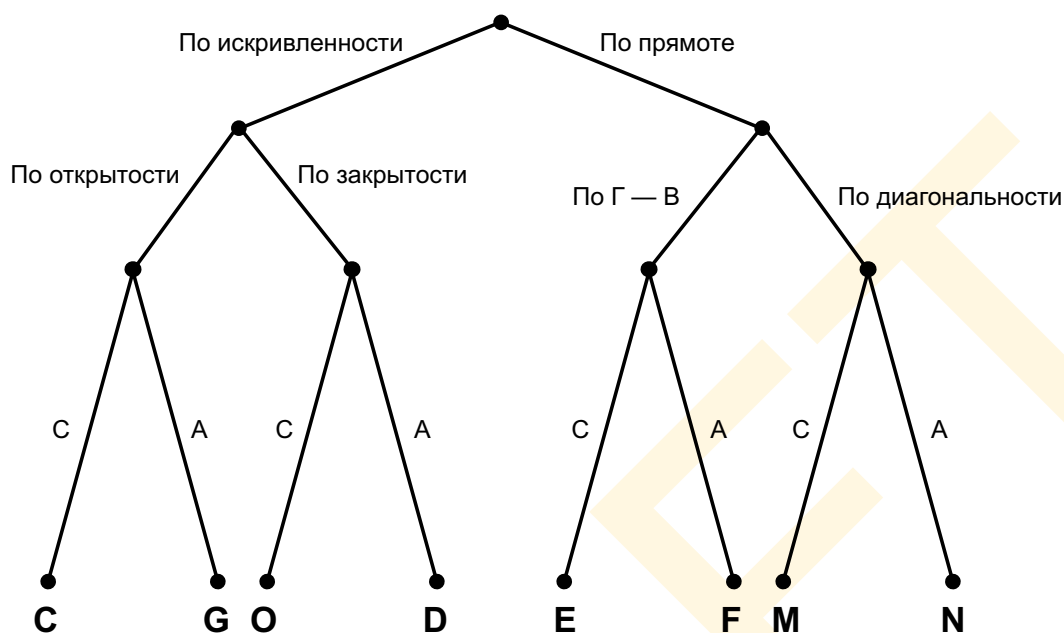


Рис. 8.2.12 Применение многомерного шкалирования в решении вопроса о подобии форм простых цифр.

Shepard, Kilpatrick & Cunningham, 1975.



**Рис. 8.2.13** Иерархическая кластеризация подобия перцептивных форм некоторых прописных букв латинского алфавита. Концептуальные узелки внизу представляют отдельные буквы, а верхние узелки — группы букв сходных формных признаков. («Г — В» — горизонтальные и вертикальные линии, «С» — симметричные, «А» — асимметричные).

различных признаков и идущих к *неконцевым узелкам*, представляющим различные уровни непересекающихся подмножеств объектов. Ветвление продолжается вплоть до *концевых узелков*, представляющих отдельные объекты. То есть, каждый объект оказывается определен множеством признаков связей терминального узелка с корневым. В данном примере (рис. 8.2.13) буква «С» представлена кривыми, открытостью и симметричностью. Чтобы отличить ее от других букв (скажем «U», которая тоже искривлена, открыта и симметрична), понадобятся дополнительные признаки. Однако мы в праве добавить в дерево дополнительные узелки и связи, дабы усилить его представительскую способность: используя разные связи для вертикальной и горизонтальной симметрий, к примеру, мы дадим вполне удовлетворительное определение разным концевым узелкам для «С» и «U».

Как и MDS-представления, многопризнаковые концепты оказываются работоспособными, только когда наличествует множество признаков, по которым идентифицируется интересующий объект. Чтобы реализовать сие существуют т.н. *программы иерархической кластеризации (иерархической группировки)*, конвертирующие  $n \times n$ -матрицу близости (подобия) в максимально подходящее дерево признаков того или иного типа. Базовый посыл этих программ в том, что расстояние между узелками дерева соответствует степени подобия между соответствующими объектами. Расстояние между двумя узелками в дереве оценивается как «большое» или «маленькое» между узелками (по связи между ними).

Первые кластеризующие программы строили иерархические древовидные структуры наподобие той, что на рис. 8.2.13, давая на выходе дерево, оптимально подогнанное под входные данные подобия. Но программы эти не могли дать *интерпретации* признаков, характеризующих связи в этом дереве. Пользователь вынужден был сам искать наилучшее толкование связям, характеризующим природу объекта (а заодно и смекал-

ку самого пользователя). Отметим, что интерпретация эта аналогична интерпретации размерностей в выходном продукте MDS.

Стоит сказать, что многопризнаковые представления позволяют уйти от некоторых проблем представлений многомерных, например: в отличие от практических ограничений MDS, состоящих в том, что пользователь без напряжения может осознать лишь три размерности, большое число признаков, представленных древом, пользователь с легкостью закладывает в работу программы кластеризации. Также в отличие от симметричных представлений MDS определенные типы многопризнаковых древ (скажем, объекты с отдельными связями по каждому направлению между парами узелков), могут представить ассиметричные отношения подобия. Но даже в этом случае трудно отрицать интуитивную привлекательность того, что многомерные представления дают возможность дискретного, количественного ввода данных в модель.<sup>1</sup>

### 8.2.3.3 Плюсы

Одной из причин популярности признаков представлений является то, что не так-то просто опровергнуть их действенность, в отличие от ущербных шаблонов или даже схем Фурье. И хотя до сих пор неясно, насколько признаковые концепты применимы на практике, выглядят они многообещающе.

Признаковые представления свободны от специфики входного сенсорного режима представления формы, поскольку не включают в себя каких-либо признаков, непосредственно относимых к типу периферической стимуляции, определяющей края или линии. Список признаков, которые мы ранее в этом разделе положили для квадрата, к примеру, не имеет отношения к его окраске, текстуре, расположению или размеру. Это всего лишь абстрактное символическое описание всех видов квадратов.

Признаковое представление, судя по всему, способно решить вопрос дольчатости структуры, просто включив различные части объекта в список. Список для квадратов, к примеру, включает в себя также такие признаки как «наличие четырех сторон» и «наличие четырех углов». Признаковое представление человеческого тела может включать в себя такие долевыe признаки как «наличие головы», «наличие туловища», «наличие двух ног» и т.д. В свою очередь, список признаков головы включает в себя «наличие двух глаз», «наличие носа», «наличие рта» и т.д. То есть дольчатость смотрится здесь куда менее проблемным фактором, нежели в представлениях шаблонных или Фурье-представлениях.

Концепция признаков также выглядит способной решить проблемы, проистекающие из трехмерности. Внутренняя размерность, которая ставит в тупик шаблонные представления и схему Фурье, не проблема для признакового подхода. Виды признаков, включенных в представление формы столь же относимы к трехмерным качествам и частям, что и к двумерным и вполне могут использоваться для формирования входных данных как по трехмерным, так и по двумерным объектам. К примеру, форму объекта так же легко описать как имеющую признак «сферическая», как «круговая», как «несущая пирамиду» или как «несущая треугольник». Следовательно, признаковый подход не ограничен двумерностью. Действительно, позиционировать трехмерную структуру оказывается столь же просто, что и двумерную.

<sup>1</sup> Дискретную структуру можно ввести в нечеткое многопризнаковое представление, но выполнение сего превращает это представление в формально эквивалентное представлению многомерному.

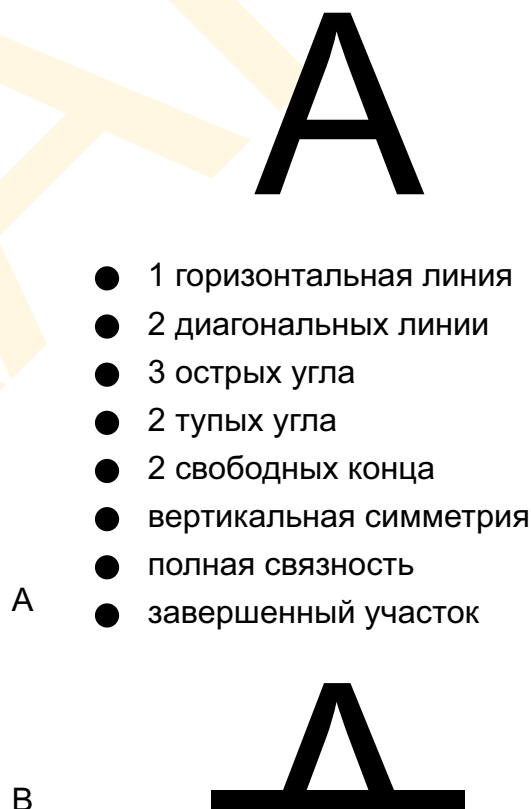


## 8.2.3.4 Минусы

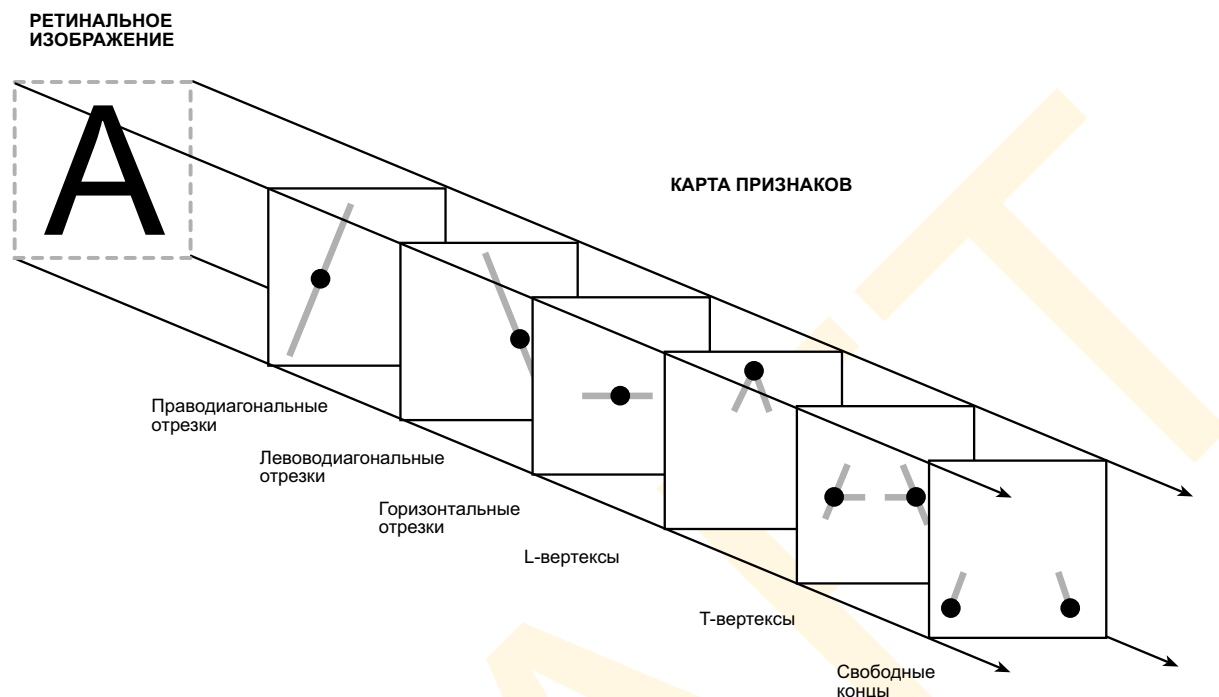
Несмотря на очевидные плюсы, у концепции признаков есть и очевидные минусы. Мы уже говорили об одном из них — о том, что теория инвариантных признаков в эквивалентности форм не работает, когда человеческое восприятие отклоняется от эквивалентности объективных форм. К примеру, идея инвариантных признаков пасует в случае махова квадратодаймонда (прямоугольного квадрата и квадрата, повернутого на  $45^\circ$ ): ежели добавить к системе специфические признаки ориентации, то у обеих фигур станут разными списки признаков, но ведь существует множество других фигур, которые мы воспринимаем обладающими одной и той же формой, несмотря на изменения в их ориентации (рис. 8.1.1 В).

Весьма серьезным минусом является и то, что зачастую невозможно понять, как количественно определить, имеются ли у объекта признаки, способные организовать представление его формы. Простые долевы́е признаки двумерных ретинальных изображений, такие как линии, края и пятна можно рассчитать по соответствующему шаблону (см. ранее), но даже здесь следует как-то абстрагироваться от специфических периферических каналов, по которым проходят эти линии, края и пятна. Более того, сии простейшие оптические признаки — это лишь верхушка очень большого айсберга. Они не покрывают всего изобилия различных атрибутов, закладываемых исследователями в представления формы. Легко сказать, что аспекты трехмерной формы можно представить в виде множества признаков, но куда сложнее их фактический количественный расчет. Такие признаки, например, как «наличие цилиндра» или «наличие носа» очень трудно высчитать из ретинальной стимуляции и нам неизвестна схема шаблона, позволяющего проделать такой трюк. До тех пор, пока процедура извлечения признака вызывает такие затруднения, концепция представления формы по признакам остается неполноценной.

Ежели все-таки исходить из того, что количественные сложности, возникающие при экстракции признаков, преодолимы, остается еще одна серьезная проблема — проблема описания того, какими могут быть собственные признаки представления формы: одно дело допустить, что некоего множества признаков формы достаточно для ее восприятия, но совсем другое дело точно сказать, каковы эти признаки. Машинные методы, такие как многомерное шкалирование и иерархическая кластеризация, иногда в силах помочь, но они по-прежнему не могут предложить жизнеспособную схе-



**Рис. 8.2.14** Признаки литеры А. В части А рисунка даны признаки прототипической литеры А, несущей в себе стандартные признаки литер алфавита. В части В дана фигура ровно с тем же набором признаков, но явно иной формы.



**Рис. 8.2.15** Представление литеры «А» картой признаков. Различные признаки «А» (несущие составляющие ее отрезки, вертексы и свободные концы) ассоциированы с их локалями в множестве пространственно-заданных карт признаков. Пространственные позиции этих признаков вносят дополнительные ограничения в описания формы, компенсируя тем самым ущербность представления простым списком признаков.

му представления формы в логике списка признаков. Чтобы понять, почему сие столь сложно, давайте рассмотрим пример.

Возьмем подходящий список признаков, которым можно охарактеризовать форму литеры «А» (рис. 8.2.14 А). Мы включим в него как общие признаки (симметрия, связность и завершенность), так и частные (три прямых линии, три острых угла и два тупых). На первый взгляд, вполне добротное представление формы. Но лишь до тех пор, пока мы не обнаружим, что прочие фигуры, с совсем иными формами располагают ровно тем же набором признаков. Один из списка дан на рис. 8.2.13 В: у фигуры ровно те же признаки, что и у представляемой литеры «А», но, как видим, форма объекта радикально иная. Сие неожиданно, поскольку в принятой логике, ежели две фигуры располагают одинаковыми наборами признаков представления формы, они должны бы восприниматься обладающими одинаковой формой.

Исследователи видят как минимум два варианта решения проблемы. Первый — обогатить представление парами признаков с их пространственными координатами в ретинальном изображении (см. рис. 8.2.15). Несмотря на то, что два разных объекта могут обладать одинаковыми наборами признаков, крайне маловероятно, чтобы признаки эти находились в одних и тех же пространственных локалях. Например, в данном случае описание расположения двух свободных оконцовок в ретинальном изображении окажется достаточным для того, чтобы отличить одну фигуру от другой (рис. 8.2.14 А и 8.2.14 В). Такие локально-специфические представления признаков (которые мы назовем *картами признаков*, дабы отличать их от простого списка)

дают дополнительные факторы формы, поскольку, напомним, информация о расположении никак не представлена в стандартном списке. Значимые для нас выводы можно сделать из того, как мы определили расположение признаков разных видов (к примеру, линий — через их центр; углов — через их вершины; общие свойства — через гравитационный центр объекта).

Идея, что форму объекта можно представить картой (а не простым списком) признаков, несет в себе определенный физиологический смысл, поскольку многие отделы зрительной коры действительно содержат карту ретинальных локалей (см. раздел 1.3.3). Если же разные признаки кодируются разными корковыми картами, как предполагают то многие исследователи зрения, сборный паттерн нейронной пульсации по всем картам зрительной коры явит собой функциональный эквивалент того, что мы называем *представлением картой признаков*. Как мы увидим в гл. 9 и 11, такие представления стали в последнее время весьма популярны в теориях детекции объектов (например, Lowe, 1985; Ullman, 1989, 1996), а также в теориях внимания, основанных на признаках, сцепленных с ретинальными локалями (например, Treisman & Gelade, 1980).

Второй путь — это предложить дополнительные признаки, позволяющие отличать две неидентичные формы с одинаковыми списками признаков представления формы. Например, мы можем добавить к списку тот признак, что горизонтальная линия была короче наклонных или что концы горизонтальной линии были соединены с их серединами. И хотя сие, фактически, признаки отличия, они явно специфичны для данной пары объектов. Даже если бы признак «наличие горизонтальной линии, которая короче наклонных» в целом был бы признаком, пригодным для представления формы (сие, однако, маловероятно) — число таких специфических признаков моментально взлетит до астрономических цифр. Еще проблематичнее то, что такие признаки составят новый класс, поскольку описывают *отношения* между некоторыми другими долевыми признаками — в данном случае, три образующих линии.

Итак, налицо слабое место представлений списками признаков: адекватному описанию формы объекта часто требуются не только описания частей и общих свойств, но также пространственные отношения между этими частями. Сии новые признаки не стыкуются с концептом списка как простого множества независимо описываемых атрибутов, поскольку внутренне зависят от признаков, представляющих части. Очевидно, требуется более сложная форма представления, непосредственно кодирующая информацию об отношениях между эксплицитно представленными частями. То есть, мы подошли к разговору о роли структурных описаний в представлении формы объектов.

### 8.2.4 Структурные описания

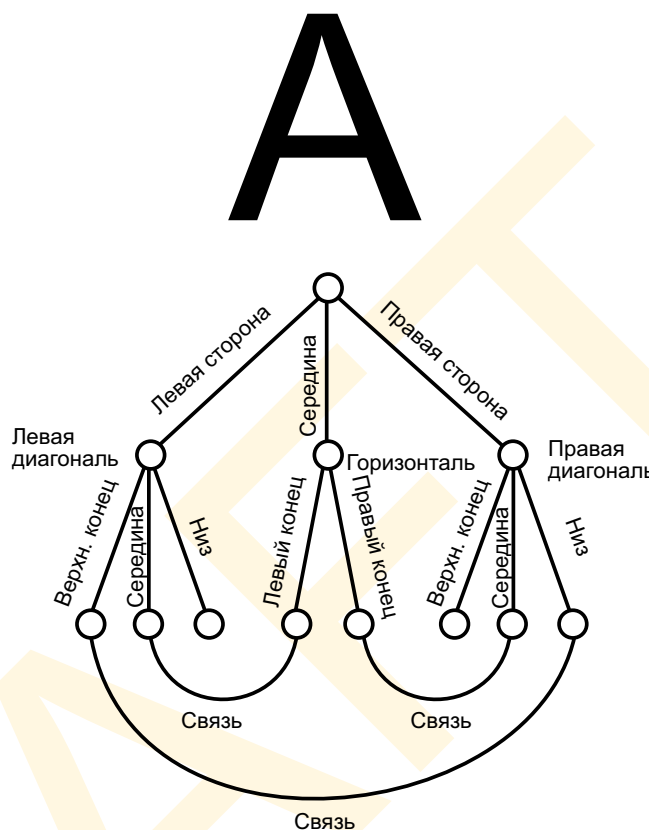
*Структурные описания* — это представления, в явном виде несущие информацию о частях объекта и взаимоотношениях между этими частями. Обычно структурные описания изображают в виде неких сетей, в которых узелки представляют цельный объект, а связи между узелками — отношения между частями. Чтобы проиллюстрировать то, как структурное описание справляется с задачами, неразрешимые списком признаков, рассмотрим структурное описание формы литеры «А», показанной на рис. 8.2.16. Верхние узелки представляют цельную фигуру, нижние — ее части. Относительные признаки, позволяющие максимально экономно отличать «А» от не-«А», представлены дугами между узелками (например, то, что оконцовки центральной линии крепятся к серединам наклоненных, эксплицитно представлен узелками для концов и середин

линий, а также отрезков линий, плюс связями между соответствующими узелками).<sup>1</sup>

Такое представление относительной информации намного экономичнее, соответствующих представлений в концепции признаков, поскольку отпадает нужда в списках этих признаков, да еще и во всех возможных комбинациях отношений между ними: сетевая структура позволяет пометать отношения отдельным символом (дуга), соединяя узелки друг с другом. Компромисс же заключен в том, что структура сети оказывается ощутимо сложнее, чем простой список признаков.

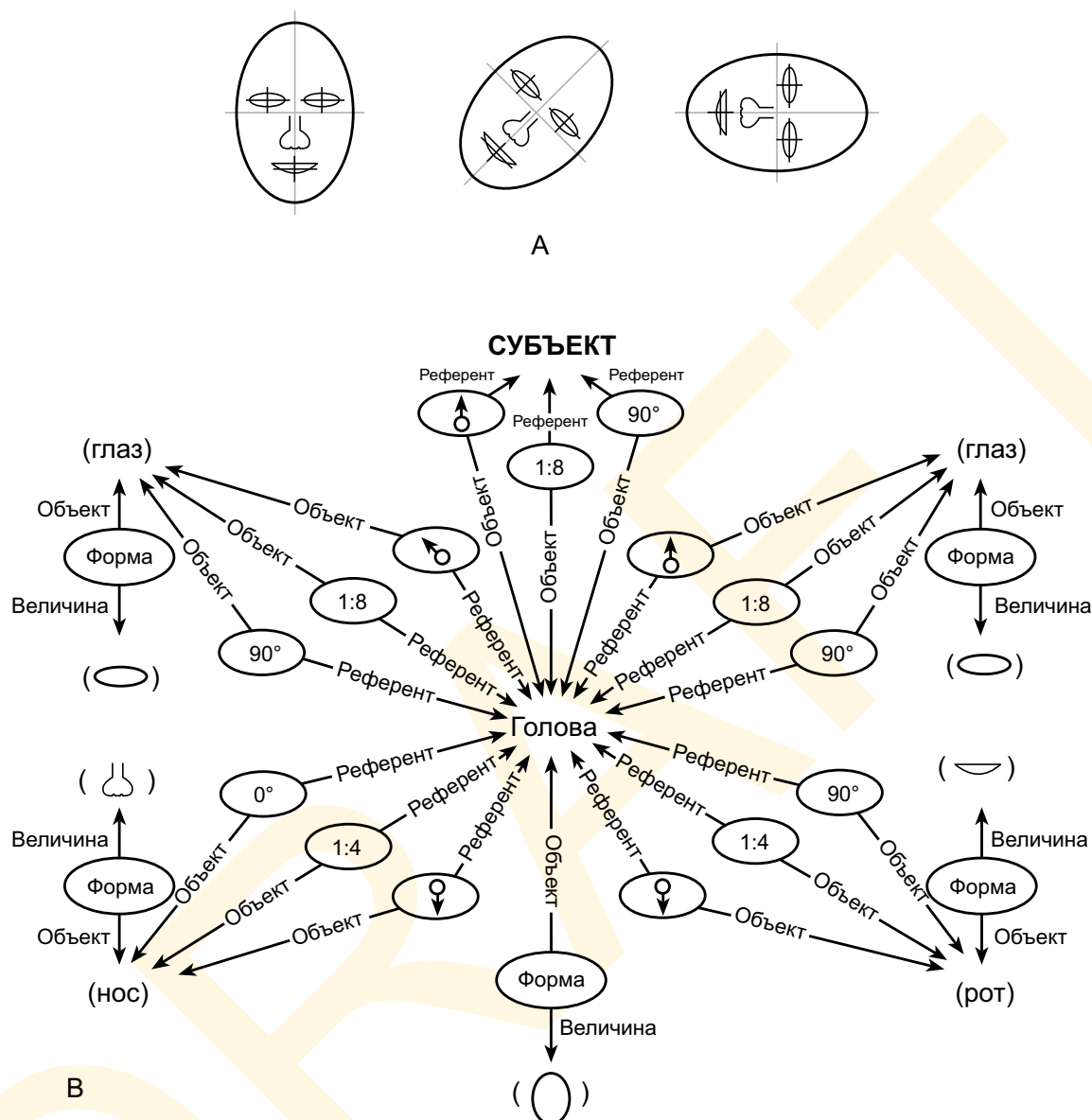
Стоит отметить также, что базовые структурные описания вносят очень ценные подбодия в иерархию шаблонов, и, главным образом потому, что одновременно представляют и части, и иерархию. То есть, структурные описания безусловно эффективнее при эксплицитном кодировании взаимоотношений между частями объекта. Более того, сети структурного описания можно строить по-разному, к примеру: общие признаки отдельных компонент можно представить, просто присвоив соответствующей информации узелок признака. Сие позволяет эксплицитно кодировать имплицитную информацию: симметрию объекта, его открытость, пропорции и т.д.

Особой ценностью в представлении формы обладает информация о внутреннем репере каждого структурного компонента. Рис. 8.2.17 продемонстрирует нам, как Пальмер (1975 b) применил эту идею в своем представлении схематической рожицы: голова как целое представлена овалом, чей внутренний репер определен позицией ее центра, ориентацией длинника и его длиной; аналогичным образом глаза представлены в виде овалов со своими собственными внутренними реперами, определенными их расположением, ориентацией и длинниками. Отношения между реперами можно весьма экономно закодировать, описав трансформацию подобия, которая соотнесет один репер с другим (т.е. комбинация переносов, поворотов, растяжек и отражений, приводящих реперы в полное соответствие). Результатом явится схема кодирования, представляю-



**Рис. 8.2.16** Структурное описание литеры «А». В сетке даны представления различных частей литеры «А» (компоненты линий и их концы) совместно со структурными отношениями между ними. (Подробнее в тексте)

<sup>1</sup> То, что наклонные линии длиннее центральной, можно также представить связью «длиннее чем» между соответствующими узелками.



**Рис. 8.2.17** Применение структурных описаний и реперов в представлении схематической рожицы. Часть А демонстрирует инвариантность отношений между внутренним репером головы и реперами глаз, носа и рта при поворотах; часть В — структурные описания, которые кодируют эти отношения путем специфических переноса (стрелки), поворота (размер угла) и масштабирования (пропорция), что соотносит каждый репер с каждым компонентом описания (голова, глаза, нос и рот).  
Pflmer, 1975b).

щая пространственные отношения между частями, причем в целом инвариантно при изменениях в абсолютном расположении, ориентации, размере и направлении, поскольку отношения между реперами кодированы друг относительно друга. Например, расположение, ориентация и размер глаз относительно головы не меняется, когда голова поворачивается, наклоняется или даже увеличивается (люминг). Сие позволяет структуре отношений объектно-центрированных реперов внутри сети оставаться инвариантной при любых преобразованиях подобия.



Отметим также, что нет никаких преград к тому, чтобы в структурных описаниях представлять и вариативные признаки: в данном случае структурное описание одного и того же объекта будет разниться при разных ориентациях, что и было показано в классических опытах Рока (1973) на фигурах со слабыми внутренними осями.

#### 8.2.4.1 Примитивы формы

Одним из замечательных свойств структурных описаний является их способность представлять общую форму компонентов объекта: полагаются т.н. *примитивы формы*, применимые как минимум для приблизительного кодирования общей формы каждого объекта или его частей. Вспомним, что в конце 7-й главы, рассуждая о доле сегментации, мы говорили о том, что в основе концепта формных примитивов лежит та мысль, что каждый сегмент можно толковать как элемент некоего базового словаря форм. По двумерным паттернам (в частности, таким как пальмеровы схематические лица) овалы могут выступить в роли первого примитива, а кривые линии — второго. По трехмерным объектам (например, таким, как реальные лица людей) примитивами явятся трехмерные фигуры — эллипсоиды, цилиндры и конусы.

Существует несколько теоретических подходов к тому, как следует определять множество волюметрических формных примитивов. В разделе 2.4.4 мы рассказали о том, как Марр (1982) на основе бинфордской идеи — т.е. через комбинации типизированных цилиндров — представлял сложные формы (см. рис. 2.4.8). Как следует из самого термина, типизированные цилиндры — это некие обобщенные, стандартизированные цилиндрические фигуры. Стандартный цилиндр можно построить, взяв в качестве основы круг определенного диаметра и протянув его по вертикальной оси на определенное расстояние. Переменными параметрами в данном описании явятся лишь диаметр основания и высота (длина). Типизированный же цилиндр позволяет некоторым *фиксированным* параметрам стандартного цилиндра стать *переменными*:

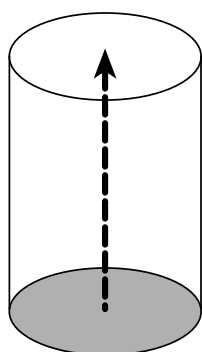
1. *Переменная основания*. Двухмерное основание фигуры (именуется *поперечным сечением* или *эффективным сечением*), протянутое по оси, вовсе не обязано быть кругом, но может явиться эллипсом, прямоугольником или любой иной завершенной двухмерной фигурой (см. рис. 8.2.18А). Кубики, например, — это типизированные цилиндры, чье эффективное сечение — квадрат, а треугольные призмы — это также типизированные цилиндры, чье сечение — треугольник.

2. *Переменная оси*. Аналогично ось, вдоль которой протянуто основание, не обязана быть прямой, но может стать искривленной как угодно. U-образная фигура на рис. 8.2.18 В, к примеру, — это типизированный цилиндр с круговым сечением, чья ось изогнута в «U».

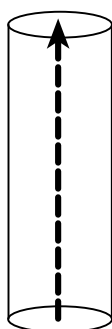
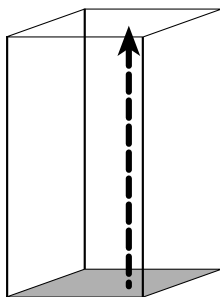
3. *Переменная протягивания*. Сама форма не обязана оставаться полностью фиксированной, но может меняться (рис. 8.2.18 С). Конус, к примеру, — это типизированный цилиндр, в котором диаметр кругового сечения линейно падает по мере протягивания вверх.

В приведенных примерах за раз меняется только один параметр (для простоты), но все три могут отличаться от стандартных значений в классическом цилиндре, давая тем примеры цилиндров типизированных.

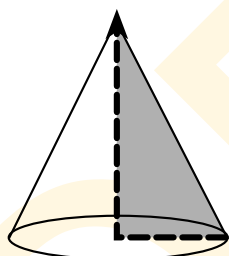
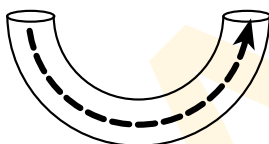
Существует множество иных вариантов трехмерных примитивов, из которых в принципе можно построить структурные описания. Некоторые из них весьма тесно связаны с типизированными цилиндрами (в частности, такими как бедермановы «геоны», которые мы обсудим в гл. 9), а некоторые скорее отличны. Как один из примеров — рис. 8.2.19, на котором даны два объекта более «органичных форм», чем те, что



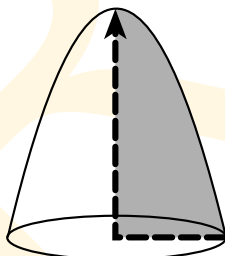
А. Переменная основания



В. Переменная оси



С. Переменная протягивания



**Рис. 8.2.18** Типизированные цилиндры как трехмерные примитивы формы. Типизированные цилиндры получены протягиванием основания (А) вдоль определенной оси (В), при этом по мере этого протягивания радиус основания может меняться (С).

Perez, 1987). Во-вторых, приходится искать более или менее работоспособный набор примитивов и отношений между ними (взять, к примеру, тонкости многих формозависимых перцепций, таких как различение знакомых лиц). В-третьих, формные примитивы и отношения между ними должны быть вычленимы из реальных изображе-

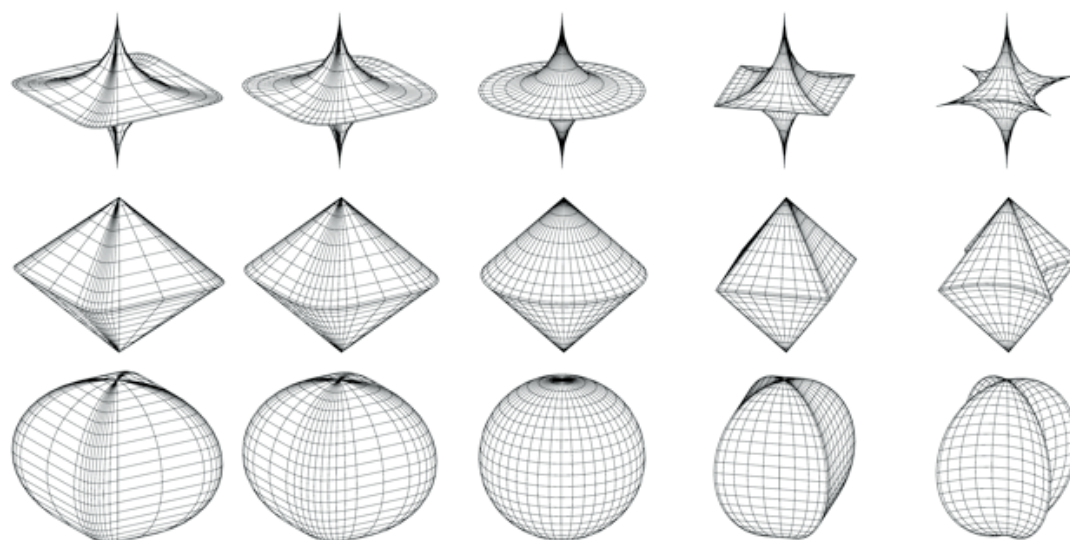
были синтезированы из примитивов, названных в свое время *суперквадриками* (Pentland, 1986). Несмотря на упорную многолетнюю работу с различными наборами волюметрических формных примитивов, нам до сих пор остается неясным, являются ли эти примитивы (независимо от их характера) наилучшим способом построения представлений формы. По сей день не найдено множества трехмерных примитивов адекватных представлению формы всех сложных объектов, но которые, тем не менее, с легкостью распознает зрительная система человека.

#### 8.2.4.2 Плюсы

Структурные описания способны решить вопросы, неразрешимые посредством шаблонов, спектров Фурье и признаков. Структурные описания инвариантны в разных сенсорных каналах, в частности, в каналах восприятия световых краев, текстур и движения. Структурные описания могут учесть эффекты пространственных трансформаций, действуя в рамках объектно-центрированных реперов. С долевой структурой объектов структурные описания работают эксплицитно, оперируя квазинезависимыми представлениями частей и пространственных отношений между долями. Наконец, структурные описания способны представить трехмерную форму через волюметрические примитивы и трехмерные отношения в представляемых объектах.

#### 8.2.4.3 Минусы

Достоинства формных примитивов, увы, бесплатны, и платить приходится тем, что представления оказываются довольно сложными (Grimson & Lozano-



**Рис. 8.2.19** Два примера сложных форм, построенных из суперквадриковых примитивов. Пентланд предложил суперквадрики как альтернативу трехмерным примитивам на базе типизированных цилиндров. Отрисованные с соответствующим сглаживанием, яркостными градиентами и тенями, суперквадрики способны дать удивительно реалистичные изображения естественных форм.

ний: одно дело, предположить, что все объекты можно представить данным множеством примитивов, другое — описать алгоритмы, способные к детекции этих объектов. Однако даже при таком раскладе остается надежда на то, что структурные описания — это базис восприятия формы.

\* \* \*

Итак, последняя концепция представления формы, которую нам надлежит разобрать, оперирует своеобразным типом структурного описания, основанным не на формных примитивах, и не на правиле границ: описания такого рода не формализуемы в логике сетей из узелков и связей между ними, и это будет подход, определяющий доли объектов контекстно-сенситивным, «оптимальным» образом, о необходимости чего толковали гештальтисты еще много лет назад. Однако для понимания сей концепции мы прежде сделаем небольшое отступление, в котором объясним два базисных принципа в гештальт-толковании восприятия формы: *фигуральное совершенство* и *прягнаниц*.

### 8.3 ФИГУРАЛЬНОЕ СОВЕРШЕНСТВО И ПРЯГНАНЦ

Гештальт-психология выделяет некий аспект зрительного ощущения, возникающего от объекта, называя этот аспект *гутегештальтом* («gute Gestalt») или *совершенной формой*. Однако в современной литературе по восприятию чаще употребляют термин *фигуральное совершенство*.

⇒ Фигуральное совершенство — это свойство перцепта, которое точнее всего трактуется как совокупность простоты, последовательности и системности объекта.

Точные формы, наподобие круга на рис. 8.3.1 А — это простые, последовательные и регулярные формы. Любое изменение, вносимое в круг, снижает ощущение «совершенства». В противоположность сему кучка линий на рис. 8.3.1 В выглядит сложно, нерегулярно и бессистемно.

Перед тем как дать детальный разбор понятия фигурального совершенства, давайте обсудим то, почему оно может стать значимо в восприятии формы, особенно феноменологический компонент этого понятия. Одна из причин в том, что «совершенные» фигуры зрительная система может обработать с большей эффективностью нежели «несовершенные».

Психолог Йельского университета Вендель Гарнер (1974) горячо отстаивал сию точку зрения в своих работах по зрительному восприятию простых зрительных паттернов и памяти на них. Ученый поставил множество изящных экспериментов, продемонстрировавших то, что решение человеком зрительных задач плотно соотносимо с тем, как мы оцениваем субъективное «совершенство» фигур. К примеру, Гарнер обнаружил, что мы:

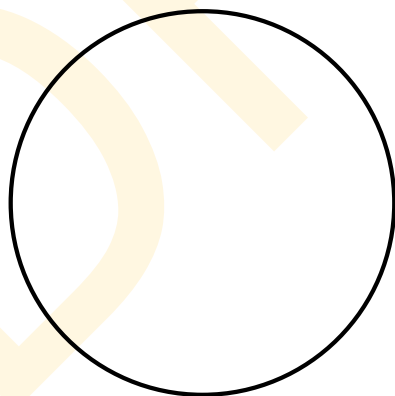
1. Сравниваем пару «совершенных» фигур на предмет физической идентичности гораздо быстрее, чем «несовершенных».

2. Помним «совершенные» фигуры гораздо лучше и четче «несовершенных».

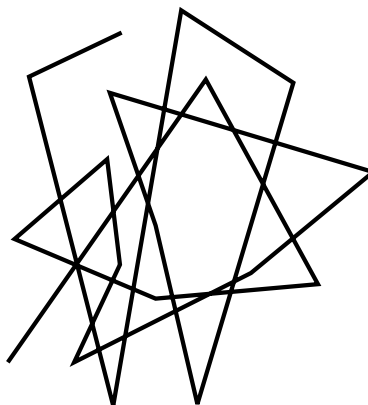
3. Описываем «совершенные» фигуры меньшим числом слов.

4. Изучаем «совершенные» фигуры гораздо быстрее «несовершенных».

Чем легче и экономичнее представляются и обрабатываются «совершенные» пат-

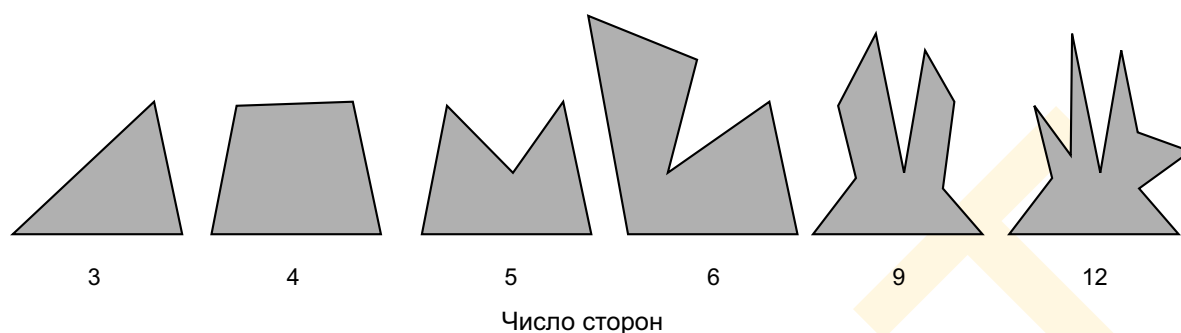


А



В

**Рис. 8.3.1** Фигуральное совершенство. Фигуру А мы воспринимаем как более простую, более регулярную и более упорядоченную, чем фигуру В.



**Рис. 8.3.2** Влияние сложности на фигуральное совершенство. При прочих равных фигуры с меньшим числом сторон склонны восприниматься «совершеннее» многосторонних.

терны, тем перечисленные факторы лучше стыкуются друг с другом. Очевидно, зрительная система специфически чувствительна к структуре «совершенных» паттернов и пользуется ими при обработке информации о форме. Вскорости мы познакомимся с некоторыми видами структуры, дающей паттерны очевидно «совершеннее» других.

Стоит сказать, что эффективности нейронной обработки гештальт психология уделяет большее внимание, нежели фигуральному совершенству. Как было показано в разделе 2.1.2, гештальтисты полагали, что обработка определяет путь, по которому идет всё наше зрительное восприятие. Напомним, что сия точка зрения коротко выражается принципом прыгнанц:

⇒ Восприятие настолько «совершенно», насколько позволяют превалирующие условия.

Фактически сие значит, что у любой фигуры может быть несколько перцептов и что фигуральное совершенство определяет, каким будет этот перцепт. Сие стало теоретическим прорывом того времени. При этом, к сожалению, гештальт-психология никогда не формулировала эксплицитной теории восприятия формы, основанной данном выше определении. Позднее голландский психолог Эммануэль Левенберг (Buffart & Leeuwenberg, 1981; Van der Helm & Leeuwenberg, 1991) попробовал заполнить этот пробел.

Итак, в оставшейся части главы мы рассмотрим два вопроса, связанных с концепцией фигурального совершенства: во-первых, поговорим о том, какие факторы определяют меру «совершенства» перцептивной формы данного объекта; во-вторых о том, как фигуральное совершенство соотносится с нашим зрительным восприятием формы. Последнее будет строиться на структурных описаниях формы силами т.н. *левенберговой теории структурной информации*.

### 8.3.1 Теории фигурального совершенства

Один из самоочевидных факторов, влияющих на фигуральное совершенство объекта — это количество частей, составляющих этот объект. На рис. 8.3.2 показано, что по мере роста числа сторон усложняется перцептивная форма фигуры. Вспомним, что в гл. 6 мы коротко коснулись того, что количество частей объекта — это далеко не всё, что следует учесть нашему восприятию, поскольку объекты могут различаться по критерию фигурального совершенства, даже если у них одинаковое число физиче-



ских частей. К примеру, фигурки на рис. 8.3.3: наблюдатели обычно располагают их по порядку от простых и регулярных слева к сложным и нерегулярным справа. Сей порядок не может зависеть только от числа физических частей в фигуре, поскольку у всех них только по четыре стороны. Визуально фигурки разнятся тем, как организованы и взаиморасположены образующие их линии. Сие и есть влияние общей структуры на восприятие формы.

Гештальт-психология полагает фигуральное совершенство атрибутом перцептивного опыта, идентифицируя его как сходное с субъективной простотой/сложностью перцепта. «Совершенство» полагают неким качеством зрительного ощущения, вызванного цельной фигурой. Гештальт-психология утверждает при этом, что сие качество в принципе нередуцируемо до кусочковых свойств (т.е. до частей и связей между ними), и что оно относимо к неким глобальным качествам, наподобие билатеральной симметрии. Однако нет уверенности в том, что самой симметрии достаточно для учета всего, что касается фигурального совершенства, но очевидна ее холистическая окраска. При этом нельзя не сказать, что кое-какой прогресс в нашем понимании критериев фигурального совершенства наметился в результате работ Гарнера (1974) и Пальмера (1991), о чем мы коротко поговорим ниже.

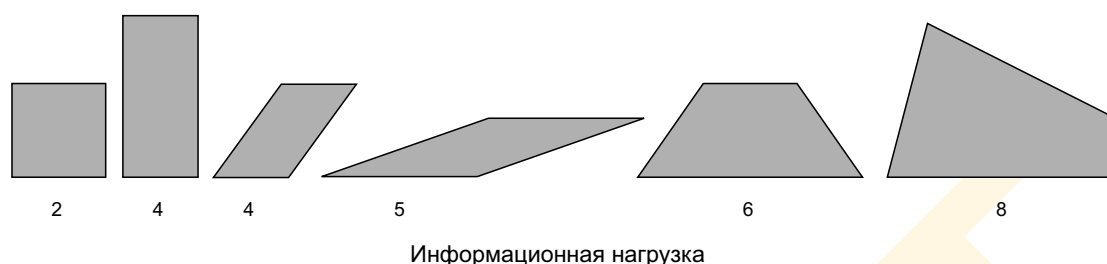
### 8.3.1.1 Классическая теория информации

Первым существенным шагом в развитии концепции фигурального совершенства было признание теории информации Клода Шеннона. Теория информации — это математическая теория обмена данными. Шеннон окрестил эти данные «информацией», хотя их отношение к повседневному и общему пониманию информации не столь очевидно, чем могло бы быть. Атомарная единица информации — т.н. *бит* (Shannon, 1948; Shannon & Weaver, 1949).<sup>1</sup>

Связь между теорией информации и гештальт-концептом фигурального совершенства была практически одновременно сформулирована американцами Фредом Эттнивом, Джулианом Хочбергом и Роджером Мак-Алистером (1953-54). Ученые полагали, что ежели перцептивная система кодирует фигуру оптимально, т.е. удаляя из нее все избыточности (например, симметрии и повторы) «совершенные» фигуры можно закодировать и сохранить с большей эффективностью, чем «несовершенные». Термин «эффективность» они относили к кодированию формы «совершенной» фигуры всего лишь несколькими «битами» информации, тогда как для описания «несовершенных» понадобится множество бит.

Информационный анализ фигурального совершенства основывается на том, что фигура дробится на составляющие, такие как углы и линии, которые затем анализируются на предмет структурных системностей, в частности на равенство углов и длин отрезков внутри одной и той же фигуры. И у квадрата, и у неправильного четырехугольника (рис. 8.3.3) по четыре стороны и по четыре угла, но квадрат «совершеннее», поскольку все его стороны одной длины, а все его углы — по 90°. Очевидно, что такие системности для описания их требуют меньшего числа бит (хотя теоретикам информации так и не было до конца ясно, на что же должны походить эти описания или сколько бит информации в данной фигуре). Однако же информационный анализ оказался

<sup>1</sup> «Бит» — это единица информации равная информации о наступлении одного из двух равновероятных исходов некоторого события (например, о падении монеты одной или другой стороной вверх).



**Рис. 8.3.3** Демонстрация влияния иррегулярности на фигуральное совершенство. Несмотря на то, что у всех шести фигур по четыре стороны, мы отмечаем, что сложность фигур растет слева направо. Цифрами обозначена величина информационной нагрузки согласно теории структурной информации (см. раздел 8.3.2).

большим шагом вперед, нежели вольные тезисы гештальтизма, поскольку показал, что «совершенные» фигуры *объективно* проще «несовершенных».









































### 8.3.1.2 Подмножества поворота и отражения

Несмотря на то, что исследователям удалось-таки найти объяснение гештальт-конструкции фигурального совершенства в объективных показателях, объяснение это слабо стыкуемо с гештальт-позицией оппонентной кусочковой теории восприятия (Wertheimer, 1924/1950). Формулировку, в большей степени подходящую холистическому подходу гештальтистов, дал Гарнер (1974), предположивший, что фигуры «совершенны» до тех пор, пока остаются трансформированными версиями самих себя. Он сформулировал это положение в своей концепции подмножеств поворота и отражения — R&R.<sup>1</sup>

Множество пространственных трансформаций фигуры дает множество *трансформационных вариантов* (далее — *трансвариантов*) этой фигуры. По Гарнеру «совершенные» фигуры дают меньшее число трансвариантов, чем «несовершенные». В оригинальной формулировке Гарнера и Клемента (1963) в качестве примера используются паттерны, состоящие из пяти точек внутри квадратной матрицы 3×3. К этим паттернам применяются восемь трансформаций: четыре центральных поворота (0°, 90°, 180° и 270°) и четыре центральных отражения (относительно вертикальной, горизонтальной, леводиAGONальной и праводиAGONальной осей). После чего определяют *множество поворотов и отражений* (некоторые примеры которых даны на рис. 8.3.4). Мы видим, что внутри этого множества имеется подмножество явно отличных фигур, названное как *R&R-подмножеством*, и пронумерованное на рис. 8.3.4 арабской цифрой под каждым трансвариантом. Гарнер и Клемент (1963) обнаружили, что паттерны с рейтингом «совершенные» (такие как фигуры верхнего ряда на рис. 8.3.4) имеют всего лишь несколько трансвариантов, тогда как у фигур с рейтингом «несовершенные» (нижний ряд) — трансвариантов множество. Разумеется, исследователи предположили, что совершенство фигуры — это функция, обратная размеру R&R-множества.

Отметим, что анализ Гарнера — Клемента целиком опирается на гештальт-позиции, поскольку оперирует цельными фигурами, и не бывает ситуаций, когда применение его потребовало бы дробления фигур на составляющие. Фактически гештальт-прототипом совершенства являлись те паттерны, что остаются сами собой после отра-

<sup>1</sup> Rotation & Reflection — *Прим. пер.*

Повороты				Отражения				Размер подмн-ва R&R	Подгруппы симметрии
0	90	180	270	В	Г	П	Л		
 1	 1	 1	 1	 1	 1	 1	 1	1	(0, 90, 180, 270, V, H, R, L)
 1	 2	 3	 4	 1	 3	 2	 4	4	(0, V)
 1	 2	 3	 4	 4	 2	 1	 3	4	(0, R)
 1	 2	 1	 2	 3	 3	 4	 4	4	(0, 180)
 1	 2	 3	 4	 5	 6	 7	 8	8	(0)

**Рис. 8.3.4** Стимулы, применявшиеся в опытах по изучению фигурального совершенства. По идее Гарнера — Клемента (1963) фигуральное совершенство каждого из пяти паттернов в левой колонке можно вывести из размера их R&R-подмножества, а также — уже по идее Пальмера — из трансформаций в их подгруппы симметрии (1983).  
Palmer, 1983.

жения относительно линии, проходящей через их центры, т.е. паттерны, обладающие билатеральной симметрией. Но даже несмотря на это Гарнер не стал формулировать свой концепт в логике симметрии, поскольку обнаружил, что поворотная инвариантность тоже весьма значима и не стыкуется с идеей симметрии как минимум в своей повседневной и общеупотребительной форме. Однако же поворотная инвариантность четко согласуется с современным математическим определением симметрии, что стало краеугольным камнем пальмерова пересмотра идеи фигурального совершенства в логике подгрупп симметрии, о которых и пойдет теперь разговор.

### 8.3.1.3 Подгруппы симметрии

Гарнерова идея, основанная на концепции R&R-подмножеств, способна учесть большое число отличий, возникающих из-за разницы в уровнях совершенства. Однако отметим, что идея эта сводит вместе разные виды качественно отличных структур. Рассмотрим, например, три центральных ряда в рис. 8.3.4: у всех строго по четыре разных фигуры в их R&R-подмножествах, но первая билатерально симметрична по вертикальной оси, вторая — по леводиагональной, а третья — несимметрична. Большинство

из нас находят вертикально-симметричную фигуру «совершеннее» двух других, и множество опытов подтверждают сей факт (например, Chipman, 1977; Palmer, 1991; Palmer & Nemenway, 1978; Royer, 1981). Эти отличия невозможно объяснить в логике R&R-подмножеств, поскольку у таких фигур одно и то же число трансвариантов.

Однако возможное решение следует из анализа, который мы только что дали в логике симметричности структуры. Каждую фигуру можно охарактеризовать через ее вид симметрии, различая билатеральную симметрию относительно вертикальной, горизонтальной и диагональных осей. Но что происходит с паттерном в четвертом ряду рис. 8.3.4, в котором нет билатеральной симметрии? Очевидно, такие фигуры все же обладают симметрией, но иного свойства.

Современная математика толкует данную симметрию в логике т.н. *трансформирующей инвариантности* (Weyl, 1952): фигура будет симметрична при данной трансформации, если приложение этой трансформации оставит фигуру неизменной. В основе стандартного примера билатеральной (или зеркальной) симметрии лежит трансформация, представляющая собой отражение относительно оси симметрии. Так, например, вертикально симметричный паттерн — скажем, прямостоящие литеры «А» и «Т» — останутся неизменными после зеркального отражения относительно вертикальной оси, проходящей через их центр. Аналогично горизонтально симметричный паттерн — скажем, литеры «С» и «В» — остаются неизменными после отражения относительно горизонтальной оси, проходящей через их центр.

В рамках сей общей идеологии поворотная симметрия — это точно такая же форма билатеральной симметрии, но лишь с тем отличием, что в преобразовании участвует центральный поворот на определенный угол. Скажем, паттерны с поворотной симметрией на  $180^\circ$  (например, «S», «N» и «Z») остаются неизменными после поворота на  $180^\circ$  относительно их центров (четвертый ряд рис. 8.3.4). Фигуры с поворотной симметрией на  $90^\circ$ , например, свастика, остаются неизменными после поворота на  $90^\circ$  относительно центра.<sup>1</sup>

Следовательно, *подгруппу симметрии* данной фигуры можно охарактеризовать как подмножество пространственных трансформаций, оставляющих эту фигуру инвариантной.<sup>2</sup> Самая правая колонка рис. 8.3.4 дает нам подгруппы симметрии по гарнеровым пятиточечным паттернам. Отметим, что «совершенные» фигуры обладают бóльшим числом симметрий (т.е. бóльшим числом возможных трансформаций в их подгруппах симметрии), чем «несовершенные». По этой причине Пальмер и предположил, что совершенство фигуры можно идентифицировать по ее подгруппе симметрии.

Если принять во внимание плотную связь между гарнеровыми R&R-подмножествами и пальмеровыми подгруппами симметрии, неудивительным станет то, что и у первых, и у вторых количественные структуры тоже близки. Фактически число трансфор-

<sup>1</sup> Отметим, что фигуры с двумя симметриями отражения (скажем, литеры «Н», «О» и «Х») всегда обладают симметрией поворотной.

<sup>2</sup> Сие обнаруживает тот факт, что обсуждаемые подмножества обладают некоторыми специальными структурными свойствами, из-за чего не все логически возможные подмножества трансформаций могут явиться подмножествами симметрии. К примеру, фигура не может обладать одновременно и вертикальной, и горизонтальной симметрией (скажем, «Н» и «Х»), не обладая при этом поворотной симметрией на  $180^\circ$ . Аналогично фигура не может иметь поворотной симметрии на  $90^\circ$  (скажем, свастика) без поворотной симметрии на  $180^\circ$  и  $270^\circ$ . Сия дополнительная структура подмножества симметрии характеризует их как примеры математических *групп* и *подгрупп*, а не множеств и подмножеств. Подробнее см. в литературе по абстрактной алгебре.

маций в подгруппе симметрии — это величина обратная размеру R&R-подмножества по отношению к полному множеству трансформаций:

$$R \times S = T,$$

где  $R$  — это размер R&R-подмножества,  $S$  — размер подгруппы симметрии и  $T$  — общий размер всех групп трансформации. Сие соотношение действительно имеет смысл, поскольку оба концепта отличны друг от друга лишь теоретической обратностью фигура/фон: Гарнер сосредоточился на *паттернах*, образованных группой трансформаций, тогда как Пальмер — на *трансформациях*, при которых паттерны инвариантны.

Однако же, несмотря на столь тесную связь у анализа, построенного в логике подгрупп симметрии, имеется серьезное преимущество. Дело в том, что элементы подгрупп симметрии появились из фиксированного множества трансформаций, одних и тех же по всем возможным фигурам. Следовательно, подгруппы симметрии можно напрямую сравнивать на предмет идентичности трансформаций в подгруппе при разных фигурах. Сие невозможно с R&R-подмножествами, поскольку элементы подмножеств — это фигуры, которые нельзя сравнивать, коль скоро фигуры разные. Сравнимо только *число* фигур в R&R-подмножестве. Поэтому у трех средних паттернов на рис. 8.3.4 один размер подмножества R&R (и одно и то же число симметрий), но разные трансформации в их подгруппах симметрии.

### 8.3.2 Теория структурной информации

Итак, мы подготовлены теперь к тому, чтобы, наконец обсудить теорию представления формы, плотнее всего соотносимую с гештальт-концепцией фигурального совершенства и принципом прягнанц. Сей концепт, исходно названный *теорией кодирования*, а позднее *теорией структурной информации*, был в 1971 г. сформулирован голландским психологом Эммануэлем Левенбергом (1971), а спустя время усовершенствован им в сотрудничестве с Гансом Буффартом и другими исследователями (например, Buffart & Leeuwenberg, 1981; Van der Helm & Leeuwenberg, 1991). Теория дает метод построения различных описаний формы одного и того же объекта и соотносит их с восприятием через гештальт-принцип прягнанц. В этом смысле она совсем непохожа на прочие теории представления формы, рассмотренные нами ранее.

Описания формы, полученные силами теории структурной информации через создание и последующее упрощение перцептивных описаний, именуются *кодами*. Фигуральное совершенство и прягнанц встраиваются в систему через такой показатель как *информационная нагрузка*.

⇒ Информационная нагрузка — это мера сложности максимально подходящего кода представления формы.

То есть, информационная нагрузка — это показатель, более или менее согласуемый с гештальт-принципом прягнанц. Левенберг показал, что перцептивным выбором большинства наблюдателей является то, что имеет простейший код представления формы, т.е. наименьшую информационную нагрузку. Отметим, что теория хорошо работает не только в отношении оценки фигурального совершенства, но и вполне удовлетворительно объясняет многие из интересных организационных эффектов, обсуждавшихся в гл. 6, в частности таких как частично закрытые объекты и объекты полупрозрачные.

Общее описание концепта структурной информации таково:



1. Построение *примитивного кода* — символьного описания контура фигуры как последовательности отрезков той или иной длины и углов между ними.

2. Применение множества *семантических операторов* (или *правил подстановки*) для упрощения примитивного кода путем удаления всех возможных избыточностей. В результате упрощенный код получает название *кода редуцированного*.

3. Вычисление *информационной нагрузки* каждого редуцированного кода через подсчет числа параметров (числовых величин) в них. Сие значение приблизительно соответствует фигуральному совершенству перцепта фигуры, описанной этим (редуцированным) кодом.

4. Редуцированный код с минимальной информационной нагрузкой именуется *минимальным кодом*, и описывает то, что теория структурной информации прогнозирует как наиболее вероятный путь перцепции. Прочее может восприниматься с вероятностью, зависящей от величины информационной нагрузки.

Заметим, что хотя описанные процедуры больше смахивают на некую машинную программу, теория структурной информации не нацелена на моделирование фактических процессов в зрительной системе: в логике Марра — это количественная, но не алгоритмическая концепция, поскольку она просто описывает отношения между входными паттернами (и возможное представление формы), но не процесс, в результате которого получается это представление.

### 8.3.2.1 Примитивные коды

Чтобы понять, как теория структурной информации создает описания формы, начнем с примера кодирования квадрата. Вначале фигура конвертируется в символьную последовательность: выбирается стартовая позиция и направление по контуру (рис. 8.3.5) в виде стрелки с точкой. Символьная последовательность примитивного кода представляет собой последовательность чисел, представляющих, во-первых, длину линии, которую нужно вычертить в текущем направлении и, во-вторых, угловое направление, которое следует принять после того, как линия отрисована.<sup>1</sup> Таким образом, примитивный код квадрата с длиной стороны  $L$ , который начитается с левого верхнего угла, будет иметь следующую символьную последовательность:

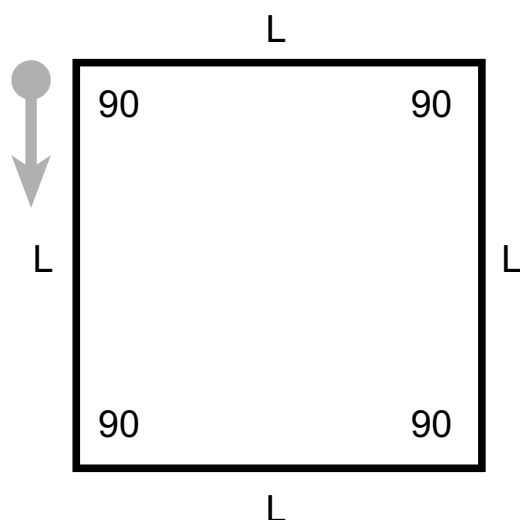
$L\ 90\ L\ 90\ L\ 90\ L\ 90.$

Сей код «вычерчивает» квадрат, заканчиваясь исходными позицией и направлением. Отметим, что данная запись — это частный случай общего описания любого произвольного четырехугольника:

$W\ a\ X\ и\ Y\ c\ Z\ d,$

где  $W, X, Y, Z$  представляют длины четырех сторон, а  $a, b, c$  и  $d$  — величины четырех углов. Заменяя эти переменные соответствующими значениями, можно описать любой четырехугольник в т.ч. с рис. 8.3.3.

<sup>1</sup> Отличие между этими двумя параметрами мы подчеркнем тем, что длину символов станем отмечать курсивом, тогда как углы оставим в прямом начертании.



Примитивный код:  $\perp 90 \perp 90 \perp 90 \perp 90$

Редуцированный код:  $4 \times \perp 90$

**Рис. 8.3.5** Кодирование квадрата согласно теории структурной информации. Начав с левого верхнего вертекса и двигаясь вниз, примитивный код генерится через описание длины пути движения в направлении угла, где мы поворачиваем перед началом каждого следующего сегмента. Удалив избыточности из кода, получаем код редуцированный.

В теории структурной информации существуют несколько *семантических операторов* (правил перевывода), позволяющих понизить сложность символьных строк. Три важнейших из них:

1. *Оператор повторов* — исключает повторяющиеся последовательности и заменяет их на  $n$ -параметр, т.е. число итераций:

$$S \rightarrow n \times (X).$$

Пример работы оператора:

$$a b c a b c a b c \rightarrow 3 \times (a b c).$$

2. *Оператор симметрии* — исключает из последовательностей зеркальную симметрию и заменяет ее на символ SYM:

$$S \rightarrow SYM (X).$$

Пример работы оператора:

$$a b c c b a \rightarrow SYM (a b c).$$

### 8.3.2.2. Устранение избыточностей

Чем проще и регулярнее четырехугольник, тем проще его описание, благодаря устранению из примитивного кода разного рода избыточностей. К примеру, код квадрата несет в себе четыре повтора ( $L 90$ ): сию паразитную регулярность можно исключить из последовательности, применив т.н. *оператор повторов* — *правило перевывода*, сокращающее примитивный код до кратчайшей, простейшей строки символов. Оператор повторов заменяет символьную строку  $S$ , состоящую из  $n$ -повторов в последовательности  $X$  на символьную строку  $n \times X$ . То есть, правило перевывода можно записать так:

$$S \rightarrow n \times (X),$$

где стрелка означает «перевыведено как». При описании квадрата это выглядит следующим образом:

$$L 90 L 90 L 90 L 90 \rightarrow 4 \times (L 90).$$

Сей код соответствует перцептивной поворотной симметрии квадрата (на  $90^\circ$ ), но не билатеральной симметрии.

3. *Оператор распределения* — исключает перемежение между подстроками и заменяет его на пару подстрок в угловых скобках:

$$S \rightarrow \langle X \rangle \langle Y \rangle.$$

Пример работы оператора:

$$a b a c a d \rightarrow \langle a \rangle \langle b c d \rangle.$$

Отметим, что фактические правила много сложнее, и поэтому на практике используются совсем другие операторы, но при этом показанные составляют их базовое множество (Van der Helm & Leeuwenberg, 1991), на примере которого мы можем понять, как, согласно теории структурной информации, следует избавляться от разного рода избыточностей в примитивных кодах.

### 8.3.2.3 Информационная нагрузка

Итак, мы уже догадываемся, что простота или «совершенство» редуцированного кода соответствует информационной компактности его описания. Формально, совершенство можно определить в логике структурной информации как информационную нагрузку на данную строку кода —  $I(S)$ , то есть число параметров в описании. Когда мы говорим о примитивном коде квадрата, информационная нагрузка равна 8 — одна единица на каждый отрезок и одна — на каждый угол. Нагрузка же редуцированного кода того же квадрата равна 3 — одна единица на число итераций (4), одна на длину ( $L$ ) и одна на угол (90). То есть, мы видим, что редуцированный код много проще исходного примитивного и, в гештальт-логике, соответствует «совершенному» восприятию, поскольку отражает перцепт регулярности в форме поворотной симметрии. Но является ли этот код простейшим? Вопрос далеко не праздный, поскольку, согласно гештальт-принципу прягнанц, теория структурной информации предполагает, что наиболее вероятный перцепт будет обладать наименьшей информационной нагрузкой.<sup>1</sup>

Существует множество других кодов квадрата, соответствующих разным аспектам видения его структуры. Как мы уже говорили ранее, код  $4 \times (L \ 90)$  эксплицитно кодирует поворотную симметрию квадрата на  $90^\circ$ , но при этом не касается его наиболее выраженного свойства — тетрасимметрии на отражение. Последняя кодируется иначе: начнем с центра одной из сторон и закодируем стороны как два соседних сегмента с длиной  $L/2$  и углами как два соседних поворота на  $45^\circ$ . Запись будет выглядеть так:

$$L/2 \ 45 \ 45 \ L/2 \ L/2 \ 45 \ 45 \ L/2 \ L/2 \ 45 \ 45 \ L/2 \ L/2 \ 45 \ 45 \ L/2,$$

и, хотя сей примитивный код фактически длиннее, чем тот, что был в начале разговора, с помощью трех операторов симметрии его можно перекодировать даже в еще более простой редуцированный код:

<sup>1</sup> На самом деле, теория структурной информации чуть сложнее и исходит из того, что вероятность перцепта данной интерпретации — это функция отношения информационной нагрузки этой интерпретации к нагрузке интерпретации альтернативной. Мы сознательно ушли от обсуждения этой тонкости, хотя она и важна в количественном прогнозировании эмпирических данных.

(SYM (SYM (SYM (L/2 45))))).

Сие описание дает информационную нагрузку всего лишь 2: одна единица — параметр длины ( $L/2$ ), а вторая — угловой параметр ( $45^\circ$ ).<sup>1</sup> Редуцированный код фактически описывает зеркальные симметрии квадрата его минимальным кодом, что четко согласуется с перцептом билатеральной симметрии, возникающим у большинства из нас.

Теория структурной информации применима к разным четырехугольникам с рис. 8.3.3 и позволяет прогнозировать их фигуральное совершенство. Информационная нагрузка минимального кода по каждой фигуре дана под ней. Мы видим, что у квадрата эта нагрузка минимальна (2), затем идут четырехугольник и ромб (4), затем параллелограмм (5), трапеция (6) и, наконец, неправильный четырехугольник (8). То есть, порядок роста информационной нагрузки четко согласуется с порядком восприятия совершенства этих фигур.

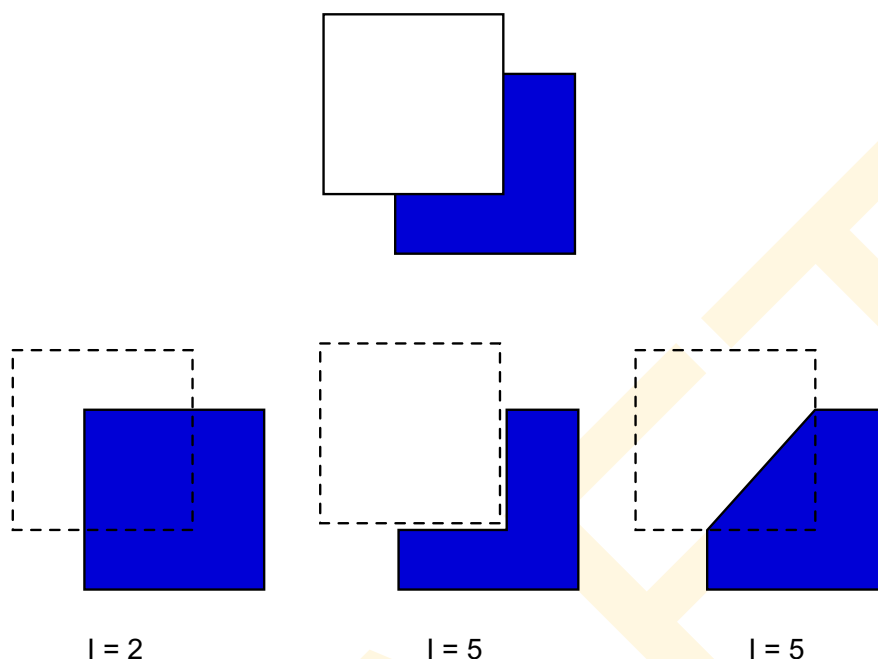
### 8.3.2.4 Приложения к перцептивной организации

Теория структурной информации пригодна не только для оценки перцептивного фигурального совершенства форм, но также и для объяснения перцептивной организации во многих потенциально неопределенных ситуациях, в частности, таких как завершение частично закрытых объектов и в ситуации полупрозрачности (например, Buffart, Leeuwenberg & Restle, 1981). В качестве примера разберем случай квадрата, частично закрытого другим квадратом (рис. 8.3.6). Вначале мы рассмотрим три возможных перцептивных интерпретации стимуляции: квадрат позади другого квадрата; мозаичную интерпретацию (квадрат, упирающийся в L-образную фигуру); квадрат с обрезанным углом (рис. 8.3.6). Итак, ясно, что код белого квадрата слева вверху — один и тот же во всех трех случаях, поэтому мы можем проигнорировать его, сосредоточившись лишь на кодах альтернативного завершения черной фигуры в правой нижней части рисунка.

Как показано на рис. 8.3.6, минимальный код у квадрата будет иметь меньшую информационную нагрузку, чем L-образная фигура или фигура с обрезанным краем. То есть, мы видим, что теория структурной информации способна спрогнозировать тот факт, что мы скорее воспримем вариант «квадрат позади квадрата», нежели какую-то из двух других возможных альтернатив. Так происходит потому, что квадрат — это «совершенная» фигура (т.е., обладающая наименьшей информационной нагрузкой) в сравнении с мозаичной, равно как в сравнении и с любыми другими вариантами завершения. Любопытно и то, что теория структурной информации способна прогнозировать нашу склонность к перцептивному завершению восьмигранников (рис. 6.4.7 А), но не крестов (рис. 6.4.7 В): причина в том, что у примитивного кода симметричного креста информационная нагрузка будет меньшей, чем у его заверщенного варианта.

Теория структурной информации плотно соотносится с тремя теориями фигурального совершенства, описанными выше. Подобно информационному теоретическому анализу Эттнива — Хочберга — Мак-Алистера (1953) она посредством эффективных описаний идентифицирует «совершенство» фигур, из которых удалены все избыточности. И подобно концептам Гарнера (R&R-подмножества) и Пальмера (подгруппы симметрий), избыточности, удаляемые из кода, связаны с поворотной и зеркальной симметриями. Однако теория структурной информации выдается далеко за пределы этих

<sup>1</sup> Сам по себе оператор симметрии (SYM) не имеет информационной нагрузки, поскольку не существует параметра, за ним закрепленного. При этом оператор повторов требует наличия дополнительного параметра — числа необходимых итераций.



**Рис. 8.3.6** Толкование зрительной завершенности теорией структурной информации. Композиция наверху обычно воспринимается как синий квадрат позади белого. Теория структурной информации объясняет это тем, что именно при квадрате система обладает наименьшей информационной нагрузкой. Литера «I» указывает на степень информационной нагрузки при разных перцептах.

концептов, потому что описывает *фактическое* представление формы и использует его для прогнозирования реальных перцептов в неопределенных ситуациях.

### 8.3.2.5 Плюсы

Теория структурной информации легко управляется с трансформациями подобия, поскольку минимальный код не зависит от расположения, размера, ориентации и отражения фигуры. Т.е. мы говорим, что эти переменные эффективно факторизуемы (выносимы за скобки) при генерации кода.

Еще один плюс состоит в тесной связи с фигуральным совершенством и с весьма изящным толкованием различных (и весьма важных) перцептивных феноменов. Связь реализуется через анализ кодируемости различных интерпретаций формы.

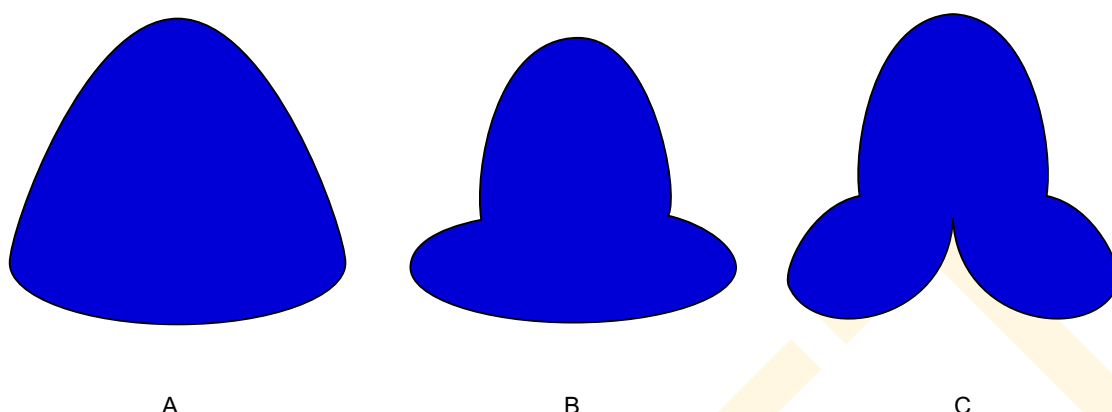
### 8.3.2.6 Минусы

Несмотря на очевидные плюсы, у теории структурной информации имеется ряд серьезных недостатков и основные из них следующие:

1. Теория неприменима непосредственно к полутонным изображениям. Подобно многим психологическим концептам, она работает только с идеализированными упрощенными штриховыми прорисовками. Следовательно, требуется предобработка, конвертирующая полутон в черно-белые контуры.

2. Теория не позиционирует эксплицитных кодов частей объекта, кроме простых линий. К примеру, выпуклости и вогнутости обрабатываются одним и тем же путем





**Рис. 8.3.7** Дольчатость структуры. Обычно мы воспринимаем эти фигуры состоящими из одной (А), двух (В) и трех (С) частей, хотя минимальный код всех трех фигур один и тот же.

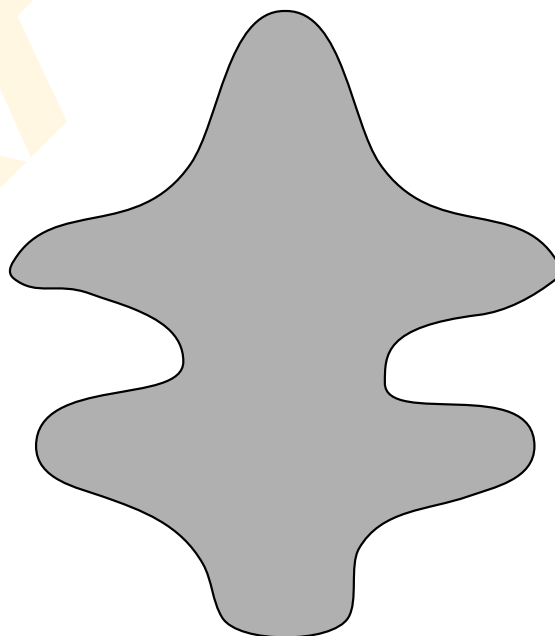
и, несмотря на то, что все структуры представлены минимальными кодами, ни одна из них не относится к частям: на рис. 8.3.7 даны три фигуры с явно разным числом частей, но что никак не отражается на их минимальных кодах.

3. Нетрудно убедиться в том, что теория структурной информации легко применима к двумерным проекциям, но намного тяжелее к трехмерным. Последние можно построить, протягиванием двумерных фигур вдоль оси, почти так же как при создании типизированных цилиндров. Однако такой вид алгоритма дает описание трехмерности без объяснения того, что перцепт получен из двумерной проекции объекта.

4. Единственный путь убедиться, что перцептивная интерпретация несет минимум информационной нагрузки — это сравнить все возможные коды. Сие очевидно затруднительно, а во многих ситуациях просто невозможно. К примеру, в случае частично закрытого квадрата имеет место бесконечное число логически возможных вариантов завершения (которые, соответственно, невозможно ни рассчитать, ни сравнить).

5. Коды крайне чувствительны к шуму. Форма представляется весьма точно, но за счет структуры большего масштаба. К примеру, приблизительная симметрия фигуры на рис. 8.3.8, не берется теорией структурной информации с применением оператора симметрии, поскольку симметрия неточна. Но даже в этом случае мы воспринимаем данный объект как более или менее симметричный относительно вертикальной оси.

6. Как было показано выше, теория структурной информации не описывает процесс. То есть, ничего не говорит нам



**Рис. 8.3.8** Приблизительная симметрия. Мы воспринимаем эту фигуру как более или менее симметричную относительно вертикальной оси, но сей факт весьма трудно истолковать силами теории структурной информации.

о фактических нейронных событиях или работе алгоритмов, посредством которых зрительная система подбирает редуцированный код. Более того, до сих пор неизвестен количественный метод построения минимального кода.

Вместе с тем, теория структурной информации — это, в контексте представления формы, наиболее определенное и наиболее успешное продолжение гештальт-идей фигурального совершенства и прыгнанц. До сих пор непонятно, решаются ли в принципе проблемы, описанные выше, но ежели да — концепция структурной информации может выступить в роли серьезного претендента на теорию представления формы объектов в зрительной системе человека.

\* \* \*

Итак, мы познакомились с основными современными концепциями представления формы объектов в зрительной системе человека, обнаружив при этом, что ни одна из них не идеальна. Возможно, не станет неожиданной та мысль, что чем эффективнее подход — тем он сложнее. Концепция кодирования формы в логике иерархически вложенных структурных описаний, основанная на некоем наборе формных примитивов и объектно-центрированных реперов, весьма сложна. Но при этом всё равно далека от совершенства.

В следующей главе мы вновь столкнемся с некоторыми из описанных здесь видов представления формы, но уже как фрагментов теории зрительной категоризации.



Давайте сделаем паузу и подытожим ранее сказанное о пространственном зрении: сперва мы познакомились с тем, как пространственная информация извлекается нашим зрением из двухмерных ретинальных изображений (гл. 4); затем рассмотрели интерпретацию этих изображений как произошедших от массива поверхностей в трехмерном пространстве (гл. 5); потом рассмотрели то, как отдельные участки ретинального изображения организуются в составляющие объектов, собственно объекты и группы объектов (гл. 6); наконец, мы разобрали то, как объекты анализируются нашим зрением по их внутренним свойствам и фрагментам (гл. 7) с исходом в перцепт формы этих объектов (гл. 8).

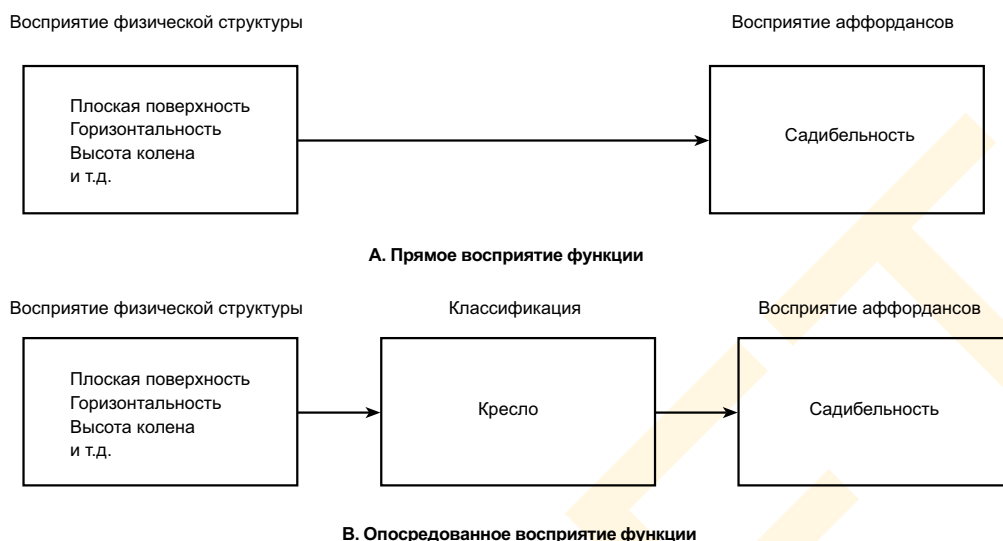
Пройдя весь этот материал, мы очень многое узнали о пространственном зрении, но мы по-прежнему не касались того, как восприятие наше достигает одной из важнейших своих целей, сформулированной нами еще в первой главе — поставке наблюдателю информации об *индивидуальном назначении* объекта сцены. То есть, в нашем разговоре мы еще ни разу не упомянули о таком важном аспекте зрения, как *детекция функции* зримого объекта. Последнее — предмет текущей главы.

## 9.1. ВОСПРИЯТИЕ ФУНКЦИИ

Зрительное восприятие функционального назначения объектов позволяет нам понимать то, каким образом эти объекты можно использовать в своих интересах. Если вдруг, паче чаяния, мы станем воспринимать окружающий мир, не оценивая функцию объектов вокруг нас, мы окажемся в положении того несчастного, которого внезапно закинули на далекую и неизвестную планету, населенную невиданными сущностями и существами. При этом мы вполне успешно воспримем их форму, расположение, ориентацию, окраску и прочие физические свойства, словом, окажемся вполне способными ориентироваться в инопланетном трехмерном пространстве, не спотыкаясь и не натываясь на все подряд. Если бы понадобилось, то мы даже смогли бы построить их трехмерные модели, окажись под рукой подходящий материал (и если бы имели соответствующий навык). Но без дополнительной информации, мы не смогли бы понять, как себя вести с ними: какие из них съедобны, а какие нет?; из каких можно построить хижину?; из каких пошить одежду? Более того, без восприятия дополнительной функциональной информации мы не сможем ответить даже на столь простой вопрос, как «Пригоден данный объект, чтобы опереться на него или нет?».

В конечном счете, все обсуждавшиеся нами ранее процессы, подчинены одной главной цели — обеспечивать эволюционную миссию зрения в контексте восприятия функции объектов. Восприятие функции объектов человеком — это невероятно сложный феномен, организующий культурную и межперсональную структуру современного общества. Мы даже пытаться не станем анализировать социокультурный базис огромного потока функциональной информации, которую мы извлекаем из окружающих нас объектов, но лишь засвидетельствуем его существование, сосредоточившись на том, как воспринимается зрением функциональное назначение объектов окружающего мира.

Стоит отметить, что множество классических материалов по восприятию исключают детекцию функции из собственно перцепции, полагая эту детекцию неким более



**Рис. 9.1.1.** Прямое и не прямое восприятие функции. Гибсон утверждал, что через аффордансы возможно непосредственное (прямое) восприятие функции объекта из его свойств. Альтернативный подход — восприятие через классификацию, когда функция объекта извлекается из памяти после того как объект классифицирован.

поздним процессом, задействующим ассоциативную память (например, Rock, 1975). Причина в том, что ежели восприятие физических свойств объектов возможно без предварительного опыта визуального общения с ними, то детекция функции — невозможна. Существуют очевидные доказательства этого тезиса, но небрежение функциональным компонентом зрения образует серьезный пробел в толковании нашего перцептивного опыта. Таким образом, вопрос зрительного восприятия функции объектов — критично важен в науке о зрении.

Существуют два базовых теоретических подхода к изучению зрительного восприятия функции объектов внешнего мира, схематично представленные на рис. 9.1.1.

1. *Аффордансы.* Прямой или, как еще говорят, — непосредственный подход, предложенный Джеймсом Гибсоном, полагавшим, что наблюдателю исследующему объекты дают возможность воспринимать некоторые их функции непосредственно, читая ее из видимой структуры в динамическом оптическом массиве. Ученый назвал эти функциональные свойства *аффордансами*.

2. *Классификация.* Второй — не прямой или опосредованный подход, проповедуется большинством исследователей, полагающих, что функция воспринимается сопоставлением перцептивной структуры объекта с внутренними представлениями известных *категорий* объектов. Функция затем определяется восстановлением связей между категорией объекта и его известным назначением.<sup>1</sup>

Начнем мы разговор с гибсоновых аффордансов, потому что он и проще и основательнее. К сожалению, о восприятии аффордансов известно довольно мало и в основном

<sup>1</sup> Строго говоря, восприятие функции и классификация независимы друг от друга. К примеру, мы вполне можем воспринимать предмет как носимый, при этом понятия не имея о том, что он из себя представляет. И наоборот: мы можем хорошо знать, что сие за предмет, не очень понимая его назначение. На практике, однако, если мы знаем, к какой категории относится объект, то в целом представляем себе его миссию, равно как и то, где его можно найти и где хранятся остальные.



потому, что особо никто не пытался исследовать их. Тому есть ряд причин. Во-первых, Гибсон умер вскоре после того, как начал разработку своей теории; во-вторых, потому что его рассуждения достаточно сложны, чтобы говорить о них коротко; третья причина в том, что концепция классификации так давно доминирует в науке, что идею аффордансов просто не воспринимают всерьез.

После короткого обзора концепции аффордансов мы в оставшейся части главы будем детально разбирать вопросы классификации объектов, что образует костяк большинства современных исследований восприятия функции объектов из оптической информации.

### 9.1.1 Аффордансы

Исторически традиционный подход к восприятию функции — это классификация в логике ассоциативности и бессознательных инференций. Проще говоря, назначение объекта читается через его внешний вид, активирующий категорийное представление, связанное с известным назначением объекта через ассоциации в памяти. Сей процесс можно рассматривать как разновидность инференции: этот объект — стул; на стульях сидят; следовательно, этот предмет, для того, чтобы на нем сидеть. Сия точка зрения предложена британскими эмпириками, а затем проработана Гельмгольцем, Вундтом и их современниками.

Первый несогласный голос зазвучал из среды гештальтистов, которые стали говорить о том, что определенные аспекты функции объектов можно воспринимать напрямую, непосредственно. Курт Коффка писал: «Первобытному человеку вещи сами говорили о том, что с ними делать: яблоко сказала “съешь меня”, вода сказала “пей меня”, гром сказал “бойся меня”, а женщина сказала “люби меня”» (Koffka, 1935, стр. 7). Ту идею, что некоторые объекты сами открывают свою миссию наблюдателю, гештальтисты окрестили *физиономическим характером восприятия*. К сожалению, концепт не был проработан, отчего оставался под спудом, пока за него не взялся Гибсон (1979), ожививший идею и углубивший ее в свете своей аффорданс-доктрины.

Аффорданс можно понимать как акт *воздействия* предмета на организм или *взаимодействия* с ним, что позволяет воспринимать этот объект непосредственно.<sup>1</sup>

Гибсон утверждал, к примеру, что мы можем видеть, как предмет позволяет себя взять, или сесть на него, или войти в него, или резать им — и при этом без предварительной классификации, как, скажем, баскетбольный мяч, стул, дверь или нож. Действительно, несмотря на значимость их функциональных свойств в созидательной деятельности, мы даже и не пытаемся относить их к какой либо категории или даже дать простое лингвистическое выражение объектам, определенным большинством аффордансов. Их можно описать лишь одной фразой, наподобие «вещь, которую можно взять в руки», «вещь, на которую можно сесть» и т.д. На стулья, в первую очередь, можно садиться, но садиться можно и на многие другие объекты — бревна, кровати, пни, табуретки и многие-многие другие прочные предметы, обладающие соответствующими функциональными признаками: более или менее ровной поверхностью на уровне колена, комфортно подпирающей наши ягодицы и выдерживающей наш вес.

В основе прямого восприятия гибсоновых аффордансов лежат два важных условия:

<sup>1</sup> В данном случае «непосредственно» означает «без промежуточной активации ассоциаций или бессознательных инференций через промежуточное категорийное представление». Сие не обязательно означает через прямое восприятие (см. раздел 2.1.3), хотя оба значения плотно соотносимы с гибсоновой экологической теорией.



**Рис. 9.1.2.** Объекты сходной структуры, но разных функций. Почтовый ящик «предлагает» отправить письмо, но ящик мусорный — очевидно, нет, несмотря на то, что обладает сходными физическими признаками: размером, вытянутостью вверх, окраской и клапаном, открыв который, можно внутрь ящика положить письмо или мелкий почтовый пакет.

1. *Функциональная форма.* Отношения между формой объекта и его аффордансом должны быть полностью «прозрачны», то есть — комбинация видимых свойств объекта должна быть достаточной для определения того, действительно ли объект этот выдает аффорданс (например, обладает должной высотой, стабильностью и уровнем верхней поверхности, подходящими для «садимости» на него). Следовательно, аффордансы не могут произвольно соотноситься со структурой объекта — функция должна логически *проистекать* из формы.

2. *Соотносимость с наблюдателем.* В контексте взаимоотношений с наблюдателем, аффордансы — это функциональные свойства объектов: они — суть возможность *взаимодействия* объекта с организмом воспринимающим, возможность, даваемая самим этим объектом. Скажем, пенек в лесу предлагает взрослому сесть, но ребенку — влезть на него, то есть, то, чего не предложит взрослому.

Согласно Нейссеру (1989) функциональные свойства объектов, удовлетворяющие условиям функциональности формы и соотносимости с наблюдателем, мы можем называть *физическими аффордансами*, связав их физическую структуру с их функционалом. Благодаря такой связи функцию объекта можно в потенциале воспринимать прямо из информации, несомой оптическим массивом (если физическая структура объекта реализует эту функцию).

Однако не все аффордансы, описанные Гибсоном, подходят под сей шаблон. Ученый полагал, что, например, яблоко предлагает съесть его, карандаш — взяться писать, почтовый ящик — отправить письмо и пр. (Gibson, 1979), однако ни одна из этих функций не выводится из одной лишь зрительной информации. Возьмем, к примеру почтовый ящик. Несмотря на то, что почтовые ящики можно воспринимать как носителей сложного физического аффорданса, предлагающего запихивать в них конверты и мелкие бандерольки, все-таки чтобы послание отправилось в путь нужно нечто большее — мусорные урны имеют порой такой же размер и отверстие, в которое пролезут и конверты и мелкие пакеты (см. рис. 9.1.2), но если бросить последние в мусорный ящик, письмо не пойдет к адресату только потому, что урна похожа на ящик почтовый. Чтобы отправление начало ход по назначению нужно, чтобы почтовые работники извлекли пакет из ящика и доставили его в почтовое отделение, а не в помойку, куда обычно отправляют содержимое уличных урн. Сие свойство почтовых ящиков никак не читается из их оптической структуры без связи с памятью, направляющей нашу

мысль. Такие функции действительно воспринимаются только через ассоциации и не являются физическими аффордансами.

В строгом смысле восприятие физических аффордансов не требует промежуточной классификации: к примеру, наблюдатель четко понимает, что на предмет можно сесть без предварительной классификации его как стула. Впрочем, говоря «прямое восприятие», мы отнюдь не предполагаем, что непременно безошибочное: аффордансы невоспринимаемы, всякий раз когда информация передается невизуально. Например, бревно может *выглядеть* вполне «садибельным», но ежели оно гнилое, то не приглашает сесть на него взрослого человека. Аналогично бейсбольный мяч может как бы приглашать бросить его, но если он прочно приклеен к полу, то вряд ли станет бросаемым.

Как уже было сказано, физические аффордансы проявляют себя только в тех случаях, когда восприятие функции возможно напрямую, то есть без посредничества классификации. Следовательно, они являются функциями, которые воспринимаемы у совершенно новых для нас объектов и на далеких планетах. У нас наверняка не вызовет затруднений восприятие поверхностей, позволяющих ходить по себе, объектов, позволяющих бросать себя или садиться на себя, даже если мы их видим впервые. Сие без сомнения важно в плане выживания популяции, поскольку касается критически важных поведенческих свойств — движения, поддержки, манипуляции и прочих форм базового физического взаимодействия непосредственно с внешним миром.

Знаменитый германо-американский психолог Ульрих Нейссер (1989) углубил гибсонов анализ функции, положив аффордансы и классификацию до такой степени различными режимами восприятия, что задействуются разные нейронные системы мозга. Вспомним (гл. 1), что мы отметили разницу между системами «Что?» и «Где?» в зрительной коре.

Л. Унгерляйдер и М. Мишкин (1982) полагали, что вентральная система «Что?» действует в направлении от зрительной коры в сторону коры инфратемпоральной, а дорсальная система «Где?» — от зрительной коры в сторону задней париетальной. Нейссер (1989) предположил, что дорсальная система лежит в основе гибсонова экологического восприятия: захвата информации о расположении поверхностей объектных аффордансов, т.е. деятельности организма исследующего. Ученый полагал систему «Что?» лежащей в основе распознавания и классификации, требующих доступа к внутренним представлениям категорий в памяти. Система «Где?», по его мнению — это система прямого восприятия, не опосредованного памятью на известные категории объектов.

Нейссерова позиция созвучна недавней реинтерпретации гипотезы двуединства зрительной системы, выполненной британцами М. Гудейлом и А. Мильнером, утверждавшими наличие двух подсистем «Что?» и «Как?» (Goodale, 1995; Milner & Goodale, 1995). Они полагали, что вентральная система как нейронный субстрат осознанной зрительной перцепции нацелена на идентификацию объектов в целях высокоуровневого планирования произвольных действий. Дорсальная система, по мнению ученых, — это параллельная система исполнения произвольных действий в реальном времени, в частности таких как хватка, перемещение и отпускание объектов в сугубо практических зрительно обусловленных задачах. Исследователи предположили также, что системам, ориентированным на действие, вовсе необязательно сознание и что они могут работать даже тогда, когда вентрально-перцептивная система разрушена.

В доказательство своей точки зрения Мильнер и Гудейл (1995) доложили о случае, бывшем с больной D.F., вентральная система которой была поражена, но дорсальная оставалась в порядке. Больная почти случайно решала простые перцептивные задачи, такие как подбор двух четырехугольников по форме, определение ориентации линий и размера объектов. Совершенно неожиданным явился тот факт, что, несмотря на оче-

видную недостаточность сознания и способности рассказать о своих ощущениях, она точно и легко могла исполнять простые зрительно наведенные моторные задачи. К примеру, больная столь же успешно, что и здоровые испытуемые, помещала литероподобный объект в лоток, предназначенный фигурам определенной ориентации, и могла плавно и точно управлять разведением пальцев, когда брала в руку объекты разной формы. У больных с повреждениями дорсальной системы обычно обратная картина: они в состоянии дать точную перцептивную оценку, но им затруднительно управление действиями, требующими зрительной обратной связи. Описанные наблюдения вполне согласуются с возможностью того, что дорсальная система может задействоваться при восприятии простых объектных аффордансов и подчинять соответствующие действия зрительному контролю.

Однако, несмотря на возможную пользу и значимость идеи аффордансов, концепция не в силах учесть всю функциональную информацию, которую мы получаем об объектах. В частности, первым условием функциональной формы является исключение объектов, чью форму нельзя определить непосредственно из их физических свойств без дополнительной информации, полученной из предыдущего опыта. Сие включает несколько случаев, которые Гибсон обсуждал в логике аффордансов (например, съедобность яблок, письмопригодность карандашей, письмоотправность почтовых ящиков и пр.), равно как и множество человеческих артефактов, чьи функции визуально скрыты: видеокамеры, CD-плееры, музыкальные центры, к примеру, обладают формой, из которой никак не читается их функция. Сказанное касается и других антропогенных объектов: лампочек, часов, музыкальных инструментов, книг, печей, холодильников и пр. И хотя мы можем вывести их функцию из пристального и детального рассматривания и взаимодействия с ними, но, ежели не узнаём их сразу, зрение позволяет укоротить процесс изучения через ассоциации с хранимым в памяти. Сие значит, что классификационный подход к восприятию функции имеет полное право на существование.

### 9.1.2 Непрямое восприятие функции через классификацию

Как было сказано выше, классификация требует, в первую очередь, восприятия внутренних свойств объекта, дабы отнести этот объект к тому или иному классу, а затем восстановить функцию объектов этого класса из памяти. Отметим, что формально запрета на получение функциональной информации таким путем, не существует. Сие безусловно возможно, поскольку связь между категорией объекта и его функцией потенциально случайна и зависима только от ассоциаций, установленных предшествующим опытом.

Мы можем спросить себя, действительно ли аффордансы или классификация наилучшим образом толкуют восприятие функции человеком? По всей видимости, используется и первое, и второе, но для разных целей, включающих в себя восприятие, целеполагание и действие. Очевидно и логически невозможно, чтобы мы классифицировали всё, попадающееся нам на глаза — многое из того, что нам необходимо знать о плотно окружающем нас пространстве в целях локомоции и базового моторного контроля, содержится в аффордансах: Ровна ли земля, чтобы идти по ней?; Действительно что-то нужно, дабы безопасно переместиться отсюда туда?; Можно ли этим объектом манипулировать, или воздействовать на него движениями иных частей тела? Для решения более сложных целенаправленных задач, таких, например, как взбивание яиц для торта, выбранные объекты классифицируются нами без всяких усилий. Сие входит в состав плана преднамеренного действия, а затем используется по установленному шаблону (и вполне вероятно, что путями, задействующими в реальном времени оцен-



ку простых аффордансов, таких, например, как взять миску и вилку для взбивания яиц). На практике отличие между аффордансами и классификацией в дальнейшем стирается, поскольку отношения формы объектов с их функцией могут меняться от очень строгих (например, объекты, на которые можно сесть) до очень слабых (например CD-плеер) с множеством промежуточных градаций. Проще говоря, и аффордансы, и классификация очевидно нужны в повседневном восприятии.

### 9.1.2.1 Четыре компонента классификации

Итак, приняв то, что мы чаще классифицируем объекты на базе исходного знания их функциональных классов, выясним, каковы необходимые условия этой классификации. Зрительная классификация объекта требует наличия четырех базовых компонент:

1. *Объектного представления.* Значимые характеристики классифицируемого объекта должны восприниматься зрительной системой и представляться в ней.
2. *Категорийного представления.* Каждое из множеств возможных категорий должно быть представлено в памяти в доступной для зрительной системы форме.
3. *Процесса сравнения.* Должен существовать путь, по которому представление объекта согласуется (сравнивается) с категорийным представлением этого объекта.
4. *Процесса принятия решения.* Должен существовать какой-то алгоритм принятия решения в отношении того, к какой категории данный объект принадлежит.

Недавно (в гл. 7 и 8) мы обсудили проблему экстенсивного представления свойств объектов и пришли к тому, что форма — это наиболее важный тип информации для классификации (Biederman & Ju, 1988). Хотя прочие свойства, такие как текстура, окраска, размер и ориентация объекта — тоже весьма важны. Как обсуждалось в гл. 8, шаблоны, спектры Фурье, списки признаков и структурные описания — это главные классы представления формы, но, разумеется, возможны и различные их гибридные варианты. Теперь же настало время обсудить некоторые аспекты процессов сравнения и принятия решения. Совместно с видами представления формы сии соображения приведут нас к той мысли, что круг возможных теорий классификации объектов в зрительной системе человека весьма ограничен.

### 9.1.2.2 Процессы сравнения

Процессы сравнения требуют согласования объектного представления с представлением категорийным. Чтобы сие состоялось, представление объекта и представление категории должны быть одного *типа*, хотя их специфическое содержимое может при этом разниться (что чаще всего и бывает). Иными словами, представление объекта может явиться шаблоном, списком признаков, структурным описанием или еще чем-то, но при этом представление категории должно быть того вида, который допускает возможность согласования. Попытка сравнить, например, шаблон со списком признаков априори обречена. Нужны процессы, конвертирующие шаблон в список признаков, который сравнивается затем с другим списком, либо же процессы, конвертирующие список признаков в шаблон, который сравнивается затем с другим шаблоном, или же конвертирующие оба представления в представления какого-то третьего вида и т.д. Но при этом всякий раз сравниваются представления только одного и того же вида.

Одним важным замечанием в отношении процессов сравнения явится то, что процессы эти могут быть как серийными (последовательными), так и параллельными (од-



новременными). То есть, нам необходимо получить ответ на два принципиальных вопроса:

1. Процессы согласования представления объекта с множеством всех категориальных представлений идут последовательно или параллельно? Вернее, действительно ли представление объекта согласуется только с одним представлением категории в данный момент времени или одновременно со всеми возможными категориями? Допуская то, что у современного человека имеется огромное число объектных категорий (по некоторым данным порядка 30000 — Biederman, 1987), вероятнее всего работа ведется параллельно. Иначе классификация будет идти слишком долго — лев успеет сожрать нас, пока мы сообразим, что перед нами лев.

2. Предположив, что представление каждого объекта состоит из множества элементов (признаки, размерности, части и пр.), мы ставим второй вопрос — Элементы данного категориального представления сравниваются последовательно (серийно) или сразу (параллельно)?

Рассмотрим, к примеру, представление списком признаков. Если представление категории несет, скажем, 20 признаков, значит ли это, что процедура займет вдвое больше времени, чем согласование с другой категорией в 10 признаков (последовательный процесс сравнения), или она займет столько же времени (как то прогнозирует процесс параллельного сравнения)? Или как-то иначе? Ответы на эти вопросы далеко не очевидны.

### 9.1.2.3 Процессы принятия решения

Финальный этап классификации — это принятие решения о том, к какой категории отнести данный объект. Решение сие — явно или неявно — базируется на результатах процесса сравнения представления объекта с одним или более представлениями категории. Давайте предположим, что каждое сравнение можно описать в логике отдельного значения, числа, представляющего степень подгонки или подобия представления объекта представлению категории. Это будет одно значение по каждой категории, что укажет нам на то, насколько плотно представление объекта согласуется с представлением категории. Как только эти величины будут определены, процесс принятия решения может начать работу в логике того, как сии значения можно использовать для выбора правильной категории.

При выборе способа принятия решения играют роль два основных фактора:

1. *Новизна*. Порой мы зрим объекты, которые нельзя отнести к какой-либо из известных категорий. Они относятся к неизвестной категории и воспринимаются as is. На рис. 9.1.3 находим пример: наше правило принятия решений способно, в принципе, распознать то, что для него сей объект новый и что можно установить новую категорию, нежели ошибочно отнести его к какой-то из известных.

2. *Уникальность*. Поскольку многие категории объектов взаимно исключают друг друга (не имеют общих членов), каждый объект является членом только одной из них. Например, зверь не может быть собакой и котом одновременно или вещь не может быть одновременно аккордеоном и скрипкой. В идеале процесс принятия решения должен позволять сему факту отражаться при классификации.

Стоит отметить, что объект может корректно идентифицироваться как член нескольких различных категорий, например: данный конкретный авиалайнер может быть марки Embraer, пассажирским самолетом, воздушным судном, аппаратом, артефактом. Однако сие никак не отменяет уникальности категорий, поскольку категории эти не являются взаимно исключающими, но вложенными или пересекающимися.

ся. И нам следует немного поговорить об иерархической структуре категорий.

Предположим, что для определения того, насколько плотно классифицируемый объект относится к каждой из известных категорий, существует два классических подхода принятия решения: *правило порога* и *правило максимума*.

**1. Правило порога.** Существует для установки критериального (порогового) значения результата сравнения, и последующего назначения обработанному объекту какой-то из категорий, когда превышено это значение.

В принципе, правило порога способно помочь распознать новизну: даже если не существует категории, превышающей порог, система может решить, что представление объекта ново и действует должным образом. Однако обратная сторона в потенциальной утрате уникальности, поскольку более одной категории могут одновременно превысить порог. Следовательно правило порога приводит к множественности классификаций между взаимно исключающими категориями. К примеру, невозможно гарантировать то, что зверь не воспримется одновременно и как волк, и как собака.

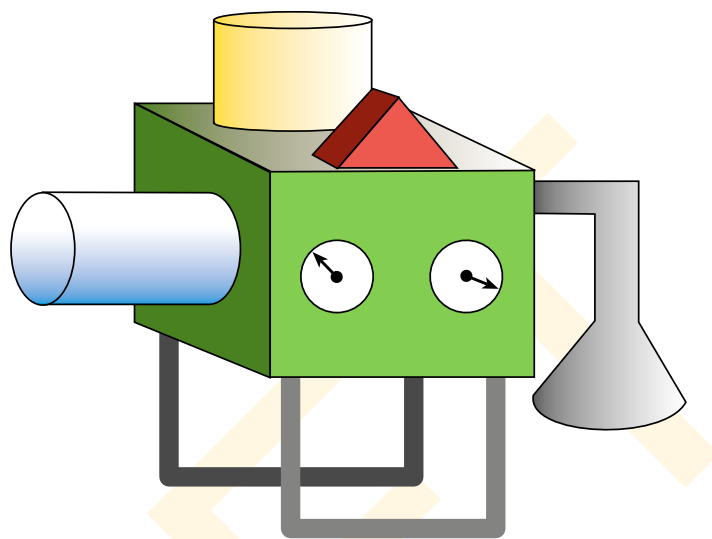
**2. Правило максимума.** Это выбор категории с наивысшим (максимальным) значением показателя стыковки.

Правило максимума непременно назначает объекту уникальную категорию, исходя из того, что почти невозможно, чтобы более чем одна категория имела максимальное значение в одно и то же время. Однако оно бессильно в детекции новизны, поскольку, по определению, всегда существует одна категория, имеющая максимальное сродство интересующему объекту. Следовательно, правило максимума всегда назначает объекту наилучшим образом подходящую категорию, не обращая при этом внимания на то, каковой может стать наихудшая стыковка.

Вместе с тем, оба правила можно объединить в третье, избавившись от недочетов первого и второго. Такое правило мы назовем *правилом максимума через порог*.

**3. Правило максимума через порог.** Суть подхода в установке порога, ниже которого объекты будут восприниматься как новые, но выше которого выбирается категория с наивысшим значением. Сия комбинация дает возможность идентифицировать объекты как новые без результирующей неопределенности, когда многие категории превышают порог.

Стоит отметить, что процесс принятия решения никак не требует наличия «внутричерепного гомункула», оценивающего выходной продукт процессов сравнения и «решающего» вопрос о том, действительно ли данное значение выше порога и которое из значений максимальное. Вполне допустимо (и действительно желаемо), чтобы решение являло собой полностью механистический процесс. К примеру, нейронная сеть класса



**Рис. 9.1.3.** Неизвестный ранее предмет, иллюстрирующий новую проблему: порой мы воспринимаем объект, но при этом не можем отнести его к какой-либо известной ранее категории.

«Победителю достается всё», состоящая из целиком взаимно подавляющих единиц, — это один из алгоритмов реализации правила максимума. Единица с наивысшей активацией будет в большей степени подавлять все прочие единицы, нежели наоборот, и, следовательно, в конечном итоге, только одна единица останется на устойчивом уровне активации (см. раздел 6.5). Ингибиторные взаимодействия между единицами весьма эффективно и без помощи гомункулов «решают», какая из единиц наиболее активна.

\* \* \*

Итак, мы рассмотрели некоторые концепции зрительной классификации объектов: различные аспекты представлений (см. гл. 8) и некоторые аспекты сравнения и принятия решений (текущая глава). В этих рамках обитает огромное число теорий, чьи свойства требуют объяснения в логике их способности учитывать феноменологию перцептивной классификации зрительной системой человека. Мы же рассмотрим лишь некоторые из наиболее значимых феноменов. После чего разберем примеры теорий объектной классификации и поговорим о том, как можно истолковать происхождение этих феноменов.

## 9.2 ФЕНОМЕНЫ ПЕРЦЕПТИВНОЙ КЛАССИФИКАЦИИ

Цель любой теории в объяснении наблюдаемых фактов. Сравнивая теоретические прогнозы с фактами, мы обнаруживаем несостоятельность многих концепций, но их дальнейший пересмотр порой дает возможность повысить прогностическую эффективность смежных научных гипотез. В случае перцептивной классификации, некоторые из требующих объяснения фактов рождаются из структуры собственных человеческих категорий, другие же — из оценки человеческой деятельности при решении задач классификации. Обе группы значимы в построении жизнеспособной концепции зрительной классификации.

### 9.2.1 Иерархия категорий

Первое, на что следует обратить внимание в контексте разговора о зрительной классификации объектов, это на тот факт, что хотя мы обычно и рассуждаем об объектах, как о принадлежащих к какой-то одной категории, объекты эти *могут* принадлежать одновременно нескольким категориям. Знаменитая Лесси<sup>1</sup> — это собака, но она же — колли, она же — млекопитающее, она же — зверь, живое существо, любимец хозяев, телезвезда и т.д. Категории человеческого восприятия и познания — суть сложные сущности с различными структурными отношениями. Полноценный анализ этих отношений дает важную информацию для понимания нами законов перцептивной классификации.

Возможно, базовым свойством категорийной структуры является ее иерархия. Один из способов представления иерархии — это древо (рис. 9.2.1 А), где узелки означают классы объектов: вышестоящие узелки относятся к более крупным классам, нижестоя-

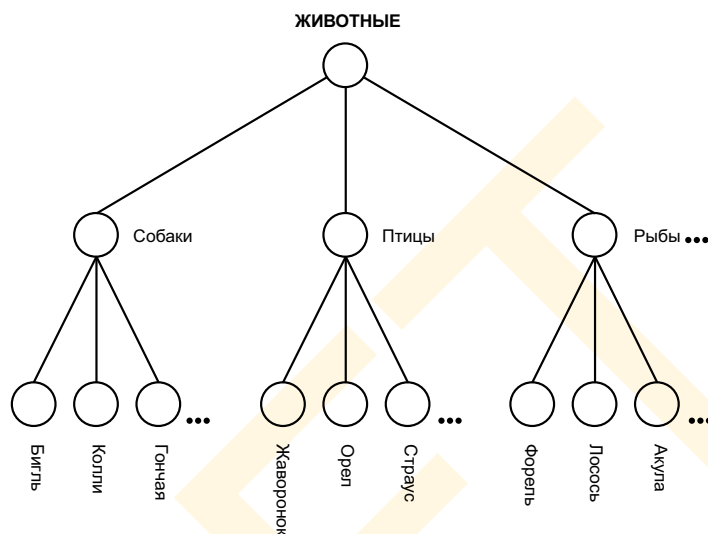
<sup>1</sup> *Лесси* (англ. Lassie) — вымышленная собака породы колли, персонаж многих фильмов, сериалов и книг. Образ Лесси был создан англо-американским писателем Эриком Найтом (Eric Mowbray Knight; 1897 — 1943). В 1938 году он опубликовал в журнале «Saturday Evening Post» рассказ «Лесси возвращается домой», а в 1940 г. развил его до романа, который был переведён на 25 языков.

ящие — к мелким, а терминальные узелки — к индивидуальным объектам. «Вертикальная» структура древа указывает, к примеру, на то, что колли — это собаки и что все собаки — животные. «Горизонтальная» — на то, что может существовать множество различных подкатегорий внутри данной категории. Например, категория собак включает в себя (помимо колли) биглей<sup>1</sup>, боксеров, такс, овчарок и т.д.; животные (помимо собак) — кошек, обезьян, птиц, рыб и т.д.

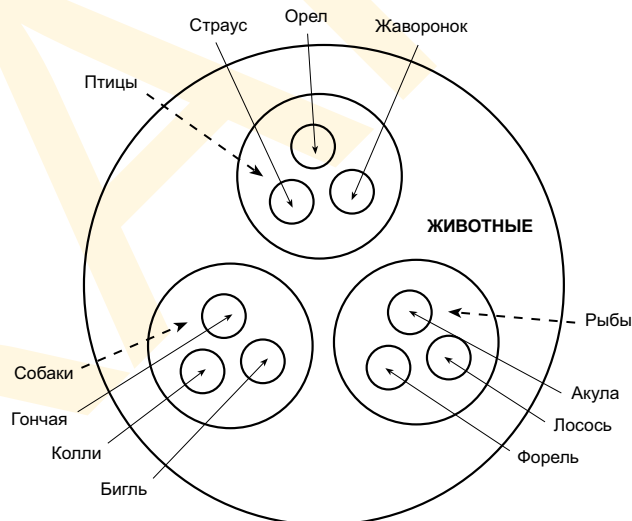
Второй способ представления иерархической структуры категорий — это отрисовка т.н. *диаграмм Венна* (рис. 9.2.1 В). Диаграммы Венна отражают категории в виде неких отграниченных областей как если бы границы заключали в себе всех членов данной категории. Тот факт, что все собаки — животные, представлен тем, что круг, соответствующий собакам целиком внутри круга животных. То есть, иерархия категорий отражена вложением участков. Таким образом и венновы диаграммы, и структурные древа можно использовать для представления иерархизма объектных категорий.

### 9.2.1.1 Прототипы

Один из наиболее серьезных вопросов в сфере объектной классификации касается статуса отдельных экземпляров по отношению к той или иной категории: не все собаки выглядят одинаково, не все птицы и не все лошади. Как следует понимать вариабельность отдельных экземпляров внутри категории? Своей фун-



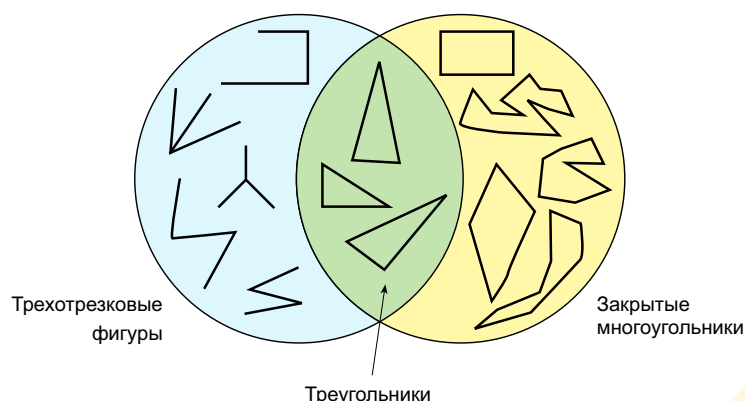
А. Иерархическое представление древом



В. Представление венновой диаграммой

**Рис. 9.2.1.** Два представления иерархической структуры категорий. В первом случае (А) представлена иерархия в логике нейронной сети, в которой верхние узелки охватывают нижние. Веннова диаграмма (В) представляет гнездовую иерархию — большие круги включают в себя меньшие.

<sup>1</sup> Бигль — английская гончая — *Прим. пер.*



**Рис. 9.2.2.** Определение категорий в логике необходимости и достаточности. Треугольники лежат на пересечении множества всех трехлинейных фигур и множества всех закрытых многоугольников (зеленоватая область).

должен обладать, чтобы претендовать на членство в данной категории. Хороший пример — геометрическая категория треугольников. Чтобы назваться «треугольником» фигура должна быть закрытым полигоном с тремя строго прямыми сторонами. Каждое из этих двух свойств само по себе — это *необходимое* условие того, чтобы нечто явилось треугольником, ибо все треугольники обладают этими свойствами. Однако каждый признак сам по себе не является *достаточным*, поскольку и другие фигуры, не являющиеся треугольниками, могут стать закрытыми многоугольниками (например, квадраты) или иметь по три стороны (например, стрелы). Однако вместе они составляют достаточное условие, поскольку любой объект, обладающий обоими — треугольник.

Если категорию можно определить, через набор необходимых и достаточных условий, сие позиционирует простую бинарную концепцию членства в категории, к примеру, фигура либо является треугольником, либо нет, без промежуточных вариантов. Сие легко проиллюстрировать венновыми диаграммами: все треугольники можно представить затемненной областью, обозначенной «Треугольники» (рис. 9.2.2), поскольку содержимое отграниченной области бинарно — каждая точка либо внутри, либо вне границ. Отношения между категориями можно представить затем наложением областей венновой диаграммы. Рис. 9.2.2 иллюстрирует то, как категория треугольников соотносится с категориями закрытых полигонов и категорией трехлинейных фигур. Как требуется по определению в логике необходимости и достаточности, треугольники окажутся на пересечении этих двух больших множеств.

С логической, математической точки зрения аристотелев подход неоспорим. Но насколько сие концепция подходит категориям перцептивным? Могут ли собаки, рыбы, стулья, деревья и пр. охарактеризоваться в логике списка необходимых и достаточных условий бинарного членства в категории?

Много лет ответ на этот вопрос был утвердительным, главным образом потому, что не существовало внятной альтернативы. Однако знаменитый австро-британский философ Людвиг Виттгенштайн (1953) раскритиковал аристотелеву точку зрения и показал, что сие принципиально неверно в отношении категорий реального мира. Ученый задал себе вопрос: Каковы необходимые и достаточные признаки характерные для всех случаев? Он доказал, что такого множества признаков не существует и что одни и те

даментальной заинтересованностью в том, как категории определить в логике перцептивных представлений, сей вопрос нацеливается в самое сердце проблемы классификации объектов.

Классический ответ на вопрос о том, как определять категории, прозвучал еще из уст Аристотеля: философ утверждал, что категории строятся набором правил, характеризующих *условия необходимости и достаточности* члена категории, то есть списком свойств, которыми объект



же аргументы приложимы к большинству естественных категорий. Как наиболее подходящую альтернативу Виттгенштайн предложил идею т.н. *семейного сходства*. Лица членов одной семьи часто распознаются как таковые не потому, что все они обладают множеством специфических общих черт, но потому, что они в целом подобны друг другу неким образом, непостижимым формальной логикой.

Вслед за Виттгенштайном, в 70-х годах прошлого века американская психолог Элеонора Рош в серии весьма изящных и убедительных исследований продемонстрировала ущербность аристотелевой позиции (Rosch, 1973b, 1975a, 1975b; Rosch & Merris, 1975). Ее данные радикально изменили позиции когнитивной науки в отношении природы перцептивной классификации зрительной системой человека. Вспомним, что в гл. 3 мы говорили о ее исследованиях структуры цветовых категорий, дающих идеальный пример цветовой классификации. Исследования, выполненные Рош в остальных сенсорных модальностях привели исследовательницу к убеждению, что все естественные категории могут структурироваться сходными путями, то есть, в логике центрального или идеального образца, названного ею *прототипом*.

Идея, лежащая в основе теории прототипов классификации, состоит в том, что прототипы представляют каждую категорию в логике наилучшего образца. В большинстве случаев прототип категории — это «средний» член. Прототипическая собака, к примеру, будет «наисобачнейшей» собакой — обычная «дворняга», средних габаритов, стандартной масти и типично собачьей формы тела. В других случаях прототип — это «идеальный» образец, каким находим фокальные цвета. Прототипический красный — это чистейший, насыщеннейший, краснейший из возможных красных, но средний среди них. В обоих случаях прототипы играют привилегированную роль в классификации.

Прототипия в структуре категорий, предложенная Рош, отличается от классической аристотелевой доктрины двумя ключевыми моментами:

1. *Логическими и эталонными представлениями.* Теория прототипов в категориальных представлениях определена относительно некоего специфического эталона внутри каждой категории (прототип), но не набора логических правил, основанных на списке необходимых и достаточных условий.

2. *Бинарными и градуальными членами.* Теория прототипов позволяет наличие промежуточных градаций (или степеней) членства в категории, в противоположность аристотелевому бинарному концепту членства, либо не членства.

Мы уже сталкивались с бинарным и градуальным членствами в гл. 3, когда говорили о категориях цветовых ощущений. Мы тогда описали отличия между множеством классических теорий, предполагающих строго бинарное различие между членством и нечленством во множестве, а также теорию нечетких множеств, предполагающую существование плавных переходов из категории в категорию (Zadeh, 1965). Таким образом, концепция прототипов в перцептивной классификации весьма близка идее нечетких множеств.

Рош проверила некоторые положения своего концепта по естественным категориям (Rosch, 1975b). К примеру, она рассудила так, что объекты как представители той или иной категории, должны варьировать по критерию их «добротности». Когда испытуемых просили распределить собак в логике «типичности», они очень высоко оценили биглей и сенбернаров, а чихуахуа — весьма низко. Аналогично испытуемые отнесли дроздов и воробьев к типичным примерам птиц, а страусов и пингвинов — к нетипичным. Сей рейтинг типичности позволяет прогнозировать то, как быстро испытуемые способны ответить «истинно» или «ложно» на вербальное утверждение наподобие «Дрозд — это птица» и «Пингвин — это птица». Дальнейшие эксперименты показа-

ли, что изображения (рисунки) объектов — представителей типичных примеры категорий, подобно вербальным, перцептивно классифицируются быстрее представителей нетипичных (Ober-Thompkins, 1982). То есть, на верификацию пингвина как птицы уходит больше времени, чем на то же в отношении дрозда. Следовательно, время, требуемое для классификации объекта как члена категории определяется тем, насколько «типичен» представитель данной категории.

### 9.2.1.2 Категории базового уровня

Рошево исследование прототипов показало, что в первые мгновения зрительного акта в ранжировке категорий задействуется некая очень важная структура, существование которой никак не выводится из аристотелевой классической идеи. Рассуждая об иерархии категорий, обычно говорят о «горизонтальной» размерности классификации, но рошевы данные о прототипах и типичности позволяют нам говорить в том числе о «вертикальной» размерности. К примеру, когда данный объект (скажем, Лесси) классифицирован, на каком вертикальном уровне иерархии сие происходило? На одном из нижних («колли»), или может быть даже на уровне «Лесси»? На одном из высоких, более общих уровней физических объектов, или на каком-то из промежуточных уровней — собаки?

Ответ звучит примерно так, что большинство из нас вначале распознает объекты на некоем *срединном* иерархическом уровне. Лесси, например, большинством из нас вначале классифицируется как собака, но не как животное или колли. Рош назвала сей уровень категорийной абстракции *базовым категорийным уровнем*. Категории уровнем выше она окрестила *старшими*, а те, что уровнем ниже — *младшими*.

Рош определила категории базового уровня по трем критериям: *подобию формы, подобию моторных взаимодействий и общим атрибутам* (Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976).

1. *Сходство форм*. Базовый уровень категорий — это наивысший уровень, на котором у членов категорий подобны формы.

К примеру, форма у всех собак более или менее одинакова; то же справедливо для категорий рыб, птиц, лошадей и т.д. Члены младших категорий (такие как колли и пудели, или лососи и окуни, или же дрозды и пингвины) обладают более или менее сходной формой, но у членов старшей категории («животные») вариабельность форм бóльшая. То есть, если мы возьмем степень подобия формы внутри иерархии категорий как функцию от уровня в этой иерархии (рис. 9.2.3), то увидим резкое падение степени подобия от базового уровня к старшему.

Рош выделяет один любопытный момент: оказывается, если формы членов категорий базового уровня «усреднить», т.е. найти некий средний контур множества собак (если смотреть в общепринятом ракурсе, в общепринятых угловом размере и ориентации и т.д.) — результирующая форма все еще будет распознаваться как член категории. Однако, если средняя форма получена тем же способом, но из членов старшей категории (например, многие виды животных), результирующий контур окажется нераспознаваемым. Опыт подтвердил сей прогноз.

2. *Сходство моторных взаимодействий*. Категории базового уровня — это высокоуровневые категории на которых мы *взаимодействуем* с их членами, задействуя сходные моторные последовательности.

Рассмотрим, к примеру, движения рук пианиста, а затем кларнетиста или гитариста. Каждое из этих движений относится к категориям базового уровня внутри старшей категории — категории музицирования. Имеют место несколько общих движений на

старшем уровне, в частности, такие как движения пальцев по определенным частям инструмента, но, что далеко от большинства подобий в моторном взаимодействии, возникающих на базовом уровне. Мы играем на пианино в целом так же, как и на рояле, но совершенно иначе, чем на гитаре или кларнете. Подобие даже больше, чем когда мы взаимодействуем с различными членами младших категорий, но только за счет той тонкости, что у рояля поднята крышка (чего нету у пианино). Если мы возьмем степень подобия моторных взаимодействий как функцию от уровня в иерархии категорий, то получим обвальное падение с базового уровня на старший (рис. 9.2.3).

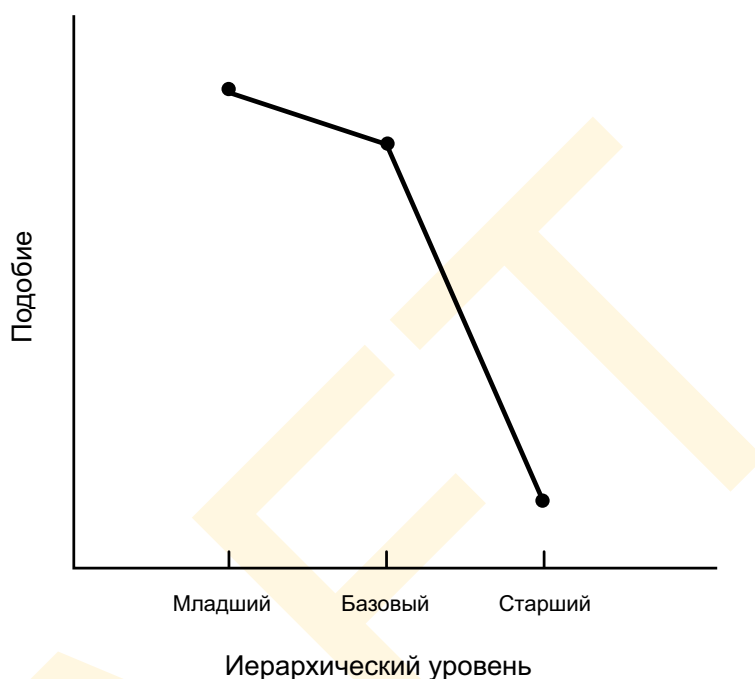
### 3. Общие атрибуты.

Категории базового уровня — это высшие категории, по которым существует большое число атрибутов общих той или иной паре членов. Первые два критерия (сходство форм и сходство моторных взаимодействий) можно фактически рассматривать как частный случай более общего критерия общих атрибутов, поскольку форма и моторные взаимодействия это лишь два очень важных атрибута объекта и того, как мы используем их. Ту же картину наблюдаем, когда любые возможные атрибуты членов представлены списком, включая позиции высокого уровня — число общих атрибутов резко падает с базового уровня на старший (см. рис. 9.2.3).

Из всех перечисленных определений категорий базового уровня подобие формы очевидно ближе всего к вопросу перцептивной классификации. Действительно, Рош по началу решила, что большинство объектов исходно распознаются как представители категории базового уровня (Rosch et al., 1976), но дальнейшие исследования показали, что вопрос существенно сложнее.

#### 9.2.1.3 Категории начального уровня

Пьер Жоликёр, Марк Глюк и Стивен Косслин (Гарвард, 1984) провели серию опытов по изучению перцептивной классификации на базовом, младшем и старшем уровнях. В одном из экспериментов исследователи просили испытуемых дать название разным картинкам, причем сходу, т.е., какое первым придет в голову. Исследователи получили мощное экспериментальное подтверждение той идеи, что *типичные* члены категорий базового уровня действительно именуются на базовом уровне. Что и прогнозирует теория Элеонор Рош. Однако *атипичным* членам свойственно классифицироваться на



**Рис. 9.2.3.** Определение базового категориального уровня через подобие. Подобие лишь слегка снижается при переходе младшего уровня на базовый (например, от колли к собакам), а затем резко падает на старший (например, от собак к животным).

младшем уровне, например, большинство из нас, кто много раз сталкивался со словом «дрозд», назовут самого дрозда скорее попросту птицей, нежели собственно дроздом. Но когда им предъявят страуса, то назовут его скорее «страусом», нежели просто птицей.

Отметим, что сия картина именовании не универсальна и наблюдается, как правило, в отношении членов категорий базового уровня, которые относительно многообразны. Рассмотрим некие категории базового уровня, взятые из старших категорий фруктов (яблок, бананов и грейпфрутов) и животных (собак, птиц и обезьян). Большинство из нас согласится с тем, что вариации формы внутри категории яблок, к примеру, гораздо менее выражены, чем внутри категорий собак. Действительно, по одной лишь форме очень трудно различить сорта яблок, бананов и грейпфрутов, но давайте посмотрим как отличаются друг от друга такса и афганская борзая, пингвин и страус, карась и акула. Неудивительно, что атипичные экземпляры из различных категорий базового уровня склонны именоваться нами согласно их младшей категории.

Поскольку категории, по которым классифицируются объекты исходно отличны от базового уровня, Жюликёр (1984) окрестил их категориями уровня *начального*. Можно подумать, что ученый переопределил базовый уровень. Но это не так: на самом деле его результаты позиционируют непреодолимую проблему в концепции этого уровня, суть которой в том, что рошева конструкция категорий базового уровня определена для полной категории. Однако, данные свидетельствуют, что некоторые члены категорий вначале идентифицируются на базовом уровне, тогда как другие — из той же категории — на старшем. По рошевым критериям категория базового уровня должна быть либо «птицы», либо же какой-то из ее возможных подкатегорий. Концепция категории начального уровня по данному объекту не подпадает под это ограничение и вполне способна истолковать результаты опытов. В отличие от категорий базового уровня категории уровня начального достаточно зависимы от результатов, кроме того уровня, на котором объекты идентифицированы. По этой причине термин «категория начального уровня» — это только название данного уровня.

Перцептивное определение объекта в категорию начального уровня часто называют *распознаванием объекта* или его *идентификацией* (или же *классификацией* этого объекта). Мы не будем употреблять термин «распознавание», прибегая к нему для той специфической ситуации, когда мы видим объект безотносительно к тому, знаем мы к какой категории его следует отнести или нет. К примеру, если мы сейчас посмотрим на абсурдный объект, изображенный на рис. 9.1.3, мы все равно окажемся способными распознать его, не относя при этом к какой-то из известных ранее категорий. Аналогично идентификация объектов лучше приспособлена к распознаванию определенных известных объектов, нежели их членства в категории. Распознавание объекта как моей собственной собаки, лодки, велосипеда — это техническая идентификация объекта, но не объектная классификация. Сие очень тонкое отличие, не имеющее универсального подтверждения. Идентификацию объектов можно отнести к объектной классификации, отметив, что отдельные конкретные объекты — это минимальная категория (содержащая только один объект), но совершенно необязательно, что второе обрабатывается теми же перцептивными процессами.

Стоит отметить, что категории начального уровня (равно как и категории уровня базового) могут варьировать между наблюдателями и, возможно, между разными контекстами. К примеру, для орнитолога или даже просто любителя птиц, крайне небольшое число их видов окажется в категории начального уровня: на протяжении жизни эти люди накапливали перцептивный опыт различения пернатых по видовым признакам, система восприятия этих людей очень тонко настроена на эти признаки. Поэтому



орнитологи воспринимают дрозда именно как дрозда, воробья как воробья, но не как птицу вообще (Tanaka & Taylor, 1991).

### 9.2.2 Ракурсы

Одним из примечательных фактов классификации трехмерных объектов является то, что их можно идентифицировать с любой точки наблюдения, в любом ракурсе. К примеру, домашнее кресло легко читается креслом вне зависимости от того, откуда на него смотреть — спереди, сзади, сбоку, сверху или снизу. Следовательно, одним из важнейших вопросов, на который должна дать ответ любая теория объектной классификации — это «Как же сие возможно?». Мы уже знаем (см. гл. 8), что весьма сложно организовать представление перцептивной трехмерной формы, поэтому вполне можем предположить, что классификация трехмерных объектов (которая, как мы помним, основывается, в первую очередь, на информации о форме), это очень сложная теоретическая проблема.

Априори приняв тот факт, что мы *можем* классифицировать объекты в различных ракурсах наблюдения, очень легко придти к тому, что классификация объектов инвариантна во всех ракурсах. Однако исследования показали, что сие верно лишь на первый взгляд. Мы можем сказать, что способность распознавать трехмерные объекты, несмотря на вариабельность ракурса — это фактор «первого порядка». Но при этом часто упускают из виду фактор «второго порядка», суть которого в том, что определенная перспективная точка наблюдения — ракурс — влияет на *скорость* и *точность*, с которой объект будет классифицирован и распознан. Если мы внимательно рассмотрим серию ракурсов лошади (рис. 9.2.4), нам станет ясно, что некоторые из них более репрезентативны, другие менее. Действительно, некоторые ракурсы совершенно бесполезны, а некоторые — неестественны.

#### 9.2.2.1 Канонический ракурс

Пальмер, Рош и Чейз (1981) системно исследовали и документировали перспективные эффекты в зрительной классификации объектов. Начали с того, что предъявляли испытуемым множество изображений одного и того же объекта и просили дать субъективную оценку схожести на исходный объект по некоей шкале (от 1 — очень похоже, до 7 — совсем непохоже). Задача оказалась вполне решаемой и, в итоге, группа выдала комплект средних рейтингов, показанных под каждым изображением. Ракурсы с наивысшим средним рейтингом даны на рис. 9.2.5.

Затем Пальмер (1981), предъявляя разные объекты в разных ракурсах, просил испытуемых дать название категории начального уровня, к которой относится каждый из объектов с рис. 9.2.5, причем так быстро, как только возможно. Результаты продемонстрировали то, что изображения, которые наблюдатели рейтинговали как обладающие наилучшим ракурсом, именовались быстрее всего, и латентный период именования рос по мере того, как «совершенство» ракурса снижалось (рис. 9.2.6). То есть, оценку «худший» мы даем медленнее, чем «лучший». Легче и быстрее всего идентифицируемые ракурсы Пальмер окрестил *каноническими*.

Существование канонических ракурсов и перспективных эффектов в целом демонстрирует ошибочность нашего интуитивного предположения, что мы одинаково хорошо распознаем объекты с любого ракурса. Существуют очевидные и систематические вариации в длительности латентного (скрытого) периода именования, что должна уметь объяснить адекватная теория объектной классификации.





Наилучший (1.60)



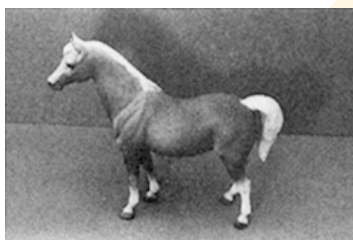
Боковой (1.84)



Передне-боковой (2.12)



Передне-верхне-боковой (2.80)



Верхне-боковой (3.48)



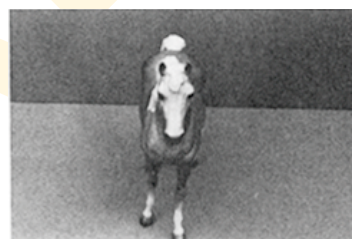
Передний (3.72)



Задне-боковой (4.12)



Задне-верхне-боковой (4.29)



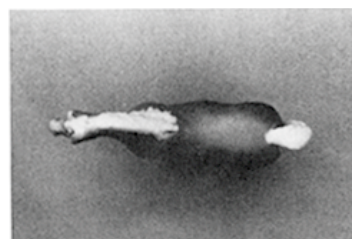
Передне-верхний (4.80)



Задне-верхний (5.56)



Задний (5.68)



Верхний (6.36)

**Рис. 9.2.4.** Канонический ракурс. Представленные здесь ракурсы лошадки сильно разнятся по схожести с исходным объектом. Цифры представляют средние наблюдательские рейтинги по шкале от 1 (очень похоже) до 7 (едва похоже).

По материалам Palmer, Rosch, & Chase, 1981.



Лошадка



Рояль



Чайник



Автомобиль



Кресло



Фотоаппарат



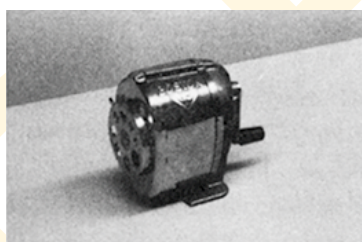
Будильник



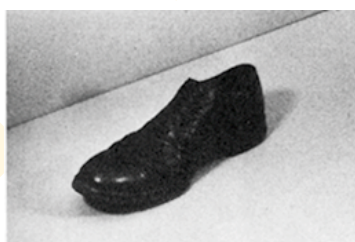
Телефон



Дом



Точилка



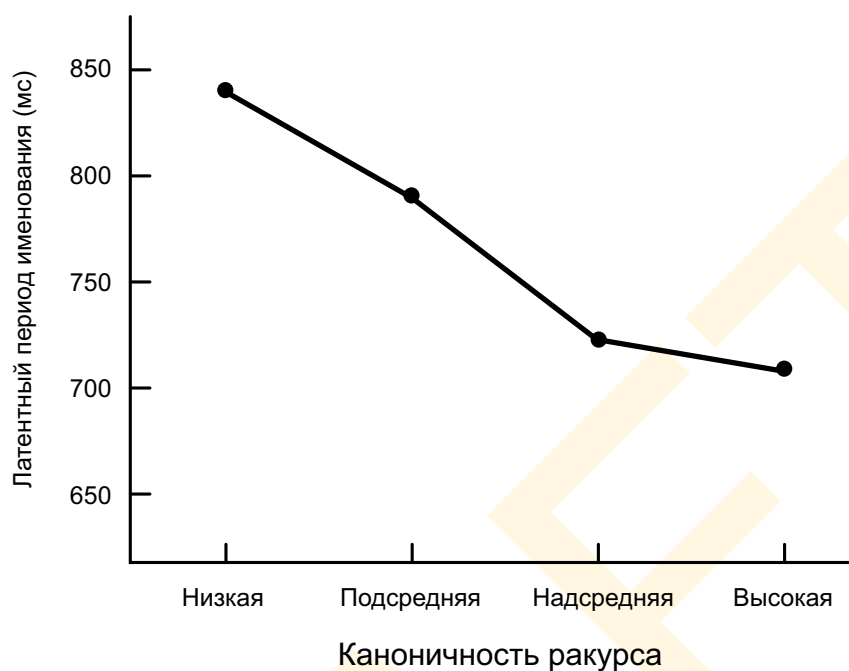
Туфель



Утюг

**Рис. 9.2.5.** Канонические ракурсы 12 объектов, получившие наивысший зрительский рейтинг в опыте Пальмера — Рош — Чейза (1981). В этих ракурсах объект узнавался быстрее всего (см. рис. 9.2.6).

По материалам Palmer, Rosch, & Chase, 1981.



**Рис. 9.2.6.** Влияние ракурса на время именования. Среднее время, которое понадобилось испытуемым, чтобы поименовать объекты, показанные на рис. 9.2.5.

Перспективные эффекты можно истолковать несколькими путями, и два наиболее очевидных из них — это *гипотеза частоты* и *гипотеза максимальной информации*.

#### 9.2.2.2 Гипотеза частоты

Суть ее в том, что скорость именования — это попросту функция от числа крат, которые мы видели объект с данной точки. Слабая версия, поскольку имеет место множество случаев, при которых прогнозы оказываются ошибочными. Чашки, например, мы чаще всего видим строго сверху, но рейтинг такого ракурса очень низкий, а время на идентификацию — большое. И наоборот, лошадей мы крайне редко строго сбоку (обычно спереди, сзади или даже сверху), но рейтинг такого ракурса очень высок, и очень высока скорость идентификации.

#### 9.2.2.3 Гипотеза максимальной информации

Суть гипотезы в том, что перспективные эффекты отражают количество информации о форме и функционале объекта поставляемой с разных точек наблюдения. Концепция способна учитывать то, что в оптимальном ракурсе видны несколько сторон объекта — передняя его часть, задняя, верх, низ, но не какая-то одна. Правда, сие противоречит тому факту, что в отношении некоторых предметов оптимальным является вид какой-то одной поверхности, например, сугубо фронтальный ракурс циферблата часов.

Судя по всему, в обеих гипотезах есть здоровое зерно, и перспективные эффекты, о которых говорил Пальмер, зависят одновременно от обоих факторов. Канонические ракурсы дают наблюдателю то, что можно назвать *диагностической* информацией об объекте: информацией, которая позволяет наилучшим образом отличать объект от дру-



гих объектов, поскольку виден он в самом типичном для него ракурсе. Сие позволит учесть тот факт, что фронтальный ракурс циферблата часов является каноническим, поскольку условно вся информация о часах, которая отличает их от других предметов в сцене, извлекается с их передней поверхности. Даже могут иметь место случаи, при которых определенные поверхности объекта несут *отрицательную* диагностическую информацию. Сие значит, что такая поверхность больше походит на представление некоторого другого объекта, нежели самого себя.

Однако на диагностику различных фрагментов зрительной информации об объекте также может повлиять типичность ракурса. Например, поскольку большинство из нас видит дома в основном с земли и редко сверху, ракурс с земли признается лучшим и распознается быстрее, чем, скажем, с самолета. Сие может стать обратным для птиц, которые чаще видят дома с высоты.

Сравнительно недавние эксперименты, посвященные перспективным эффектам и проведенные с новыми объектами, были поставлены так, что стали независимы от прочих факторов. С. Эдельман и Х. Бюльтхоф (1992) обнаружили эффект канонического ракурса на время распознавания новых объектов из свернутой бумаги (рис. 9.2.7), которые исходно предъявлялись испытуемым в последовательности из статических ракурсов, дававших видимый наклон в глубину. Каждый отдельный ракурс представлялся строго один раз в движущейся последовательности, что позволило исключить разного рода эффекты сближенности. Но даже в этом случае распознавание ощутимо варьировало, повторяя пальмеровы канонические эффекты (1981). Следовательно, такие эффекты — суть функция не только от сближенности.

Дальнейшие исследования показали, что определенное влияние оказывает осведомленность. Когда в исходной последовательности предъявляется небольшое множество ракурсов, последующее распознавание оказывается лучше по уже виденным ракурсам и падает по мере роста углового расстояния между тренировочными и тестовыми ракурсами, даже когда стереоскопическая и тональная информация, а также информация о движении, были заранее включены для оптимизации восприятия глубины пространства (Bülthoff & Edelman, 1992; Edelman & Bülthof, 1992). Результаты свидетельствуют о том, что испытуемые могут хранить в памяти специфические двухмерные ракурсы объектов и согласовывать новые с сохраненными посредством неких процессов, эффективность которых падает по мере роста несоответствия между новыми и сохраненными ракурсами.

Дальнейшие эксперименты показали то, что когда в тренировочной сессии используются множественные ракурсы одного и того же объекта, качество распознавания растет, но этот рост специфично зависим от соотношений между этими ракурсами

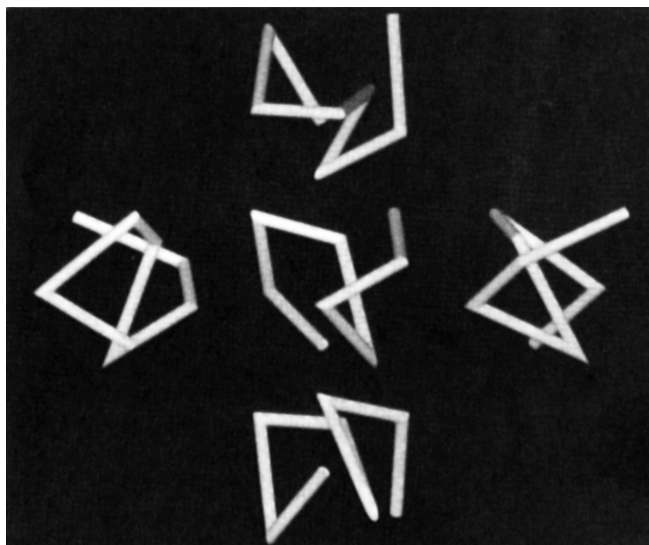
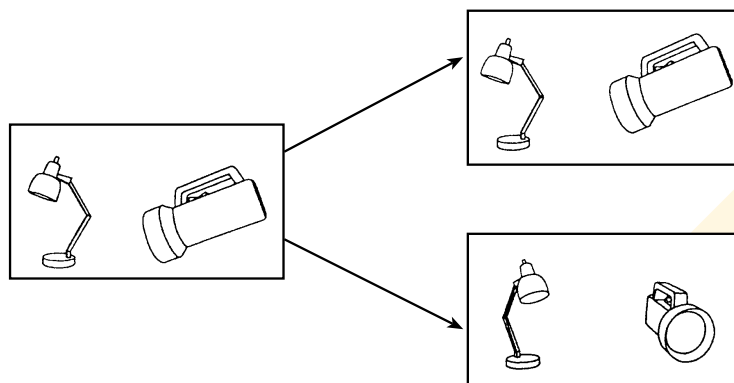


Рис. 9.2.7. Образцы стимулов, использованных в опыте по распознаванию объектов в разных ракурсах. Испытуемых просили найти виденный объект среди изображенных.

По материалам Bülthoff & Edelman, 1992.



**Рис. 9.2.8.** Стимулы в прайминг-эксперименте. Измерялась длина скрытого периода отклика на изображения объектов в первом блоке показа. Во втором блоке те же объекты предъявлялись вновь, при этом либо в том же ракурсе, либо в другом. После чего измерялась величина уменьшения времени именования относительно первого блока.

По материалам Biederman & Gerhardstein, 1993.

(Bülthoff & Edelman, 1992). В частности, если исходные ракурсы различались поворотом относительно некоей оси, качество распознавания росло сильнее в отношении тех объектов, что были повернуты вокруг той же оси, но не оси ортогональной. К примеру, если испытуемым вначале был предъявлен центральный и правый ракурсы бумажной фигурки (рис. 9.2.7), они лучше узнавали объект в левом ракурсе, нежели в верхнем или в нижнем. Сие указывает на то, что мы можем интерполировать между двухмерными ракурсами и экстраполировать за них при распознавании трехмерных объектов. Мы вернемся к этому вопросу позднее, когда поведем разговор о теориях объектной классификации, основанных на приводе и согласовании двухмерных ракурсов (Poggio & Edelman, 1990; Ullman, 1996; Ullman & Basri, 1991).

#### 9.2.2.4 Прайминг-эффект

Еще один интересный метод, но дающий несколько иные результаты, касаясь перспективных эффектов, основан на т.н. *парадигме прайминга*, предложенной Ирвином Бидерманом и его коллегами. Базовая идея, исходно озвученная британцем Дэвидом Бартрамом (1974), состоит в том, что классификация данного ретинального изображения объекта окажется точнее и пройдет быстрее, если то же самое изображение предъявить второй раз, поскольку процесс, отвечающий за эту классификацию, в состоянии повышенной готовности. Поскольку умеренно выраженный *прайминг-эффект* длится относительно долгое время — несколько часов и более — испытуемых в первом заходе просили поименовать (так быстро, как только можно) серию объектов, а затем, во втором заходе им предъявляли второй пакет объектов, некоторые из которых были взяты из первого. Цель опыта состояла в том, чтобы определить, насколько быстрее будут распознаны и поименованы повторы: разница между временем реакции в первом случае и во втором являлась мерой прайминг-эффекта.

В контексте разговора о классификации информативная ценность эксперимента заключалась в том, что повторы во втором пакете стимуляции не являлись *точной* копией объектов из первого набора, но кое-чем отличались. Например, объект оставался тем же, но иначе располагался в поле зрения, отличался ретинальным размером или зеркальностью. Особо любопытным было то, что же происходит, когда меняется ра-



курс стимуляции (рис. 9.2.8) — эффект прайминга остается или слабеет? Если эффект тот же, значит тип модификации не оказывает существенного влияния на классификацию. Если же прайминг-эффект слабеет, значит тип модификации все-таки оказывает влияние. Обе теории должны уметь объяснить, какие модификации снижают эффект прайминга, а какие нет.

Результаты опыта показали, что величина прайминг-эффекта не снижается, когда при втором показе объект меняет свое расположение или зеркально отражен (Biederman & Cooper, 1991a) или даже когда изменен его ретинальный размер (Biederman & Cooper, 1992). Однако при демонстрации одного и того же объекта в различных ракурсах обнаружилось снижение величины эффекта (Bartam, 1974). Влияние ракурса согласуется с данными Пальмера по задержке в именовании объектов (1981), а также данными Эдельмана и Бюльтхоффа (1992).<sup>1</sup>

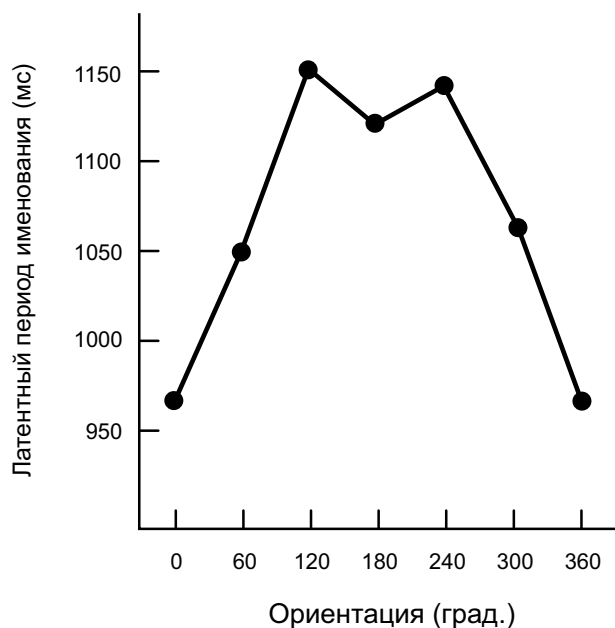
Однако сравнительно недавние исследования Бидермана и Герхардштайна (1993) показали, что при поворотах объекта в глубину существенного снижения прайминг-эффекта не происходит. Бидерман и Герхардштайн смогли показать, что когда одни и те же части объекта видимы в различных ракурсах — прайминг-эффект сохраняется. Отметим, что в опытах Пальмера разные части объектов были видны в разных ракурсах, т.е. условие видности одних и тех же частей не соблюдалось строго. Видность одной и той же части противу видности разных может объяснить то, почему в одних экспериментах перспективные эффекты имеют место, а в других — нет.<sup>2</sup>

Один из контраргументов бидермановым выводам о том, что благодаря прайминг-эффекту зрительная классификация объектов становится нечувствительной к изменениям в размере, расположении, отражении и ракурсе, звучит так, что противоречивая информация никогда не фиксируется первой. То есть, если испытуемые не обращают внимания на размер, расположение, отражение или ракурс при исходном показе, неудивительно, что изменение этих факторов не влияет на степень прайминга. Однако же Бидерман, когда проводил вторичный показ, посредством специального теста на распознающую память проверил специфическую память своих испытуемых на первый показ каждого из объектов: в контрольных условиях испытуемых просили сказать, видят ли они ту же самую картинку, что и при первом показе, или другую. Результаты указали на то, что отличия в размерах, позициях, отражениях и ракурсах были действительно восприняты, закодированы и зафиксированы в памяти. Но при этом очевидно не оказали влияния на исход классификации.

Итак, результаты опытов указывают на два варианта смены ракурса: первый влечет за собой изменения во множестве воспринимаемых частей, второй — нет. Однако прочие эффекты изменения условий просмотра не удастся объяснить в той же логике, поскольку одни и те же части видны во всех случаях. Сие справедливо не только в отноше-

<sup>1</sup> Сказанное свидетельствует о том, что величина прайминга во многом зависит от того, сосредоточено ли внимание испытуемых на первичном показе (как в исследованиях Бидермана) или нет (Stankiewicz, Hummel & Cooper, 1998). И хотя некоторый прайминг возможен, когда внимание наблюдателя не сосредоточено на объекте, все же эффект несколько сильнее и дольше, когда на объект находится в центре нашего внимания.

<sup>2</sup> Однако сие условие неприменимо к исследованиям Эдельмана и Бюльтхоффа (1992), поскольку они использовали бумажные объекты, все части которых были видимыми во всех ракурсах. Бидерман и Герхардштайн (1995) показали, что эти исследования перспективно не инвариантны, поскольку использованные в них предметы нельзя различить по их геонструктурному описанию (см. раздел 9.3.1), отличаются они лишь углом между сегментами. Интерпретация этих исследований по сей день остается предметом бурных дебатов.



**Рис. 9.2.9.** Эффекты ориентации в зрительной классификации. Испытуемые быстрее классифицируют объект в его прямом ( $0^\circ$  и  $360^\circ$ ) положении, нежели в других.

По материалам Jolicoeur, 1985.

Долгое время большинство исследователей полагали, что этих эффектов нет, но уверенность сия основывалась скорее на интуиции, нежели на четко поставленных экспериментах. Исследования канадского психолога Пьера Жоликера (1985) показали, что испытуемые действительно быстрее классифицируют изображения прямо стоящих объектов, т.е. в их нормальной ориентации, нежели когда последние повернуты в той или иной мере вокруг зрительной оси. Латентный период именования растет по мере отклонения от прямого стояния, и сие заставляет нас думать, что испытуемые мысленно поворачивают предмет в прямое положение, прежде чем распознают его (см. гл. 12). На рис. 9.2.9 показано влияние ориентации каждого из объектов при первой экспозиции в опыте Жоликера. Однако при повторном показе эффект ориентации существенно снижается.

Дальнейшие эксперименты, проведенные уже американцами Майклом Тарром и Стивеном Пинкером (1989, 1990, 1991), были посвящены ответу на вопрос, почему в реальной практике эффекты ориентации существенно ниже, чем в лабораторных условиях. Исследователи использовали новые объекты, ориентацией которых можно было легко управлять. Когда испытуемому предъявляли новую фигуру в определенной ориентации, время именования тестового объекта в этой ориентации было минимальным и росло по мере углового отклонения от нее, то есть так же, как и в опыте Жоликера (1985). Однако когда испытуемым активно показывали объекты в различных ориентациях, латентный период именования снижался по *всем* виденным ориентациям, а время отклика по остальным ориентациям в целом росло по мере отклонения от знакомой. Результаты опыта свидетельствуют о том, что мы храним в памяти множественные представления одного и того же объекта в различных его ориентациях, но не в ориентации канонической или же в каком-то ориентационно инвариантном представ-

нии бумажных объектов (Bülthoff & Edelman, 1992), но также в отношении эффектов ориентации, о которых речь пойдет ниже.

### 9.2.2.5 Эффекты ориентации

Изменения ракурса могут вызываться либо сменой точки наблюдения стационарного объекта, либо вращением объекта относительно неподвижного наблюдателя. Поворот объекта в глубину пространства часто влечет за собой изменение видности частей этого объекта, но поворот объекта вокруг зрительной оси — нет. Ровно те же части и пространственные отношения между ними будут представлены при всех поворотах в картинной плоскости. То есть, вопрос состоит в том, действительно ли изменения в ориентации объекта систематически влияют на объектную классификацию.

лении. Мы вернемся к этому разговору позднее, когда начнем разбирать видоспецифические теории классификации.

Несмотря на то, что изначально зрительная классификация объектов смотрелась как ориентационно- и ракурсонезависимый процесс, влияние этих переменных на самом деле стало весьма сложным и интересным. Отметим, что результаты описанных опытов весьма важны, ибо дают нам ключевые данные о том, каковы процессы перцептивной классификации. К примеру, стало окончательно ясно, что процессы эти ориентационно и ракурсно неинвариантны. Подробнее мы сей вопрос рассмотрим в разделе 9.3, когда будем говорить о теориях объектной классификации.

### 9.2.3 Дольчатость

В гл. 7 мы рассмотрели вопрос о том, действительно ли объекты структурируются по частям и, если да, то как эти части определить? Одним из аргументов служит тот факт, что большинство известных нам объектов исходно обладают очевидной дольчатостью: у человеческого тела имеются голова, руки, ноги и торс; у столов — столешница и ножки; у самолета — фюзеляж, два основных крыла, стабилизаторы и хвост. Ключевым вопросом теорий объектной классификации является вопрос о том, действительно ли части эти играют существенную роль в зрительной классификации объектов по категориям начального уровня?

Масштабные исследования в этой сфере выполнили (на основе обсуждавшейся ранее парадигмы прайминга) американцы Ирвин Бидерман и Эрик Купер (1991 b). Исследователи воспользовались некоей хитроумной манипуляцией, при которой часть контура объекта удалялась из показа, а визуальная значимость частей оценивалась путем манипулирования отношениями между удаленными фрагментами в первом и втором показах (см. ниже).

В первом опыте Бидерман и Купер построили парные варианты контурных объектов, т.е., удалив половину контура в каждом из вариантов. Обе половины так дополнили друг друга, что, ежели совместить одну с другой, воссоздавался исходный контур объекта (рис. 9.2.10 В). В первом показе испытуемым предъявляли некое множество картинок, на каждой из которых был изображен полуконтур объекта; во втором — предъявлялось второе множество контуров, которое определенным образом соотносилось с первым. Интерес представляло то, насколько предварительный показ того же самого (или иного) контура облегчит (то есть, *праймирует*) классификацию объекта при втором показе.

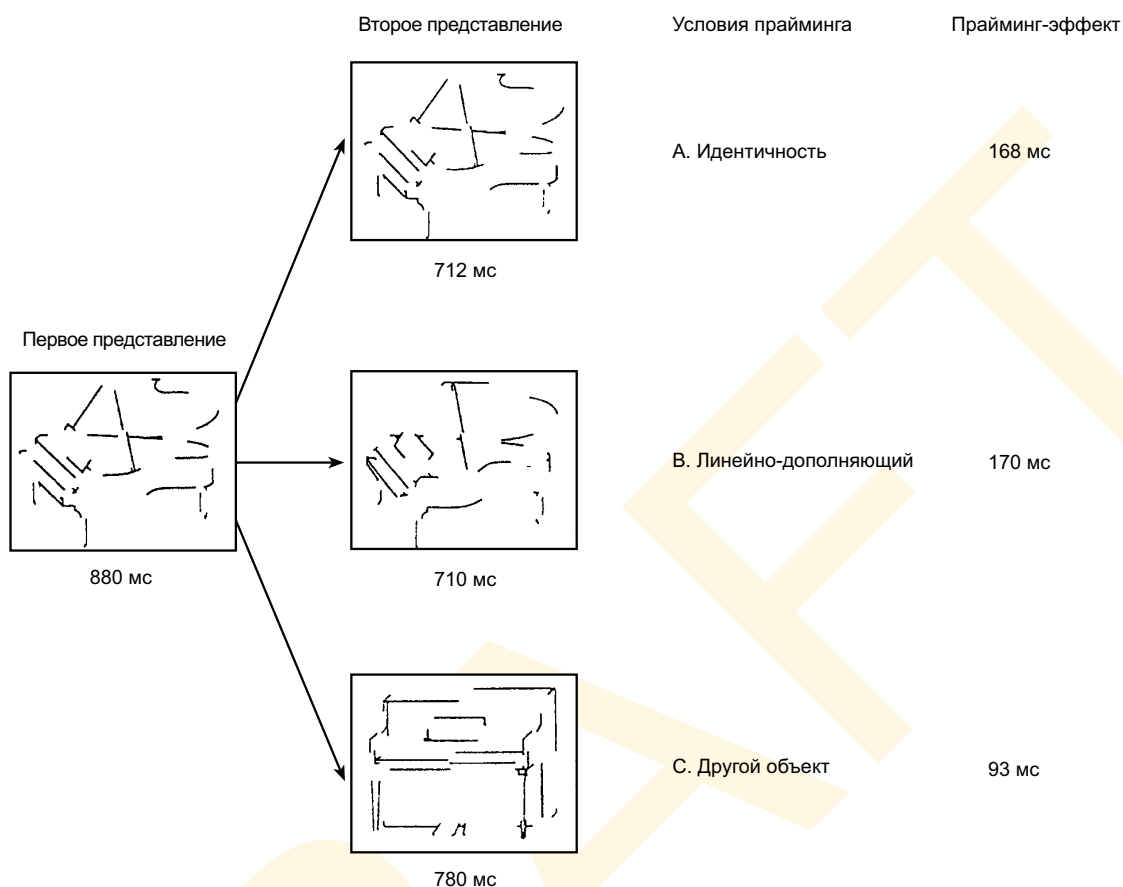
Изучались три варианта условий просмотра:

1. *Тождественный прайминг*. Предъявлялся один и тот же контур одного и того же объекта (рис. 9.2.10 А). Сие определяло максимальную величину возможного прайминга при данном задании.

2. *Линейно-дополняющий прайминг*. Предъявлялся тот же самый объект, но виден был лишь тот контур, которого не хватало при первом показе (рис. 9.2.10 В).

3. *Инообъектный прайминг*. В втором показе предъявлялся совершенно другой объект, в совершенно ином ракурсе, но из той же начальной категории. К примеру, в первом наборе картинок предъявлялся контур рояля, а во втором — пианино (рис. 9.2.10 С); в первом показе — слон, вид сбоку; во втором — тот же самый слон, но фронтально.

Бидерман и Купер убедительно показали, что в линейно-дополняющем варианте сегментов вполне достаточно для идентификации всех составляющих объекта. Коли так, сие условие является эффективной мерой прайминга через компоненты объекта, мерой, незагрязненной праймингом через специфические признаки изображения



**Рис. 9.2.10.** Опыт по исследованию линейно-дополняющего прайминга. В первом блоке латентный период именованья измерялся по изображениям, содержащим половину контуров. Во втором блоке предъявлялись те же изображения с тем же набором линий (А), дополняющим набором (В), а также изображение другого объекта (С).

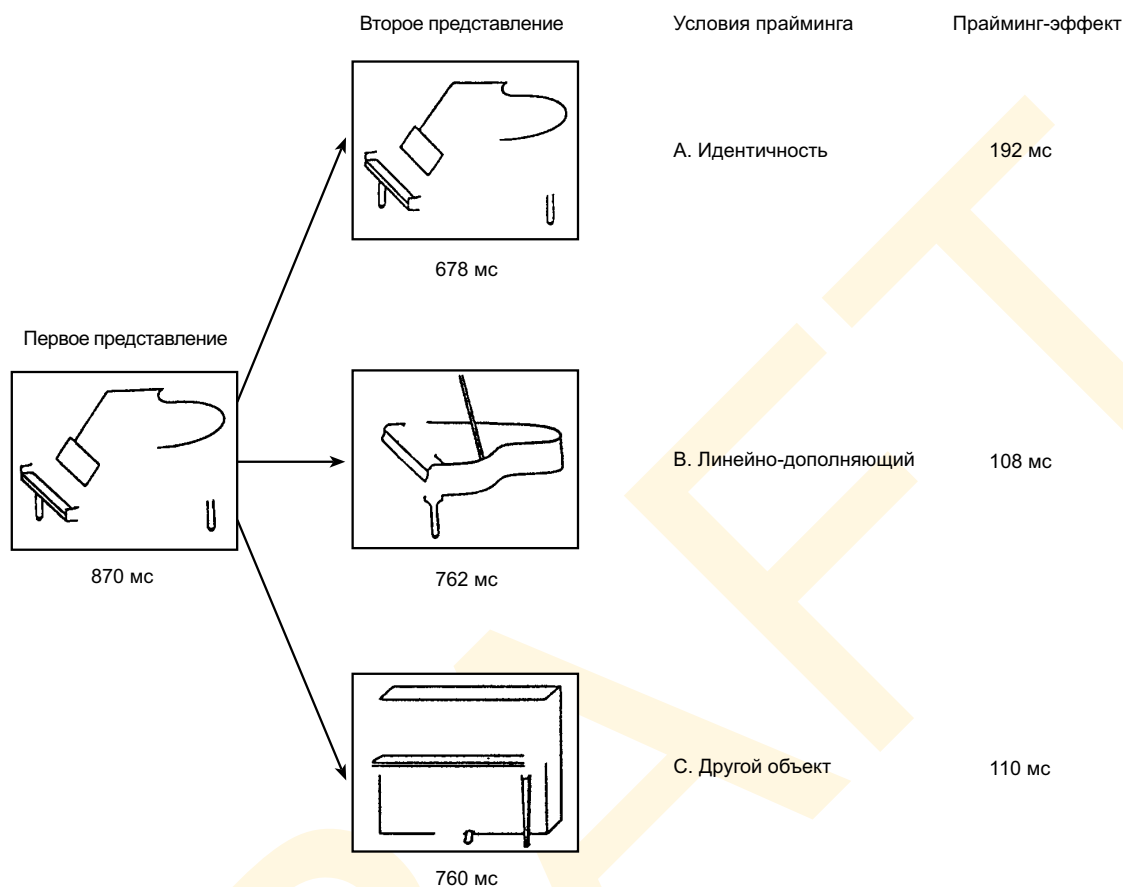
По материалам Biederman & Cooper, 1992 b.

в контуре. Им также удалось показать, что инообъектный вариант определяет величину прайминга, которой мы ждем от общепрактических эффектов и от повтора названия категории.

Результат эксперимента окончательно проясняется, если оценить время реакции (рис. 9.2.10). Повтор частей (без повтора какого-либо отдельного сегмента) в условиях линейно-дополняющего прайминга дает ровно ту же величину прайминга как повтор идентичных сегментов в идентичных условиях, то есть порядка 170 мс. Однако повтор лишь по той же категории дает много меньший прайминг — порядка 93 мс. В итоге, Бидерман и Купер пришли к выводу, что объектная классификация, базируется скорее на восприятии частей, нежели отдельных отрезков и вертексов, образующих эти части.

Однако же сие рассуждение спорно. Жесткий прайминг в линейно-дополняющих условиях может стать следствием того, что во втором показе предъявлялся тот же объект и в том же ракурсе, но не того, что восприняты были те же части объекта. То есть, прайминг мог явиться на уровне ракурса цельного объекта, нежели на уровне его частей. Ключевым является вопрос о том, что же произойдет, если один и тот же объект, в том же самом ракурсе предъявят во втором блоке, но другими частями?

Ответ дал второй эксперимент, в котором удалили половину частей (рис. 9.2.11), но



**Рис. 9.2.11.** Опыт по исследованию часте-дополняющего прайминга. В первом блоке латентный период именования измерялся по изображениям, содержащим половину частей. Во втором блоке были предъявлены те же изображения с тем же набором частей (А), дополняющим набором (В) а также изображение другого объекта (С).

По материалам Biederman & Cooper, 1992 b.

не половину линий. В данном опыте условие *частедополняющего прайминга* не повторяло части, но только *ракурс* одного и того же объекта. Следовательно, ежели сие дает прайминг, результаты второго эксперимента должны оказаться теми же. Однако, если же это прайминг от повтора составляющих объект частей, то условие частедополняющего прайминга должно давать много меньший эффект, чем условие идентичности, поскольку отсутствует повтор частей.

И вновь результат очевиден, что читается из времен реакции (рис. 9.2.11): частедополняющий прайминг изображений с удаленными частями дает на выходе много меньший прайминг-эффект, чем прайминг идентичности, фактически — примерно ту же степень снижения времени реакции, что и условия инообъектного прайминга.

Результаты обоих опытов вкупе убедительно доказывают, что восприятие частей играет ведущую роль в зрительной классификации объектов.

Ниже мы опишем бидерманову теорию объектной классификации — теорию, основанную на предположении, что объекты классифицируются через восприятие их частей и пространственные взаимоотношения между ними.



### 9.2.4 Контекстные эффекты

До сего момента все рассматриваемые нами феномены зрительной классификации касались непосредственно объекта: насколько он типичен в категории начального уровня, каков ракурс его просмотра, его размер, расположение, ориентация и дольчатость. Но перечисленными факторами влияние на объектную классификацию не исчерпывается, т.к. существуют еще т.н. *контекстные эффекты* — изменения в характере классификации под действием пространственного массива предметного окружения целевого объекта.

Один из примеров этого феномена дан на рис. 9.2.12, где перцептивная идентичность срединных букв в каждом слове, строго зависима от букв соседних. Каждый из нас прочтет рисунок как «THE CAT», и никому в голову не придет увидеть «TAE CHT», «THE CHT» или «TAE CAT», даже несмотря на то, что стимул стыкуется со всеми четырьмя интерпретациями; центральные буквы обоих слов фактически идентичны и неопределенны — наполовину «Н», наполовину «А». Но первую мы читаем как «Н» в контексте «Т\_Е», а вторую как «А» в контексте «С\_Т». Феномен столь жесткий, что никому в голову не придет заподозрить написанное в двусмысленности, даже несмотря на полную идентичность физических форм этих «букв».

На рис. 9.2.13 дан еще один пример влияния контекста на зрительную классификацию (Palmer, 1975b), иллюстрирующий еще и то, что имеет место компромисс между величиной части-структурной детали, необходимой для классификации объекта

Т А Е С А Т

**Рис. 9.2.12.** Контекстное определение двусмысленного стимула. Перцептивные «Н» и «А» физически идентичны, но воспринимаются по-разному в разных контекстах.



**Рис. 9.2.13.** Иллюстрация контекста часть/целое. Признаки лица, легко узнаваемые в профиль (A) нераспознаваемы вне этого контекста (B). Однако когда внутренняя структура признаков вербализована, они становятся легко узнаваемыми (C).

По материалам Palmer, 1975 b.

и величиной создаваемого ею контекста. Тот или иной объект, например, нос, глаз, рот можно изобразить его общей формой, когда он зрим в общем контексте. Однако, когда мы видим его отдельно, то для достижения того же уровня классификации собственную долевую структуру объекта следует описывать полнее и точнее. Судя по всему, зрительной системе для классификации объекта нужны два уровня его дольчатости, и либо один нижний (части), либо один верхний (контекст) окажется достаточным.

Возможно, наиболее эффектной демонстрацией сказанному явится один из «фруктовых портретов» кисти Джузеппе Арчимбольдо<sup>1</sup> (рис. 9.2.14), в котором специфические формы «лицевых признаков» можно с легкостью идентифицировать как фрукты. Поскольку их пространственное расположение соответствует таковому у глаз, носа, и рта человека, общая конфигурация лица воспринимается очень легко.

Действительно, сие доминирующий перцепт, о чем свидетельствует тот факт, что мы воспринимаем это стимул как необычную рожицу, но отнюдь не как странноватую комбинацию фруктов. Однако наш анализ до сих пор основывался на чисто субъективных данных. Существуют ли данные объективные, подтверждающие то, что визуальный контекст действительно влияет на зрительную классификацию объектов?

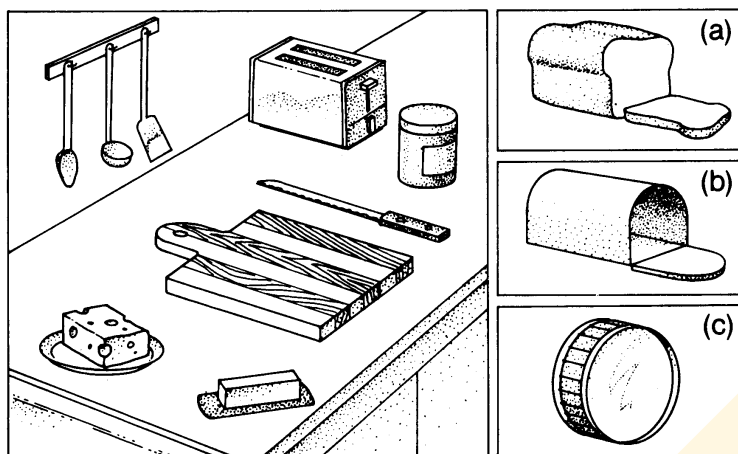
В нескольких весьма ловко поставленных экспериментах влияние контекста на объектную классификацию удалось задокументировать. Как и следовало ожидать, опыты показали, что подходящий контекст облегчает классификацию, а не подходящий — затрудняет. В одном из исследований Пальмер (1975а), после показа некоей контекстной сцены (рис. 9.2.15), ненадолго предъявлял своим испытуемым контуры цельных объектов. Затем контекстные отношения между сценой и целевым объектом менялись. В варианте кухни, к примеру, последовательно предъявляемые объекты могли либо *подходить* этой сцене (буханка хлеба), либо *не подходить* (барабан), либо *обманывать* наблюдателя в том смысле, что целевой объект оказывался зрительно схож с подходящим объектом (почтовый ящик). В контрольных — неконтекстных — условиях объекты предъявлялись вслед за белым полем (а не за сценой, несущей контекст). Через хитроумные комбинации объектов и сцен, все имевшиеся объекты были предъявлены во всех возможных контекстных условиях.

Результаты опыта даны на рис. 9.2.16: мы видим, что подходящие контексты облегчают классификацию относительно неконтекстных контрольных условий, и что



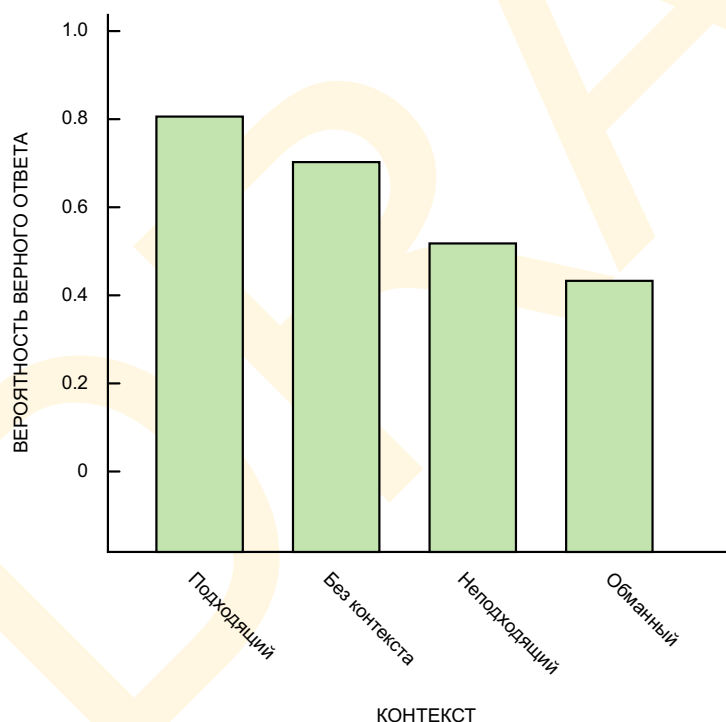
Рис. 9.2.14. «Фруктовый портрет» работы Арчимбольдо. Конфигуративная информация позволяет увидеть лицо, несмотря на то, что составляющие его признаки — фрукты.

<sup>1</sup> Джузеппе Арчимбольдо (итал. Giuseppe Arcimboldo), — итальянский живописец чинквеченто, декоратор, представитель маньеризма. В его творчестве усматривают предвосхищение сюрреализма. — Прим. пер.



**Рис. 9.2.15.** Стимулы из опыта по изучению влияния контекста на зрительную классификацию. Перед показом тестового стимула на короткое время наблюдателю предъявлялась контекстная сцена. Стимул либо подходил контексту сцены (а), либо был зрительно подобен стимулу «а» (b), либо совершенно не подходил сцене (с).

По материалам Palmer, 1975a.



**Рис. 9.2.16.** Результаты опыта по изучению влияния контекста на зрительную классификацию. Вероятность правильной классификации целевого стимула оказалась наивысшей в подходящем контексте и минимальной в неподходящем и обманном контекстах.

По материалам Palmer, 1975a.

неподходящие — затрудняют ее. Слабейшим результатом классификации оказался в обманном контексте, в котором испытуемые давали объекту название *похожего* объекта, подходящего к данной сцене. Итак, мы видим, что на точность распознавания может ощутимо повлиять характер объектов окружения. Сие крайне важный факт, который добротная теория объектной классификации обязательно должна принимать во внимание.

Бидерман при изучении контекстных эффектов использовал иной метод (1972; Biederman, Glass & Stacy, 1973): он просил своих испытуемых найти целевой объект в определенной сцене и замерял время реакции. В своем исходном исследовании ученый манипулировал контекстом, предъявляя испытуемым либо обычную фотографию сцены, либо ее случайным образом реорганизованный вариант. Как и следовало ожидать, наблюдатели заметно дольше искали нужный объект в изображении дезорганизованной сцены. В дальнейших опытах Бидерман исследовал несколько иные варианты нарушения отношений между целевым объектом и окружающим его контекстом. Изучавшиеся факторы включали то, насколько объекты были подходящими к данной сцене: была ли их позиция адекватной (например, водokolонка на

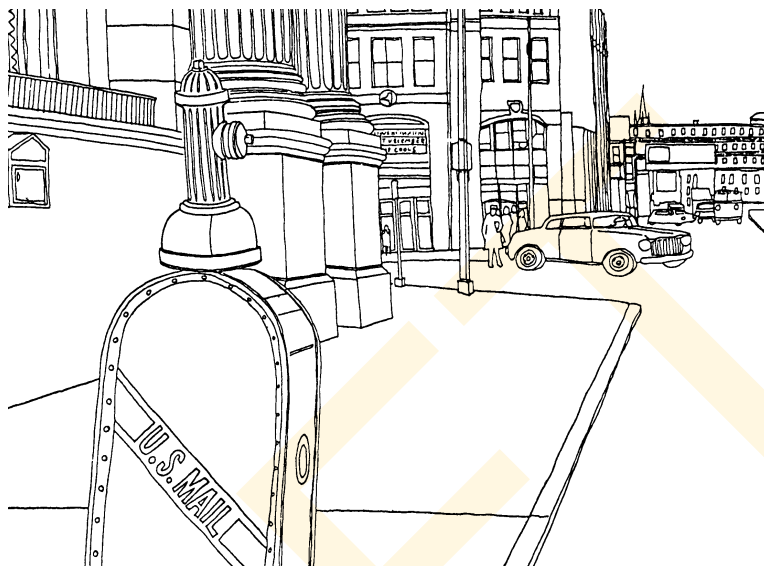
боковой дорожке) или нет (например, водоклонка на крышке почтового ящика — рис. 9.2.17); был ли объект крупнее или меньше нормального в сравнении с другими объектами сцены; получал ли объект зрительную поддержку или нет; был ли он частично закрыт или выглядел полупрозрачным. Ученый обнаружил, что любой фактор, нарушающий контекстные отношения, затрудняет детекцию целевого объекта.

Итак, наличие контекстных эффектов свидетельствует о том, что зрительной классификации отношения между объектами сцены и сложны, и важны. Очевидно, мы способны верно классифицировать объекты даже в необычных контекстах. Водоклонка

на крышке почтового ящика может распознаваться дольше, чем в норме, но все равно мы узнаём ее рано или поздно. То есть, мы говорим, что контекст влияет скорее на *эффективность* классификации: когда имеются подходящие взаимоотношения между объектами зрительная система очевидно способна использовать контекстную информацию для облегчения своей работы. В каждом случае «нормальная» ситуация обрабатывается быстро и с минимум погрешности, тогда как «ненормальная» замедляет этот процесс и дает большую ошибку. Поскольку «нормальная» ситуация по определению встречается намного чаще «ненормальной», контекстные эффекты в целом полезны организму в его родной среде, реализуя эволюционно ценную функцию.

### 9.2.5 Зрительная агнозия

Принципиально иным феноменом, ограничивающим действие теорий объектной классификации является т.н. *зрительная агнозия* — перцептивная недостаточность, возникающая при травмах головного мозга, и при которой больной не может корректно классифицировать большинство ранее хорошо знакомых объектов.<sup>1</sup> Существует множество различных форм зрительной агнозии, и взаимоотношения между ними до сих пор толком не изучены. Некоторые симптомы выглядят так, что нарушение возникло на поздних стадиях зрительной обработки (т.н. *апперцептивная агнозия Лиссауэра*, 1890/1988). Больные при этом неспособны к распознаванию объектов, поскольку соб-

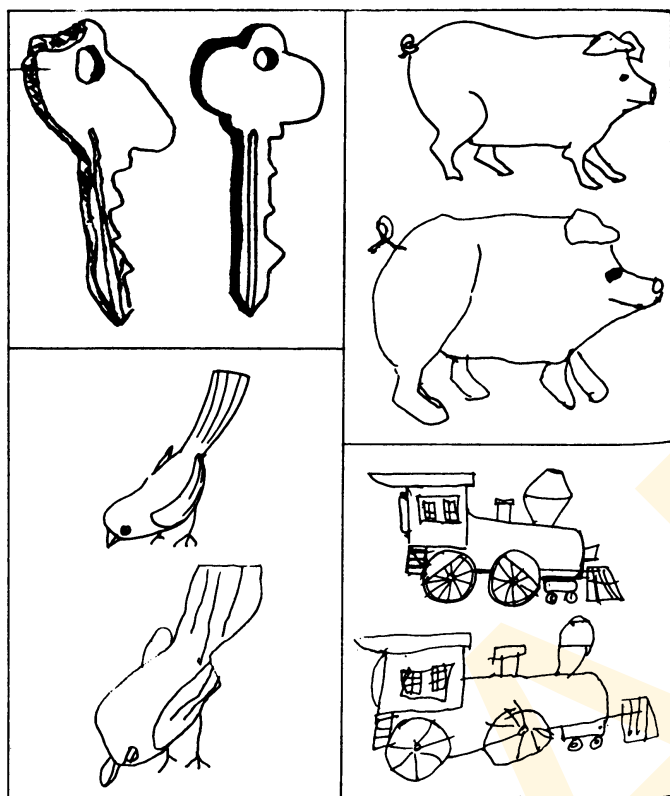


**Рис. 9.2.17.** Пример стимула, использованного в опыте по изучению влияния контекста на поиск объекта. В данном случае пожарный гидрант находится в необычном положении, увеличивая тем самым время его поиска в сцене. Увеличение времени поиска мы наблюдаем также, когда появление целевого объекта в сцене маловероятно, когда он имеет необычный размер или полупрозрачен.

По материалам Biederman, 1981.

<sup>1</sup> «Агнозия» — греческое слово, означающее «неизвестно».





**Рис. 9.2.18.** Рисунки, сделанные больными, страдающими зрительной агнозией. Несмотря на неспособность к именованию объектов, агнозики вполне способны к перерисовке (копированию) изображений, что полностью исключает фактор сенсорной недостаточности в генезисе состояния.

По материалам Rubens & Benson, 1971.

мог достать из шкафа пиджак, либо брюки. Словом, он не мог классифицировать простейшие повседневные вещи.

Пациента активно обследовали, стараясь найти причину патологии. Определенные тесты подтвердили, что имеют место нарушения именно в распознавании объектов:

1. *Сенсорная недостаточность* — отсутствовала. Состояние больного не сенсорного генеза — острота зрения (и прочие базовые его функции) — в норме. Пациент даже мог рисовать разные предметы, но без их идентификации (см. рис. 9.2.18).

2. *Когнитивная недостаточность* — отсутствовала. Состояние больного не относимо к когнитивной патологии, поскольку он мог определять объекты вербально и вполне адекватно рассуждал об их структуре.

3. *Работа прочих сенсорных систем* — в норме. Проблема состояла отнюдь не в том, что больной вообще не мог классифицировать объекты — например, он вполне адекватно определял их на ощупь.

4. *Вербальная недостаточность* — отсутствовала. Проблема относима была не просто к неспособности поименовать объект, то есть *классифицировать* его. Больной вел себя неадекватно в отношении объектов, которые не смог назвать; не мог также и жестами продемонстрировать их верное употребление.

ственно видят их ненормально. У других больных сама перцепция в норме, но они не могут *идентифицировать* увиденное — условие, названное немецким нейрофизиологом XIX века Генрихом Лиссауэром *ассоциативной агнозией*. Знаменитый американский исследователь Ганс-Лукас Тойбер (1968) описал эти состояния как задействующие «нормальную перцепцию с обнаженным смыслом» из-за неспособности к адекватной классификации.

Случай, который мы назовем «GL», — это отличный пример ассоциативной агнозии (Ellis & Young, 1988). 80-летний мужчина сильно простудил голову, после чего стал значительно хуже видеть. Состояние проявляло себя отнюдь не слепотой и даже не снижением основных зрительных функций (восприятие основных физических свойств объектов осталось в норме), но утратой способности к идентификации хорошо знакомых предметов, назначение которых здоровому человеку легко понятно. Больной, например, не



Отбросив альтернативные толкования, приходим к неизбежному выводу, что больной GL не мог именно *классифицировать* зримые им объекты. В сравнении со многими другими, при которых у больных проявляются и прочие виды перцептивной недостаточности, случай, как говорят, «чистый». Однако наличие даже одного «чистого» случая является важным доказательством того тезиса, что классификация объектов есть отдельный психологический процесс, недостаточность которого может проявить себя самостоятельно.

Больные, страдающие зрительной агнозией, выдают пеструю картину симптомов, и у некоторых имеется недостаточность в классификации объектов определенного вида или с определенными свойствами. Один из классических примеров т.н. *прозопагнозия* — неспособность к зрительной идентификации лиц людей. Прозопагностик может детально описать лицо человека, но не способен идентифицировать самую его персону, даже когда перед ним собственный супруг или ребенок. Такие больные обычно реагируют на всех, как на незнакомцев — до тех пор, пока человек не начнет говорить и пока больной не узнает его голос. Насколько жесткой может стать такая форма агнозии, читаем у знаменитого нейропсихолога К. Паллиса (1955): «В клубе я встретил незнакомца и спросил дворецкого, кто это. Он рассмеялся надо мной: «Это Вы, сударь, в зеркале!»

Исследовались и другие больные агнозией, страдавшие нарушениями классификации объектов, в частности — одушевленных. Больной JBR, к примеру, мог идентифицировать 90% изображений неодушевленных предметов, но только 6% растений и животных (Warrington & Shallice, 1984). Имеются данные даже о еще более тяжелых нарушениях, касаемых частей тела, домашних предметов, фруктов, овощей, хотя некоторые из этих видов недостаточности все же скорее когнитивные, нежели перцептивные (Farah, 1990).

Одна из проблем общих многим агнозикам и изученная экспериментально — это их специфическая неспособность классифицировать объекты, предъявленные в «необычном» ракурсе. Е. Варрингтон и А. Тейлор (США, 1973, 1978) обнаружили, что многие агнозики, которые способны классифицировать изображения большинства объектов, снятых в обычных ракурсах, не справляются с задачей, если ракурс необычен (см.



**Рис. 9.2.19.** Непривычный ракурс. Больные, страдающие агнозией, порой верно распознают предметы, показанные в стандартных ракурсах (вверху), но редко справляются с ракурсами непривычными (внизу).

По материалам Goldstein, 1989.

рис. 9.2.19). Сей агнозический феномен поразительно схож с перспективными эффектами, фиксируемыми у здоровых людей (Palmer, Rosch & Chase, 1981), за исключением того, что вместо удлинения латентной фазы распознавания, больные оказываются вовсе неспособными идентифицировать объект, причем даже в комфортных условиях просмотра. У этих феноменов вполне может стать общий корень: агнозики могут попросту не мочь преодолеть сложности, предъявляемые простым ракурсом, тогда как для здоровых людей — это вполне посильная преграда.

Ракурсы, более всего осложняющие жизнь агнозикам — это такие ракурсы, при которых длинник объекта укорочен. Сие натолкнуло некоторых исследователей на мысль, что этим больным сложно определить объектно-центрированный репер предмета (Marr, 1982). Выбор ложного репера ведет к сбою в структурном описании, что, в свою очередь, блокирует правильную классификацию. Замечено, что у больных сим специфическим недугом обычно поражена право-задняя коры теменной доли (Ellis & Young, 1988).

Существует множество иных зрительных расстройств, возникающих в результате поражения головного мозга и относимых к зрительной агнозии, которые дают множество странных симптомов, вызванных, в свою очередь, разнообразной и малоизученной церебральной патологией. Необычайно интересные истории болезней и феноменологические описания этих состояний предлагает британский нейрофизиолог Оливер Сакс (1985) в своей знаменитой «The Man Who Mistook His Wife for a Hat». Также весьма любопытны данные М. Фараха (1990).

### 9.3 ТЕОРИИ ЗРИТЕЛЬНОЙ КЛАССИФИКАЦИИ

В предыдущем разделе мы описали ряд наиболее значимых феноменов зрительной классификации объектов. Теперь же нам предстоит попытка объяснить то, каким образом идет классификация объектов зрительной системой человека, и как сии факты можно толковать в логике количественной теории зрительного распознавания и классификации.

Возможно, за последние несколько десятков лет единственным внушительным теоретическим подходом к зрительной классификации, явилась идея структурного описания в представлении форм объектов (см. гл. 8). Мысль развивалась усилиями математиков, а также количественно ориентированных американских психологов, включая Т. Бинфорда (1971), И. Бидермана (1987), Д. Марра с Нишихарой (1978), а также Ст. Пальмера (1975а). Действительно, имеется аспект, в котором концепция четырехэтапного зрительного акта (оптический этап, фасетный, предметный и категорийный) разделена на удобную для понимания последовательность стадий, исходом которых является пригодное к классификации волюметрическое структурное описание трехмерных объектов.

Из всех теорий объектной классификации, развивавшихся в лоне данного концепта, мы лишь одну рассмотрим детально — теорию Ирвина Бидермана (1985, 1987) или, как ее еще называют — теорию «расознавания через компоненты». Теория эта не сильно отличается от прочих, но в чем-то она проще и разработана с особым вниманием к экспериментальным данным. Словом, является скорее представителем «корректных» моделей, но отнюдь не «лучших». По завершении разговора о ней, мы рассмотрим и альтернативные подходы к вопросу зрительного распознавания объектов, подходы, основанные на согласовании двумерных ретинальных проекций, но не трехмерных структурных описаний (Koenderink & van Doorn, 1979; Poggio & Edelman, 1990; Ullman, 1996; Ullman & Barsi, 1991).

### 9.3.1 Теория Бидермана

Теория Бидермана или теория распознавания через компоненты (RBC — *Recognition By Components*), именуемая также теорией геонов, старается объединить описание множества зрительных процессов в единую, психологически полноценную концепцию того, как наше зрение классифицирует объекты, относя их в категорию начального уровня. Теория основана на той идее, что объект можно представить в виде массива примитивных волюметрических компонент, которые Бидерман (1985, 1987) окрестил *геонами*. Объектная классификация следует путем согласования геонного описания целевого объекта с соответствующим геонным описанием категории объектов.

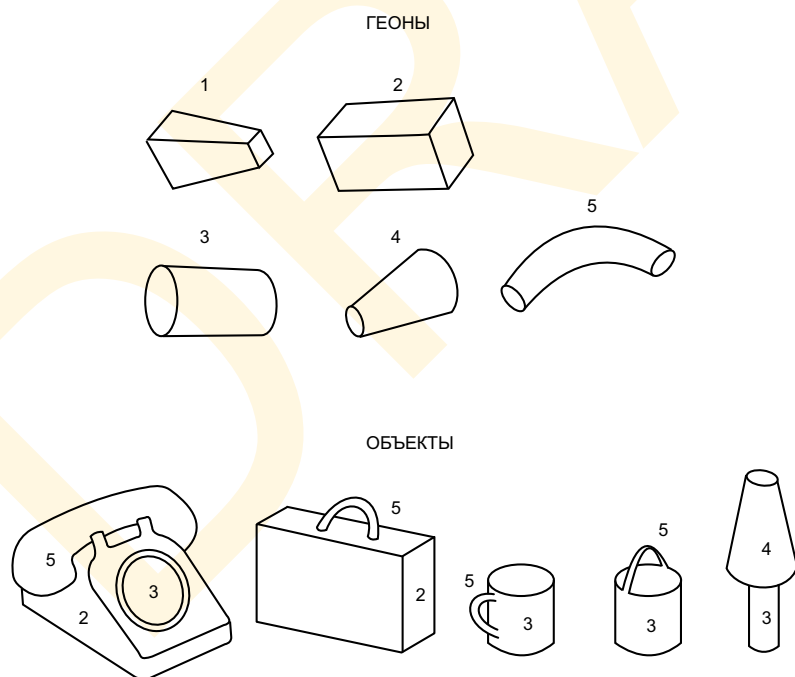
#### 9.3.1.1 Геоны

Первый базовый посыл RBC-теории таков, что сохраненные в памяти представления категории и объекта — суть волюметрические структурные описания. Сие значит, что RBC-представления — это исключительно иерархические сети, чьи узелки соответствуют трехмерным частям и чьи связи соответствуют взаимоотношениям между этими частями. Бидерман окрестил трехмерные примитивы *геонами*, что есть сокращение от *geometric ions*. Геоны — это типизированные цилиндры, чьи по сути непрерывные параметры разделены на несколько дискретных диапазонов, очень легко различимых в большинстве ракурсов. Сие дает относительно небольшое множество отдельных примитивов, из которых можно построить неимоверное число объектных представлений (подобно тому, как из ограниченного числа букв в алфавите можно построить огромное число слов и фраз). Небольшой набор геонов, а также легко распознаваемых объ-

ектов, из них собранных, дан в качестве примера на рис. 9.3.1.

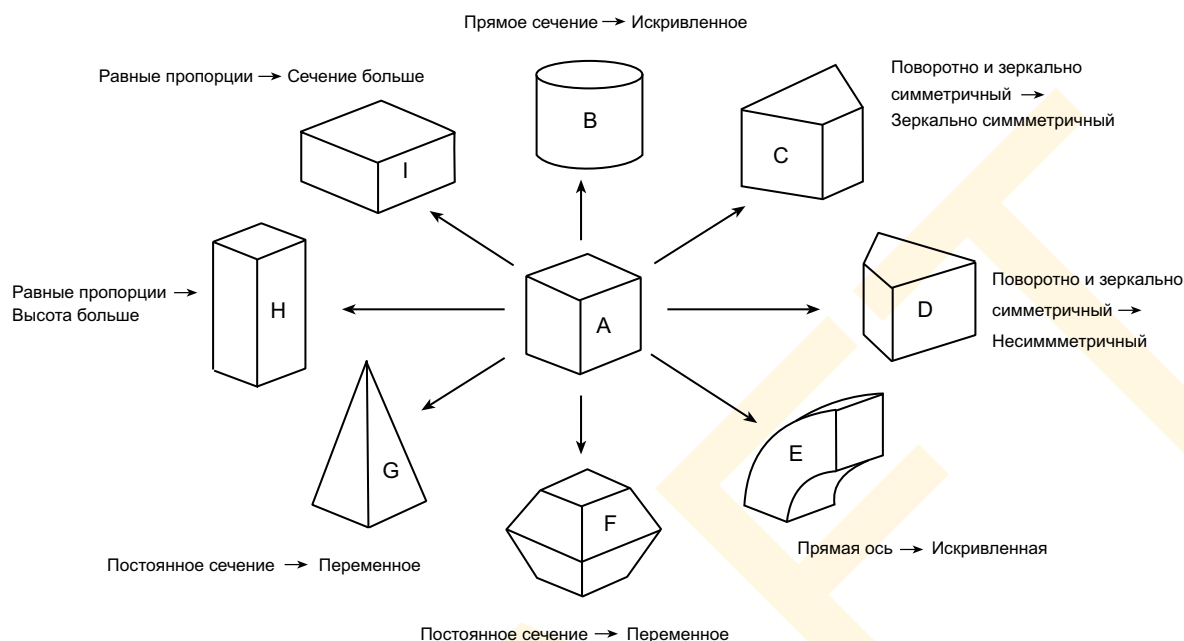
Вспомним (см. гл. 8), что типизированные цилиндры — это объемные фигуры, построенные протягиванием двухмерных форм по оси (см. рис. 8.2.17). Разнообразие форм достигается варьированием сечения двухмерной основы цилиндров и формы оси протягивания, а также режимом изменения площади сечения и отношением длины оси протягивания к сечению. Бидерман позиционировал множество из 36 качественно различных геонов, задав следующие отличия в них:

1. *Кривизна сечения*. Сечение геона может быть



**Рис. 9.3.1.** Примеры геонов и объектов из них состоящих. Пять простых геонов даны в составе нескольких объектов.

По материалам Biederman, 1987 г.



**Рис. 9.3.2.** Иллюстрация пяти переменных в конструировании типизированных цилиндров. Куб в центре (А) можно видоизменять, получая восемь новых геон, за счет изменения одного из параметров: кривизны краев сечения (В), симметрии сечения (С и D), кривизны оси протягивания (Е), формы сечения (F и G) и пропорций (H и I).

либо *прямым* (например, при квадратной базе — рис. 9.3.2 А), либо *искривленным* (как при круглой базе — рис. 9.3.2 В).

2. *Симметрия*. Сечение может быть либо *асимметричным* (рис. 9.3.2 D), либо выдавать два типа симметрии — собственно *зеркальную* (рис. 9.3.2 С), либо *зеркальную и поворотную* (рис. 9.3.2 А).

3. *Кривизна оси*. Ось протягивания может быть либо *прямой* (рис. 9.3.2 А), либо *искривленной* (рис. 9.3.2 Е).

4. *Размер сечения*. Размер сечения, протягиваемого по оси, может быть либо *постоянным*, давая параллельность сторон геона (рис. 9.3.2 А), *расширяющимся и сжимающимся* с исходом в непараллельность сторон с точкой максимальной выпуклости (рис. 9.3.2 F), либо только *расширяющимся (сжимающимися)*, давая непараллельность сторон, сходящихся к концу (рис. 9.3.2 G).

Характер каждой из этих переменных демонстрирует нам то, как единичный геон (куб в центре) можно превратить в другой геон, изменив значение какого-либо из признаков.

Смысл в выделении четырех видов отличий между геонами состоит в том, что отличия эти качественны (а не только количественны), что есть результат качественных отличий в ретинальных проекциях. Следовательно, признаки ретинального изображения, характеризующие различные геоны, относительно нечувствительны к изменениям в ракурсе наблюдения. К примеру, прямые линии в сцене проектируются на сетчатку прямыми с любого ракурса, а кривые — кривыми с большинства точек (кривая может спроектироваться в прямую, если целиком окажется в плоскости, на которой лежит зрительная ось). Благодаря качественному характеру отличий между 36-ю количественно различными геонами (2 значения кривизны сечения × 3 значения симме-



трии  $\times 3$  значения переменных размера  $\times 2$  значения кривизны оси = 36) можно сравнительно легко пренебречь отличиями в ракурсе.

Бидерман предложил еще несколько признаков геонов, один из которых дает количественно (не качественно) отличные геоны, которые он разделил по трем различным величинам:

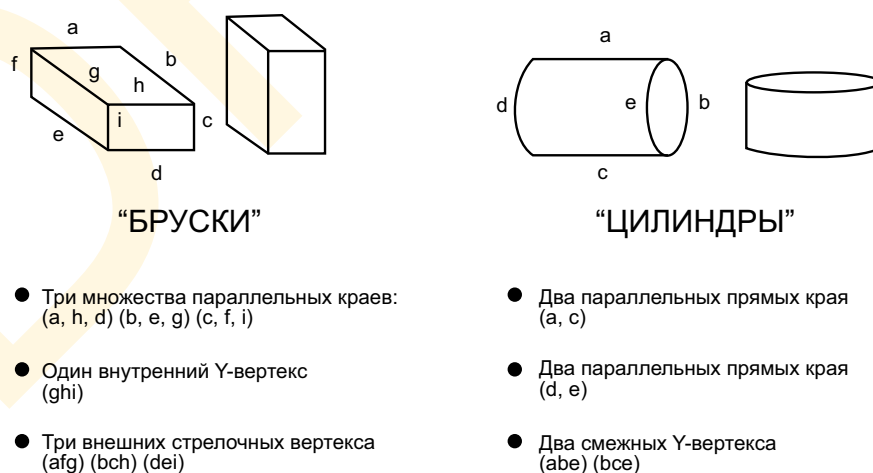
5. *Форматное соотношение*. Отношение длины оси протягивания к длиннику сечения, которое либо *примерно равно единице* (1:1) — рис. 9.3.2 А, *ось больше* — рис. 9.3.2 Н, или *длинник сечения больше* — рис. 9.3.2 I.

Понятно, что три варианта форматного соотношения увеличивают число геонов до 108, поскольку каждый из 36 исходных может существовать в трех форматных конфигурациях.

### 9.3.1.2 Неакцидентальные признаки

Несмотря на то, что геоны — суть волюметрические единицы, теория предполагает, что сами геоны идентифицируются зрением напрямую из оптических признаков ретинального изображения, в частности, краев и вертексов. Признаковые отличия между геоном «брусок» и геоном «цилиндр» даны на рис. 9.3.3. Признаки, посредством которых геон идентифицируется (1985) называют *неакцидентальными*, поскольку они отражают структуру изображения, не зависимую от «случайностей» ракурса. Например, брусок, рассматриваемый строго фронтально, дает прямоугольную проекцию на сетчатку, теряя критично важный Y-вертекс и редуцируя три множества трех параллельных краев до двух пар параллельных. Сие нехарактерный и нерепрезентативный ракурс, с которого, согласно теории, идентификация брускоформных геонов, составляющих объект, оказывается затруднительной.

Однако сей очевидный недочет может весьма негативно сказаться как минимум на теории объектной классификации. Вспомним, что исходом классификации могут явиться различные перспективные феномены, обсуждавшиеся нами ранее: в норме — эффекты канонических ракурсов (Palmer, 1981), а у больных, страдающих зрительной агнозией — ракурсов «бесполезных» (Humphreys & Ribboch, 1984; Warrington & Taylor, 1973). Ежели RBC-теория способна принять во внимание трудности, через которые про-



**Рис. 9.3.3.** Неакцидентальные свойства двух геонов. Брусок от цилиндра можно отличить по многим признакам, представленным у обоих, но лишь в нескольких ракурсах.



ходит зрение человека при естественном зрительном акте, «дефекты» эти лишь укрепляют позиции этой теории. Однако способность RBC-теории к учету перспективных эффектов до сих пор остается под вопросом.

### 9.3.1.3 Взаимоотношения между геонами

Поскольку сложные объекты понимаются RBC-теорией как конфигурации двух и более геонов в специфической пространственной организации, они кодируются как структурные описания как самих геонов, так и взаимоотношений между ними. Если геоны — это «алфавит» сложных трехмерных объектов, то пространственные отношения между геонами — это аналогия последовательности букв в словах. Подобно тому, как мы конструируем слова, организуя буквы (например, НОС и СОН), так можно конструировать и различные типы объектов, организуя геоны. Чашка и ведро на рис. 9.3.4 — тому примером.

Для представления структуры сложных объектов, состоящих из двух и более геонов, RBC задействует структурные описания, в которых можно представить 108 качественно отличных отношений между геонами. Некоторые из этих отношений отвечают за скрепление геонов (например, СТОРОНКОВСЯЗАННЫЕ и ВЕРШИНОВСЯЗАННЫЕ); прочие отвечают за свойства этих отношений, скажем, относительный размер (например, БОЛЬШЕ-ЧЕМ и МЕНЬШЕ-ЧЕМ). При 108 вариантах отношений и 108 геонах логически возможно построить более миллиона различных двухгеонных объектов. Добавив третий геон и его возможные отношения с соседями, число комбинаций вырастет до триллионов. Итак, очевидно, что геоны способны организовать огромное количество различных сложных форм. Но достаточно ли всего этого для решения вопроса зрительной классификации, мы обсудим позднее.

Наконец отметим, что коль скоро форма объекта представляется через его геоны и пространственные отношения между ними, вопрос объектной классификации в рамках RBC-теории ужимается до процесса сравнения структурных описаний интересующего объекта и множества структурных описаний известной категории начального уровня. Бидерман (1987) установил, что большинству из нас известны порядка 30000 объектных категорий, намекая тем самым на то, что существует много большее число возможных комбинаций геонов, чем фактических объектов. Сие критично, когда по нетривиальному расположению геонов зрительная система обнаруживает, что новый объект не является членом ни одной из известных ей категорий (например, один из показанных на рис. 9.1.3).

### 9.3.1.4 Этапы объектной классификации в RBC

Концепция RBC предполагает, что классификация объектов проходит в несколько этапов, некоторые из которых уже обсуждались в предыдущих главах. На рис. 9.3.5 дана оригинальная схема рабочего потока RBC (Biederman, 1985, 1987).

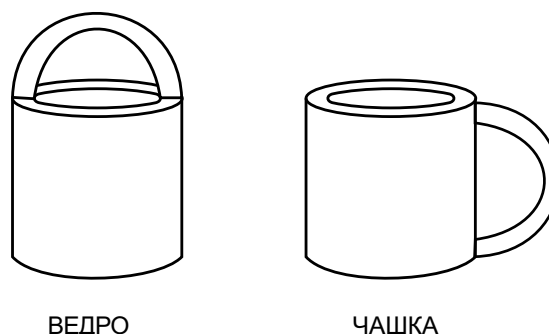


Рис. 9.3.4. Значимость отношений между геонами. Меняя соотношения одних и тех же геонов, можно получать разные объекты.

По материалам Biederman, 1987

1. *Извлечение краев.* Вначале происходит детекция краев как яркостных градиентов (см. гл. 4). При этом RBC-теория ничего не говорит о том, *как* сие происходит, но на выходе предполагается идеальная штриховая прорись всех краев, представленных в сцене.

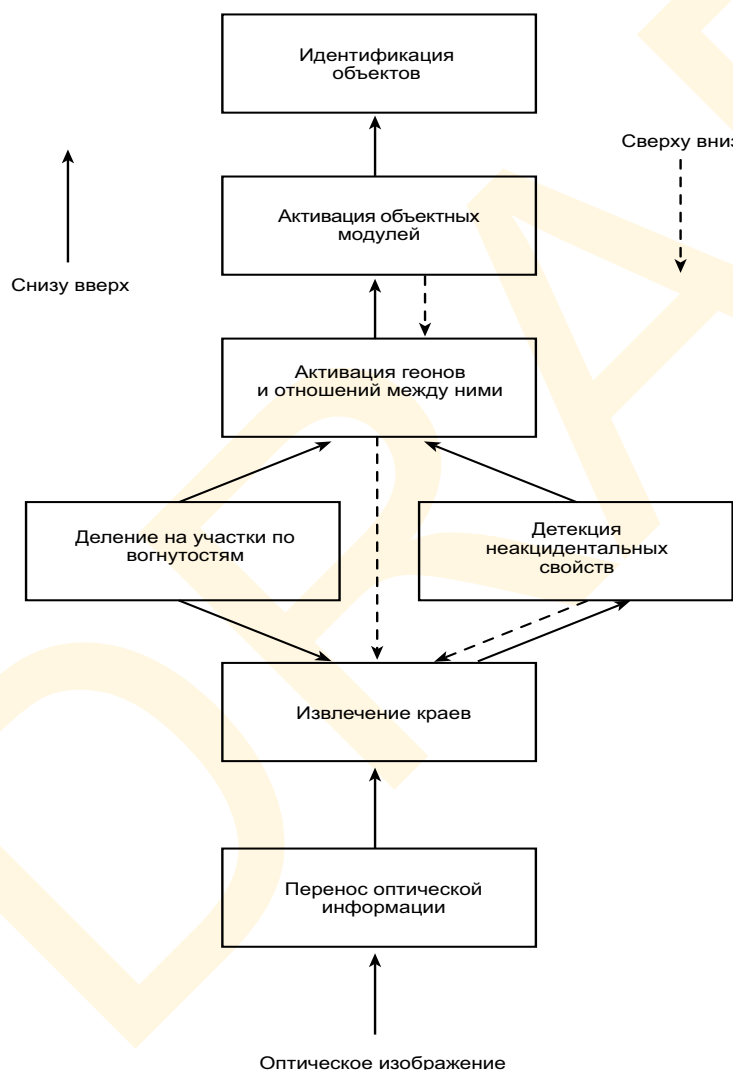
2. *Детекция признаков.* Неакцидентальные оптические свойства, нужные для идентификации геонов, извлекаются из краевой информации. Критичными являются характер краев (к примеру, искривленные или прямые), характер вертексов (например, Y-вертексы, K-вертексы, L-вертексы), параллельность (параллельные или непараллельные) и симметричность (симметричные или несимметричные). Цель сего процесса — обеспечить признаковую информацию, требуемую для идентификации геонов различных видов (см. этап 4).

3. *Разделение объектов.* Как только перечисленные признаки извлечены, система начинает разделять объекты по глубоким вогнутостям (согласно Хоффману и Ричардсу —

см. гл. 7). Целью сего процесса является разбор объекта на составляющие его геоны, но без попытки их эксплицитного сравнения их по краям и вертексам.

4. *Классификация геонов.* Совокупные результаты детекции признаков (этап 2) и разбора объектов (этап 3) используются для идентификации того, какие из специфических геонов представлены в сцене и каковы пространственные взаимоотношения между ними.

5. *Сравнение категорий.* Как только геонное описание объекта построено, оно автоматически активирует аналогичные геонные описания, сохраненные в памяти. Сравнение реализуется распространением активации по нейронной сети от геонных узелков и узелков связей, имеющихсся в представлении целевого объекта, к аналогичным геонным узелкам и узелкам связей в представлении категории. Сие сравнение — это полностью параллельный процесс, сразу же согласующий геонное описа-



**Рис. 9.3.5.** Этапы нейронной обработки согласно RBC-теории. Восходящие процессы отмечены сплошными стрелками, нисходящие — пунктирными.

По материалам Biederman, 1987.

ние объекта с со всеми категориальными представлениями и сразу использующий все геоны и все связи.

6. *Классификация объектов.* Целевой объект идентифицируется исключительно как ступень в категории начального уровня, то есть, ступень, плотнее всего классифицированная процессом сравнения, в результате которого выяснилось, что превышено некое пороговое значение.

Несмотря на то, что основной поток информации идет снизу вверх, также не исключается встречный нисходящий процессинг. Если сенсорная информация слаба (из-за зашумленности, физиологической недостаточности или по каким-то иным причинам) нисходящие процессы очевидно имеют место. В RBC имеются два участка, в которых возникновение нисходящего потока наиболее вероятно: обратная связь от геонов к признакам и обратная связь от категориального представления к геонам.

1. *Обратная связь от геонов к признакам.* Обработка на геонном уровне может повлиять на обработку, идущую на уровне признаков, поскольку не все возможные сочетания признаков формируют легальные геоны. Сказанное означает, что частичная активация геонной единицы каким-то конкретным его признаком может вызвать ожидание того, что прочие признаки также должны быть представлены, если этот геон фактически существует и в том же расположении. Сей процессинг, основанный на ожидании — это одна из форм гипотетической проверки, при которой ожидание играет роль тестируемой гипотезы.

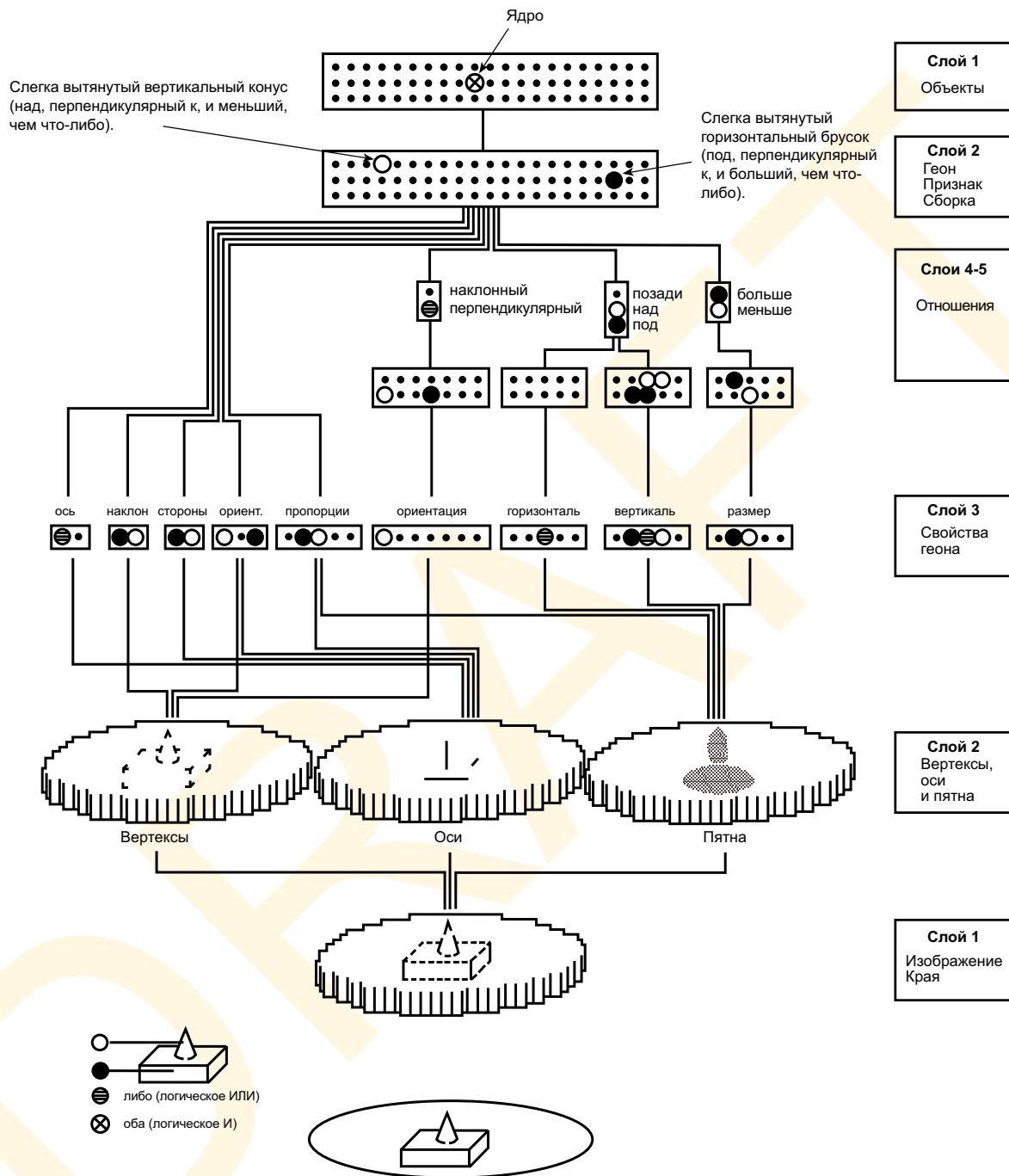
2. *Обратная связь от категорий к геонам.* Обработка категориального уровня может повлиять на обработку уровня геонного по сходным причинам: не все комбинации геонных представлений попадают в известные категории. В результате, судя по всему, объект классифицируется как принадлежащий некоей категории, выдвигающей предположение о том, что прочие геоны тоже должны быть представлены, если целевой объект фактически член этой категории. Такое предположение удобно использовать для тестирования гипотезы о типе объекта.

### 9.3.1.5 Нейронная сеть

В оригинальной формулировке RBC (Biederman, 1985, 1987) фактические механизмы описанных выше процессов не определены. Предполагалось, что они — плод распределения активации внутри сложной нейронной сети. Но при этом фактических моделей этих процессов не существовало. В итоге, Дж. Хаммель и И. Бидерман (1992) разрабатывают нейронный вариант RBC (названный ими «JIM» — John and Irv's Model), общая структура которого дана на рис. 9.3.6.

Как изначально предлагал Бидерман, на первых двух этапах идет извлечение краев (слой 1) и детекция признаков (слой 2) нейронной сети. Края извлекаются в слое 1 силами «прямых» и «кривых сегментных клеток», а также «прямых» и «кривых концевых клеток» (т.е. стопорных) во многих ориентациях и расположениях. Отметим, что сие весьма напоминает краевую концепцию рецептивных полей Хьюбела — Визеля (см. раздел 4.1.2). Слой 2 отвечает за детекцию трех типов признаков: различных видов вертексов, осевую симметрию и пятна. L-, T-, вилочковые, стрелочные, а также тангенциальные Y-вертексы весьма подходят для различения различных видов геонов (к примеру, брусков и цилиндров — см. рис. 9.3.3), а оси и пятна помогают дальнейшей спецификации атрибутов геонов, представленных в слое 3.

Атрибуты геонов кодируются паттерном активности единиц сети в слое 3, отвечающем на форму оси («Ось» на рис. 9.3.6), которая может быть прямой (s) или искривленной (c); форму сечения (X-Scn), которая может быть прямой (s) или искривленной (c);



**Рис. 9.3.6.** Общая архитектура JIM-сети, построенной на основе RBC. Описание объектов как совокупность структурных описаний геон и геонных отношений происходит на уровнях 1 — 6, а классификация достигается активацией известного категорийного представления в слое 7 (подробнее см. далее).

По материалам Hummel & Biederman, 1992.

параллельность сторон (Sides), которые бывают параллельными (p) или непараллельными (n); общую ориентацию (Orn), которая бывает вертикальной (v), диагональной (d) или горизонтальной (h); пропорцию, которая меняется от вытянутой до плоской; тонкую ориентацию (Orientation), которая меняется от вертикальной до (v) через диагональную (d) к горизонтальной (h); горизонтальность положения, которое меняется от левого (l) к правому (r); вертикальность положения от нижнего (b) к верхнему (t); размер, который меняется от большого к маленькому. Таким образом, геон определен своими значениями по каждому из перечисленных признаков.

Одним из ключевых положений JIM является то, что множество различных признаков может временно ограничиваться рамками данного геона (т.е., ассоциированы с данным геонем) за счет синхронной активности единиц сети. Сей механизм, именуемый «динамически связывающим», — это необычайно эффективный и экономичный путь представления объединения различных признаков, составляющих часть одного и того же компонента. Это также механизм, который JIM использует для деления объекта на геоны: различные признаки разнесены по разным геонам в одном и том же представлении посредством нескольких отдельных временных паттернов активации, а другое множество признаков синхронно активны в другом паттерне.

Мысль о том, что временную синхронность можно использовать для связывания признака с объектом, еще много лет назад озвучил британец Петер Мильнер (1974), а развил ее немецкий исследователь Кристоф фон дер Мальсбург (1981, 1987). Детали того, каким образом такая связь реализуется в JIM-нейронной сети, к сожалению, очень сложны для описания на страницах нашей книги. Мы лишь допустим, что сей динамический механизм связи работает и понятно, как его можно использовать для деления сцены на геоны. Что и позволяет модели исполнять объектную классификацию.

На рис. 9.3.6 различные связи «признак-геон» в слое 3 (и выше) представлены крупными кругами: белые по конусам наверху, черные по брускам внизу, с определенным уровнем яркости по каждому геону (логические «ИЛИ» конических и брусковых геонов) и X по обоим геонам (логические «И»). Таким образом, читая по представлениям в слое 3, мы можем видеть, что верхние, конические геоны (белые круги) имеют прямую ось, искривленное сечение, непараллельные стороны, вертикальную ориентацию и т.д. У нижних, брусковых геонов (черные кружки) — прямая ось, прямое сечение, параллельные стороны, горизонтальная ориентация и т.д.

В слоях 4 и 5 отношения между геонами извлекаются из геонных признаков тонкой ориентации, горизонтального положения, вертикального положения и размера. «Клетки» в этих слоях откликаются на относительные признаки, в частности — на перпендикулярность вертикальных конусов и горизонтальных брусков; на то, что конусы наверху, а бруски внизу, и что бруски крупнее, а конусы мельче. Совместно с признаками геонов слоя 3, сии геонные отношения активируют соответствующую «ассемблированную геонную единицу» в слое 6 — одну на каждый объектный геон. Активация этих единиц соответствует стадии геонной идентификации в оригинальной бидермановой теории. Единицы слоя 6 представляют множество всех возможных геонных описаний, а подмножество активированных — геонное структурное описание в JIM. Множество всех активных геонных единиц активирует затем представления категорий начального уровня в слое 7 до степени, зависящей от того, сколь близко геоны, представленные в нейронном изображении, соответствуют геонам, закодированным в единицах объектной категории. В этом случае геон-структурное описание строже согласуется с единицей «атомной установки», и сие есть механизм, по которому модель классифицирует данный конкретный объект.

Несмотря на то, что множество деталей JIM весьма специфичны, многие ее узлы



сегодня работают уже не так, как во времена Хаммеля и Бидермана, то есть в 1992 г. За прошедшее с тех пор время модель набрала силу, развилась и способна теперь классифицировать многие простые объекты. При этом система все еще не может работать с полутоновыми изображениями, хотя на это способны большинство теорий высокоуровневой перцепции и объектной идентификации.

### 9.3.2 Эмпирические феномены

Будет весьма поучительно разобрать, как эмпирические данные, которых мы касались ранее в текущей главе, толкуются RBC-теорией объектной классификации. Мы увидим, что многие из них уверенно объясняются силами RBC, другие же требуют дальнейшей спецификации и исследования.

#### 9.3.2.1 Типические эффекты

Типичные эффекты в RBC соответствуют отличиям в уровне активации данного представления категории (скажем, «птицы») посредством различных экземпляров внутри категории (скажем, «дрозд», «воробей», «пингвин» и «страус»). Вспомним, что активацию категориальных представлений полагают функцией подобия их геонного описания геонному описанию целевого объекта. Поскольку геонные описания должны быть специфическими и конкретными, категориальное представление в логике геонной структуры определяет т.н. *грубый прототип*. «Грубый», потому что геоны и отношения между геонами описываются лишь качественно. Как уже было сказано, затем следует прототипический этап (например, «дрозд»), активирующий категориальное представление («птица») сильнее, чем этап атипический (например, «страус»). Таким образом, типичные эффекты запускаются в RBC из уровня активации категории в процессах сравнения, как минимум при существенных отличиях в типичности.

Толкование тонких типических эффектов — скажем, отличия в типичности у дрозда и орла в категории «птиц» — для RBC затруднительно. Дело в том, что грубые, дискретные категории из 108 геонов и 108 геонных отношений очевидно не могут дать достаточного представительского разрешения, необходимого для различения между похожими экземплярами подчиненного уровня. Однако, если допустить тонкие градации количественных параметров в геонных описаниях категориального прототипа и целевого объекта, не имеет значения то, почему типические эффекты в целом невозможно учесть таким путем. Действительно, недавние усовершенствования RBC предоставляют такую возможность (Hummel & Stankiewicz, 1998).

#### 9.3.2.2 Категории начального уровня

Феномен классификации на начальном уровне заключен в том, что типичные члены категорий начального уровня исходного классифицированы на уровне базовом, тогда как атипичные члены чаще классифицируются на подчиненном уровне. Как сей феномен толкует RBC-теория?

Начальный уровень для данного объекта в иерархии категорий определяется RBC через то, какое из категориальных представлений сильнее активировано на этапе сравнения категорий. Неверно думать, к примеру, что все члены одной категории базового уровня (например, все птицы) будут одинаково активировать птичье представление. Действительно, есть основания полагать, что напротив — геонное описание дрозда активирует геонное описание категории «птиц» сильнее, чем геонное описание пингви-

на или страуса. То есть, мы ожидаем, что у атипичных экземпляров их субкатегория (например, пингвины или страусы) даст существенно большее соответствие, чем их базовая категория. При более тонком различии, нужном для отличия дрозда от воробья или от трясогузки (или подобной птички) и более грубом геонном описании, весьма вероятно, что RBC не справится с задачей (как минимум без дополнительного более тонкого представления базовых геонных параметров). Следовательно, в RBC вначале активируются геонные описания птиц, тогда как пингвины и страусы будут активировать геонные описания, соответственно, пингвинов и страусов. Как при этом проходит классификация на подчиненных уровнях — открытый вопрос.

### 9.3.2.3 Условия просмотра

Перспективные эффекты (эффекты ракурса) могут возникнуть как минимум на двух этапах действия RBC — этапе геонной классификации и/или этапе сравнения категорий. Как было показано выше, в некоторых ракурсах геонной классификации можно помешать, поскольку оптимальный набор краевых признаков геона оказывается недоступным из-за самоокклюзии. Простой вид сверху на автомобиль, например, даст лишь одну поверхность геонов (из которых состоят его корпус и крыша), что чрезвычайно затрудняет восприятие трехмерной формы этого объекта. Плюс к тому, правильные пространственные отношения между геонами могут стать слишком сложными для распознавания.

Однако даже если все видимые геоны идентифицированы безошибочно, проблемы классификации иногда возникают из-за неудачного ракурса, поскольку критично важные геоны оказываются полностью закрытыми геонами видимыми. Все тот же вид сверху на автомобиль исключает видность его колес, что напрочь мешает процессу сравнения его геонных описаний с представлением автомобиля в памяти (поскольку колеса — это наиболее заметные элементы автомобилей). Следовательно, даже если RBC будет базироваться на представлениях трехмерных волюметрических структур, то есть ракурс окажется инвариантен настолько, насколько возможно, нельзя исключить того, что перспективные эффекты не возникнут из-за окклюзии.

Описанные выше эффекты были исключены из опыта Бидермана — Герхардштайна (1993), поскольку исследователи заранее позаботились о том, чтобы одни и те же геоны (и отношения между ними) были видны в разных ракурсах объекта. Таким образом, изучалось только подмножество перспективных отличий, при которых не было отличий в видности геонов и не было явных отличий в прайминге. Результат вполне объясним силами RBC: эффекты видности геонов должны давать отличия в классификации при смене ракурса.<sup>1</sup>

С эффектами ориентации RBC-теория должна работать по-другому, поскольку одни и те же геоны по определению видимы при любом повороте плоскости проекции объекта вокруг зрительной оси. Эффекты ориентации возникают потому, что некоторые отношения между геонами, зашитые в RBC, не инвариантны по ориентации. RBC-теория, к примеру, определяет отношения выше/ниже скорее в гравитационной

<sup>1</sup> Объяснение перспективных эффектов ныне весьма противоречиво. Жесткий критический комментарий к материалам Бидермана — Герхардштайна (1993) опубликовали Тарр и Бюльтхов (1995), показавшие, что абсолютно все доступные данные указывают на ракурсо-зависимость процессов распознавания, т.е. не требующих деления объекта на части. Бидерман и Герхардштайн (1995) решительно опровергли эти аргументы и пришли к противоположным выводам.

логике, нежели в логике объектно-относительной, внутренней (Hummel & Biederman, 1992). Например, в представлении лампы, хранимом в памяти, колба располагается *над* цоколем. Если идентифицируется лампа, повернутая на  $45^\circ$ , представление колбы окажется и *над*, и *около* цоколя в JIM-схеме кодирования (см. рис. 9.3.6, слой 5). Введение псевдоотношения *около* понижает соответствие представления изображения содержимому памяти, тем самым замедляя распознавание. Дальнейший поворот лампы усугубляет разницу между изображением и содержимым памяти. Любопытно, что поворот на  $180^\circ$  несколько повышает соответствие, поскольку псевдоотношение *около* в этом случае исключается, и мы наблюдаем некоторое уменьшение времени реакции в данной ориентации (см. рис. 9.2.9).

Однако данные ряда опытов указывают на то, что, идентифицируя объекты, испытуемые начинают поворачивать их из дисориентированного положения в прямостоящее (Jolicoeur, 1985; Tarr & Pinker, 1989, 1990). RBC никак не объясняет сей феномен: если объект был согласован с категорией посредством ориентационно-инвариантной информации, почему мы стремимся повернуть его в прямостоящее положение? Возможно, причина кроется в том, что поворот — это исключительно верификационная процедура, исполняемая для того, чтобы убедиться в полноте соответствия (Corballis, 1988). Вторая версия звучит так, что только очевидная часть объекта (например, его верх) сравнивалась до поворота и что подлинная классификация не имеет места до тех пор, пока представление целого объекта не трансформируется до состояния конгруэнтности представлению категории. Наконец, процессу поиска верха фигуры может понадобиться дополнительное время (Biederman, 1987). К сожалению, мы до сих пор не знаем, какой из вариантов верный.

#### 9.3.2.4 Долевая структура

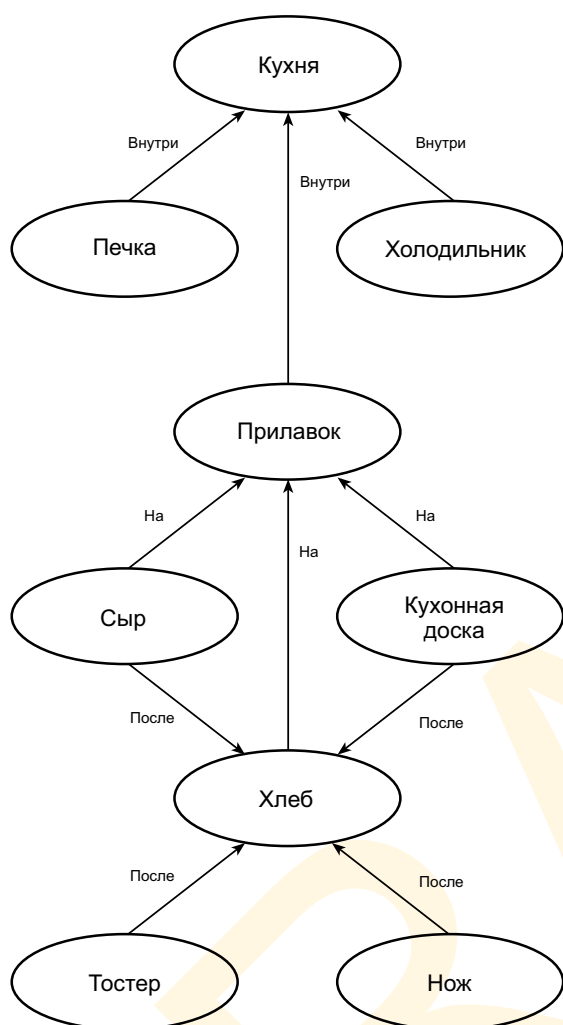
RBC легко справляется с прогнозом долевых эффектов, наблюдаемых в опыте Бидермана — Купера (1991 b). Сие неудивительно, поскольку краеугольным камнем концепции является допущение, что объекты распознаются скорее сравнением их частей (например, геонов) с представлениями категорий, нежели целно. Собственно, эта мысль и вдохновила Бидермана к проведению эксперимента.

В первом опыте по обезконтуриванию, описанном выше, RBC учитывала недостаточность отличий в прайминг-эффектах между условиями идентичности и линейной дополненности, поскольку в рамках RBC имеются компонентные геоны, активирующие категорийные представления, но при этом отсутствуют индивидуальные оптические признаки, такие как линии и края. Эти признаки служат лишь для активации геонов. Но с момента как те активированы, признаки утрачивают значение.

RBC учитывает выраженные прайминг-отличия в условиях идентичности и линейной дополненности с геон-удаленными стимулами во втором эксперименте, поскольку в частедополняющих условиях во втором представлении активируется совершенно иное множество геонов. Таким образом, результаты обоих экспериментов составляют очень важный тест долевых теорий классификации, среди которых RBC — наипервейшая.

#### 9.3.2.4 Контекстные эффекты

Контекстные эффекты нельзя объяснить непосредственно силами RBC, поскольку — это теория распознавания единичных объектов. Однако же ее можно поднять до уровня, на котором оказывается возможным учет некоторых контекстов. Недостающий ингредиент — представление структуры сцены, то есть, перцептивных «объектов», что не-



**Рис. 9.3.7.** Долевая схема сцены кухни (см. рис. 9.2.15). Отношения даны условно для иллюстрации недостатка специфичности при создании базы данных о сцене.

Palmer, 197 b.

гут в разных местах станции. Вопреки всем этим различиям представления отдельных объектов и полноконтекстных сцен схожи тем, что и первые, и вторые часто относят к схемам: схема лица, схема человека, схема птицы и т.д., равно как и схема автозаправки, и детской площадки. Отметим, что в описании всех представлений сцен особое значение придают единству и взаимосвязанности всей перцептивной информации.

Иерархические отношения между частями объектов, объектами и сценами (рис. 9.3.7) демонстрируют нам то, что, к примеру, схема кухни кодируется в виде иерархической сети (Palmer, 1975b): объектная схема — это компонент такой же схемы сцены, равно как геон — компонент схемы объекта. Схемная сеть может объяснить

сколько больше, чем отдельные объекты. Представления, несущие мультиобъектную структуру цельных сцен обычно именуют *схемами* (например, Palmer, 1975 b) или *фреймами* (например, Minsky, 1975).<sup>1</sup> Важным аспектом схем является то, что они кодируют зависимости между объектами, возникающими в значимом окружении. Например, может существовать схема кухни, содержащая геонные описания плиты, холодильника, кухонного шкафа, тостера и пр.; бывают схемы детских площадок, несущие описания качелей, песочницы, каруселей и т.д.; а также многие другие схемы сцен, встречающихся нам в повседневности.

Идея схем — это всего лишь продолжение структурной иерархии объектов и частей (Palmer, 1975b). Например, колпак — это часть колеса, которое, в свою очередь, — часть автомобиля. Сам автомобиль может явиться фрагментом иных сцен, многие из которых несут в себе другие схемы представления: схема гаражей, схема городских улиц, схема паркинга, схема шоссе, схема АЗС, схема мойки и т.д. Первейшее отличие состоит в том, что составляющие объекты и пространственные взаимоотношения в схемах сцен по сути более вариабельны и вероятны, чем таковые в объектном представлении. У машин всегда есть колеса и они всегда расположены в одних и тех же местах автомобиля. Но на заправках автомобиля могут находиться в данный момент, а могут и не находиться, а когда они есть, то располагаться могут в разных местах станции. Вопреки всем этим различиям представления отдельных объектов и полноконтекстных сцен схожи тем, что и первые, и вторые часто относят к схемам: схема лица, схема человека, схема птицы и т.д., равно как и схема автозаправки, и детской площадки. Отметим, что в описании всех представлений сцен особое значение придают единству и взаимосвязанности всей перцептивной информации.

<sup>1</sup> Во избежание смыслового конфликта с концепцией эталонных реперов, описанной в главах 7 и 8, термин «фрейм» мы более употреблять не будем.

контекстные эффекты следующим образом: предположим, что целевой объект в сцене — это буханка хлеба, показанная на вставке А рис. 9.2.15. Контекст являет собой сцену кухни, в которой имеется кухонный стол, хлебобрезка, тостер, нож и т.д. Каждый узелок сети, соответствующий таким объектам, активируется процессами объектной классификации. Активация узелков, в свою очередь, растекается по ближайшим узелкам, подготавливая их к распознаванию. Так, к примеру, кухонный стол, тостер, хлебобрезка и нож будут косвенно активировать внутреннее представление хлеба, соответствующее перцептивному ожиданию того, что хлеб тоже может оказаться представленным в такой сцене. Словом, прайминг, вызванный контекстной активацией, повышает вероятность того, что зрительная система в данном контексте станет интерпретировать любой буханкоподобный предмет как хлеб. Сие как раз то, что демонстрируют нам результаты опытов (Palmer, 1975 b).

### 9.3.2.5 Зрительная агнозия

Универсального объяснения феномену зрительной агнозии, вероятно, нет, поскольку существуют целые группы симптомов, находящихся в рамках одной общей категории (Farah, 1990). Как было показано ранее, зрительно агнозический феномен «бесполезного ракурса» можно истолковать теми же сложностями в обработке, каковая в норме имеет место при перспективных эффектах. Другой процесс, который может участвовать в нарастании зрительной недостаточности — это процесс выбора соответствующего объектно-центрированного репера вдоль оси длинника (например, Humphreys & Riddoch, 1984, 1985). Словом, мы еще очень мало знаем о зрительной агнозии, чтобы уверенно указать этап, на котором происходит сбой классификации у этих больных. Большинство же текущих исследований посвящены тонкому анализу того, как сие можно учесть в рамках той или иной научной парадигмы наподобие марровой (1982) или бидермановой (1987).

### 9.3.2.6 Минусы

При всех достоинствах у концепции RBC как теории объектной классификации начального уровня, имеется ряд очевидных недостатков. Один из наиболее серьезных описан в гл. 8 и ранее в текущей главе: недостаточность представительской способности. Частично сие проистекает из ограничений, свойственных типизированным цилиндрам как примитивам представления информации о форме. Но эти проблемы усугубляются дискретной классификацией типизированных цилиндров всего лишь из 108 геонов и редукцией всех возможных отношений между этими геонами до 108 бинарных связей, что и далее сокращает число зрительных деталей в коде геонных описаний. Даже триллионы логически возможных комбинаций недостаточны для описания малой толики отличий по форме между объектами, которые мы легко и регулярно распознаем при классификации и идентификации их в повседневности. Бидерманова набора геонов может стать достаточно для большинства ситуаций классификации на начальном уровне, но и в этом окончательной уверенности нет. Отличие собаки от кошки или осла от лошади требует, например, относительно тонкого количественного различения по форме компонент и по пространственным отношениям между ними. Такое отличие очевидно выходит за рамки возможностей RBC-теории, по крайней мере в ее оригинальной формулировке.

В подчиненных категориях потребуется даже еще большая различительная способность. Зрительные отличия между собаками разных пород, к примеру, потребуют



ощутимо большей количественной информации, чем то, что дает геонное структурное описание. Рассмотрим, например, как 108 геонов и 108 соотношений между ними участвуют в классификации при распознавании наших знакомых по лицам. Даже если лица можно естественным путем анализировать по конфигурациям геонов (однако нет уверенности, что это возможно), то следует помнить, что человеческие лица весьма схожи и, теоретически, должны иметь одно и то же геонное описание.

На критику сию есть несколько вариантов ответа. Дж. Хаммель и Б. Станкевич (1998) пошли самым прямым путем и предложили пересмотр теории, который дал более тонкие градации количественных представлений. Сам же Бидерман заявил, что RBC не предназначена для распознавания лиц, поскольку полагал, что процессы объектной классификации качественно отличны от процессов детекции лиц. Действительно, имеются данные о наблюдениях, подтверждающих эту мысль и включающих в себя доказательства того, что обработка лиц идет в разных отделах головного мозга (например, Tanaka & Farah, 1993; Wallace & Farah, 1992). Однако ж данные эти — предмет различных интерпретаций, не исключающих, в частности, наличие в коре того, что называют «зона лица» и что фактически задействовано, всякий раз когда тонкое визуальное различие проходит в сугубо практической ситуации (например, Gauthier & Tarr, 1997). В любом случае еще очень рано делать какие-либо выводы о том, насколько восприятие лиц специфично или отлично от иных видов зрительного распознавания.

Еще одной потенциальной сложностью в RBC-теории объектной классификации мы полагаем то, что до сих пор неясно, способна ли RBC классифицировать объекты на основе их оптических проекций так, как это делает зрение человека. Есть надежда, что доказательство придет из сферы машинного моделирования, когда компьютер сможет идентифицировать простые объекты из серых полутоновых изображений. Несомненно, что RBC звучит как вполне работоспособная концептуальная схема во многих отношениях, но основная особенность машинного зрения в том, что даже сильная теория часто не может работать так, как хотелось бы, а ежели и работает, то в весьма и весьма упрощенных условиях.

Отметим также, что нейронная сеть Хаммеля — Бидермана (1992), построенная на базе RBC, увиливает от большинства сложностей, работая с описаниями признаков, закодированными вручную, но не с полутоновыми картинками. Сие один из главных недостатков теорий, построенных на волюметрических структурных описаниях, и у RBC-теории не будет прочного основания, пока вопрос этот не будет решен.

### 9.3.3 Ракурсоспецифические теории

RBC наглядно иллюстрирует ход научной мысли 70-х — 80-х годов XX века по вопросам объектной классификации, будучи во многом схожей с предшествовавшими ей моделями, предложенными Бинфордом (1971), Марром и Нишихарой (1978), а также Пальмером (1975b) и Винстоном (1975). В частности, подобно марровой модели, RBC исходит из того, что объекты представляются в зрительной системе иерархическими структурными описаниями волюметрических примитивов, в основе которых лежат типизированные цилиндры. Несмотря на многолетние усилия по совершенствованию этих моделей, их машинное воплощение так и не достигло приемлемого уровня при работе с необработанным видеоматериалом. Из-за чего наука стала искать иные подходы.

Итак, нам предстоит рассмотреть три альтернативных концепции: *графический подход* (Koenderink & Van Doorn, 1978), *приводку к трехмерным моделям* (Huttenlocher & Ullman, 1987; Lowe, 1985; Ullman, 1989) и *расположение с двухмерными ракурсами* (Poggio & Edelman, 1990; Ullman & Basri, 1991).

### 9.3.3.1 Множественные ракурсы

Отправной точкой многих альтернативных концепций являются перспективные эффекты, описанные нами ранее (например, Palmer, 1981). Тот факт, что распознавание и классификация не инвариантны в различных ракурсах объекта, выявляет возможность его идентификации непосредственным сравнением двухмерных ракурсов этого объекта с каким-либо представлением ракурсо-специфической категории. Идея канонического ракурса подводит наши рассуждения к т.н. *ракурсо-специфической теории сравнения шаблонов*: каждый объект должен представляться в памяти единичным каноническим двухмерным ракурсом, и двухмерные проекции должны сравниваться с ним напрямую. Затем канонические ракурсы быстро и точно классифицируются, поскольку их можно сравнивать непосредственно с двухмерным каноническим представлением в памяти. Прочие входные ракурсы не должны сравниваться *as is*, но для успешного двухмерного сравнения должны поворачиваться в каноническое расположение.<sup>1</sup> Отметим, что для выполнения таких преобразований требуется больше времени и повышается вероятность ошибок (Palmer, Rosch & Case, 1981; Edelman & Bülthoff, 1992; Bülthoff & Edelman, 1992).

Однако сие весьма однобокое толкование перспективных эффектов влечет за собой как минимум три серьезных проблемы:

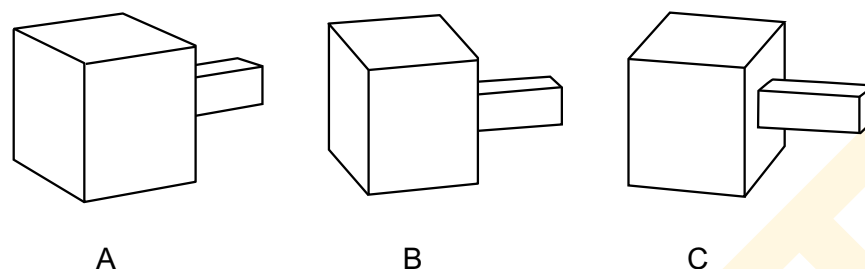
— во-первых, оно логически недостаточно для решения вопроса. Не может двухмерный ракурс без дополнительной информации обеспечить точное распознавание трехмерного объекта или его классификацию из множественных ракурсов. У объектов всегда имеется множество поверхностей, которые из-за самоокклюзии просто не представлены в единичном ракурсе, даже каноническом. Закрытые поверхности невозможно «вычислить» из единичного ракурса за исключением специальных случаев, таких, в частности, как симметричность объекта (Vetter, Poggio & Bülthoff, 1994). Следовательно, в большинстве ситуаций полную трехмерную форму объекта нельзя восстановить из единичного двухмерного ракурса;

— во-вторых, ежели мы допустим, что прочие ракурсы распознаются через приво-  
дку в соответствие ракурсу каноническому, в памяти должно существовать некое представление или процесс, реализующие поворот, открывающий ранее закрытые поверхности. Как мы покажем далее, для выполнения этой задачи существуют пути применения двух и более ракурсов, но не существует единственного ракурса, достаточного в общем случае;

— в-третьих, гипотеза приво-  
дки в канонический ракурс не стыкуется с данными экспериментов. Скрытые периоды именован-  
ия — это не просто растущая функция от отличия по ориентации между входным изображением и единственным каноническим ракурсом (Palmer, 1981; Edelman & Bülthoff, 1992) — обычно имеются несколько подходящих ракурсов с разных точек наблюдения, далеко отставленных друг от друга, с серией неподходящих ракурсов между ними.

Вероятнее всего, существуют множественные двухмерные представления с различных точек наблюдения, которые зрение может использовать при распознавании объектов. Эти множественные ракурсы, вероятно, — это те ракурсы, с которых объект мы видим чаще всего. Как было показано в разделе 9.2.2, Тарр и Пинкер (1989б 1990) в серии опытов, посвященных зрительной идентификации двухмерных фигур в различ-

<sup>1</sup> Альтернативно канонический ракурс в памяти можно повернуть в положение входящего стимула.



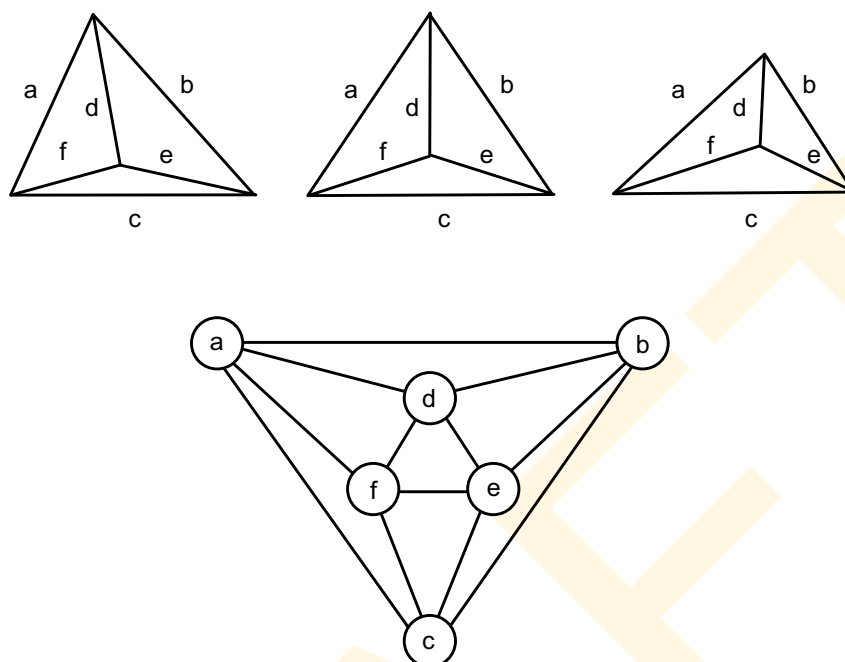
**Рис. 9.3.8.** Качественное подобие одного и того же объекта в различных ракурсах. Ракурсы А и В предъявляют наблюдателю одни и те же поверхности в одних и тех же качественных соотношениях, тогда как в ракурсе С видны другие поверхности.

ных ориентациях во фронтальной плоскости, представили доказательства возможности такого хода событий. После того, как на предварительном этапе опыта испытуемым несколько раз в нескольких ориентациях были предъявлены незнакомые ранее литероподобные фигуры, на втором этапе опыта их распознавание в ранее виденных ориентациях пошло существенно быстрее. Результат явно указывает на то, что образы были сохранены в памяти. Испытуемые могли идентифицировать те же фигуры и в тех ориентациях, которых ранее не видели, но времени на это уходило больше, и время отклика росло как функция отличия по ориентации между тестовым изображением и ориентацией, в которой объект был предъявлен на первом этапе эксперимента. По мере накопления визуального опыта, эффект ориентации на время распознавания плавно понижался. Сие наблюдение указывает на то, что распознавание может фактически выступить посредником между фактическим объектом и более или менее полным каталогом различных ракурсоспецифических двумерных представлений.

Сия гипотеза напоминает нам о шаблонных представлениях формы, которые тоже ракурсоспецифичны. Отличие же состоит в том, что множественноракурсным представлениям не требуется шаблонов, карты признаков или даже структурных описаний. Критично важным моментом является то, что это представления двумерной оптической (т.е. пикториальной) структуры, но не трехмерной предметной (т.е. волюметрической).

### 9.3.3.2 Аспект-графы

В 8-й главе мы говорили о том, что шаблоны в целом не могут служить представлением трехмерных форм, поскольку для распознавания объекта потребуется несчетное число различных ракурсов этого объекта. Сей тезис формально верен, но многие ракурсы одного и того же объекта фактически очень схожи, отличаясь лишь немногими метрическими деталями. Рассмотрим, например, ракурсы объекта, показанного на рис. 9.3.8. Все они немного отличны друг от друга, но в вариантах А и В наблюдателю видны одни и те же поверхности, края и вертексы, в строго одних и тех же конфигурациях двумерных отношений. Часть С, однако, качественно отлична тем, что открывает те поверхности, которые не были видны ранее, но при этом закрывает другие. Возможно, два первых ракурса можно распознать сравнив с единичным ракурсоспецифичным представлением, сохраненным в памяти, используя некий более абстрактный формат, чем шаблон с различными ракурсно-специфическими представлениями, необходимыми для распознавания третьего. Ежели так, то критично важными вопросами явятся: Как сие можно выполнить? И как много качественных отличных



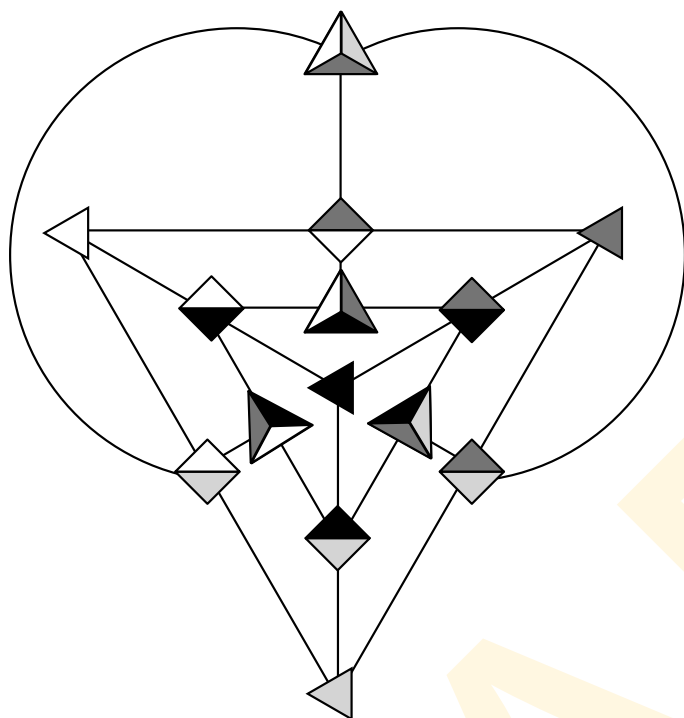
**Рис. 9.3.9.** Аспект-графы тетраэдра. Структуру фигуры во всех трех ракурсах можно представить единичной аспект-графой, демонстрирующей взаимосвязанность краев через отрезки и узелки.

ракурсов могут понадобиться для распознавания объекта из различных ракурсов, типичных для повседневного зрения?

Идея распознавания объекта путем сравнения его текущего ракурса с множеством качественно различных представлений с разных точек наблюдения была предложена голландскими психофизиками Жаном Кёндеринком и Андреа ван Дорн (1979). Ученые назвали свою идею *концепцией аспект-граф*. Аспект-графа — это нейронная сеть представлений, содержащих все топологически отличные двухмерные ракурсы (*аспекты*) одного и того же объекта: каждый аспект представлен структурным описанием согласно его топологии — множеству краев и вертексов, означающих границу между видимыми и закрытыми поверхностями в проекционном изображении, как показано на рис. 9.3.9 по одному из аспектов тетраэдра. Отметим, что поскольку все три ракурса вверху имеют одну и ту же топологию — то есть, одни и те же края и вертексы, качественно связанные одним и тем же образом в отношении специфической ориентации краев и длин этих краев — они представлены одним и тем же аспект-описанием.

Представления двух различных аспектов соединены внутри одной аспект-графы объекта, если имеется непрерывная смена ракурса наблюдения. На рис. 9.3.10 даны связи в множестве из 14 топологически отличных ракурсов (аспектов) тетраэдра. Двигаясь по какой-либо траектории ракурсов вокруг объекта, мы прочертим линию, соединяющую аспекты. В графе, однако, смещение от одного аспекта к другому возникает в момент изменения ракурса, вызывая окклюзию или наоборот — дискклюзию краев и вертексов. Таким образом, проход граф оказывается дискретен, даже если изменение ракурса — непрерывно, поскольку изменение между узелками возникает в момент, когда топологическая структура двухмерных краев в ретинальной проекции меняется.

Несмотря на то, что полное граф-представление с равной доступностью ко всем ас-



**Рис. 9.3.10.** Четырнадцать аспектов тетраэдра. Все возможные аспекты простого тетраэдра представлены в данной сети. Либо одна, либо две, либо три поверхности можно видеть одновременно, а линии между различными аспектами демонстрируют нам физическую возможность переноса. Разные поверхности отмечены разными тонами.

По материалам Koenderink & van Dorn, 1979.

мых для захвата формы объекта. Частично так происходит потому, что хотя аспект-графы и являются ракурсно-специфическими представлениями, они намного более абстрактны, чем шаблоны, поскольку являются структурными описаниями соединенности краев и вертексов и переходов между разными ракурсами. Но даже в этом случае остается открытым вопрос о том, действительно ли этого сокращения достаточно для построения адекватной теории распознавания объектов зрением человека.

Как легко догадаться, сложность аспект-граф растет по мере усложнения представляемого объекта. Понятно, что у шара один аспект, но даже уже тетраэдру, для полного описания его структуры нужны 14 разных аспектов (см. рис. 9.3.10). Число аспектов у более или менее сложного объекта, такого как невозможная фигура на рис. 9.1.3 — исчисляется тысячами. В сравнении с аналогичным числом шаблонов — это уже прогресс, но все равно очень много.

Вторая сложность применения идеи аспект-граф состоит в том, что, коль скоро они абстрактны, то не могут держать отличия между объектами с различными количественными показателями при общей двухмерной топологии краев и вертексов в их проекционных ракурсах. На рис. 9.3.11 даны несколько объектов, имеющих одно и то же

спектам даст ориентационно-инвариантное распознавание объекта, перспективные эффекты при этом все же могут иметь место. Возможно, наиболее очевидно то, что представление наиболее частых аспектов заполучить проще, чем редких. Скорость доступа может также зависеть от числа краев и вертексов: аспекты, в которых больше краев — быстрее доступны.<sup>1</sup> Отметим, что обе сии вероятности — суть исключительно вариации ранее предложенных гипотез учета перспективных эффектов (гипотезы частоты и гипотезы максимальной информации) заявленных в логике специфической структуры аспект-граф.

Аспект-графы совершенствуют теорию шаблонов в представлении объектов, радикально сокращая число ракурсно-специфических представлений, необходи-

<sup>1</sup> На первый взгляд, сие выглядит парадоксально и имеет смысл, когда сравнение краев идет параллельно, поскольку дополнительные края создают дополнительные ограничения, делая топологическую структуру объекта ракурсно более однозначной.



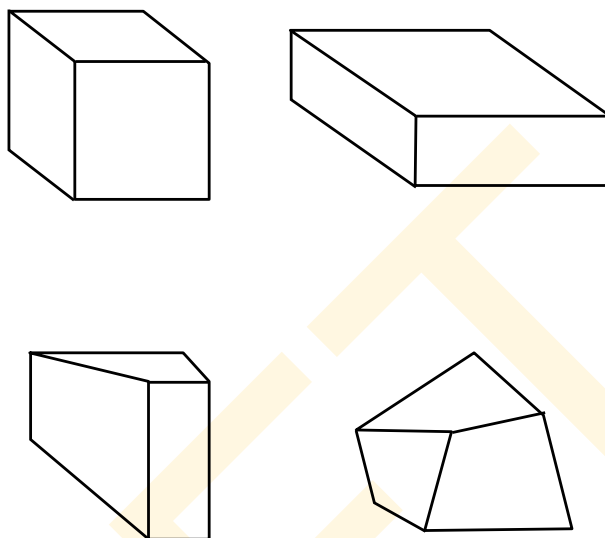
аспектное представление, но разные перцептивные формы. Сие не проблема для геонного представления, поскольку геоны, вопреки их глубоко качественной природе, включают в себя и такие признаки, как пропорции, симметрию и параллелизм краев, то есть чувствительны к метрическим показателям. Стандартные аспект-графы не содержат сей количественной информации.

Третья проблема в том, что даже если наблюдателю предъявлено лишь одно представление ракурса объекта, он обычно догадывается о трехмерной структуре этого объекта и способен узнать его во многих новых ракурсах. Сия способность, напрямую не учитывается аспект-графами для подлинно новых объектов, поскольку система

аспект-граф — это система топологических описаний двухмерных ракурсов, с которых объект был виден. Несмотря на то, что количественно распознать новые ракурсы можно из относительного малого числа должным образом ограниченных двухмерных ракурсов (Ullman & Basri, 1991), сие не учитывается концептом аспект-граф, который хорош для распознавания объектов только из качественных ракурсов, представленных в памяти. Новое изображение объекта будет распознаваться системой аспект-граф только если оно получено из того же аспекта, что и предыдущий ракурс.

Вторая проблема задает ключевое направление совершенствования ракурсоспецифических теорий, указывая на то, что больше структурности содержится в частных вариациях внутри ракурсов единичного аспекта, нежели схвачено топологией его краев. Сие обнаруживается, когда мы понимаем, что все ракурсы внутри аспекта соотносятся через отличие двухмерных проекций одного и того же трехмерного объекта. Поскольку проективные трансформации намного более ограничены, нежели трансформации топологические, теория ракурсов может найти отличия между объектами, показанными на рис. 9.3.11, если для использования этой дополнительной структуры придумать какой-то свой метод.

Третья проблема свидетельствует о том, что должен существовать некий путь экстраполяции от нескольких ракурсов к полноценной трехмерной структуре, или, как минимум, к тому, как будет выглядеть объект в других ракурсах. Эти рассуждения привели некоторых исследователей к сходным идеям, касаемо того, как можно распознать объект, сравнив входящие двухмерные ракурсы с напрямую сохраненными в памяти как модели, либо выработанными процессами, объединяющими две или более из этих моделей (Poggio & Edelman, 1990; Ullman, 1996; Ullman & Basri, 1991). Давайте посмотрим, каким образом эти идеи могут помочь в развитии более сложных ракурсоспецифических теорий распознавания объектов.



**Рис. 9.3.11.** Топологически эквивалентные ракурсы четырех различных объектов. Поскольку топология краев и вертексов одна и та же во всех четырех вариантах, мы получим одно и то же аспект-представление, вне зависимости от очевидных метрических отличий.

### 9.3.3.3 Приводка к трехмерным моделям

Отправной точкой к таким теориям явилась количественная работа по отработке методов прямого и холистического сравнения двухмерных проекций, с трехмерными моделями объектов, сохраненными в памяти (Lowe, 1985; Huttenlocher & Ullman, 1989). Базовым посылом являлось решение этой задачи серией процессов трансформирующей приводки (аналогично тому, что происходит при решении вопроса эквивалентности форм — см. раздел 8.1). Ключевыми в данном случае являются следующие процессы:

1. Поиск соответствия между очевидными признаками в оптической проекции и признаками в модели.

2. Определение точки наблюдения, которая наилучшим образом приведет эти признаки изображения в соответствие признакам трехмерной модели.

3. Расчет проекции полной трехмерной модели на двухмерное изображение в ракурсе, определенном нами в п. 2, включая *все* точки (а не только те, что самоочевидны).

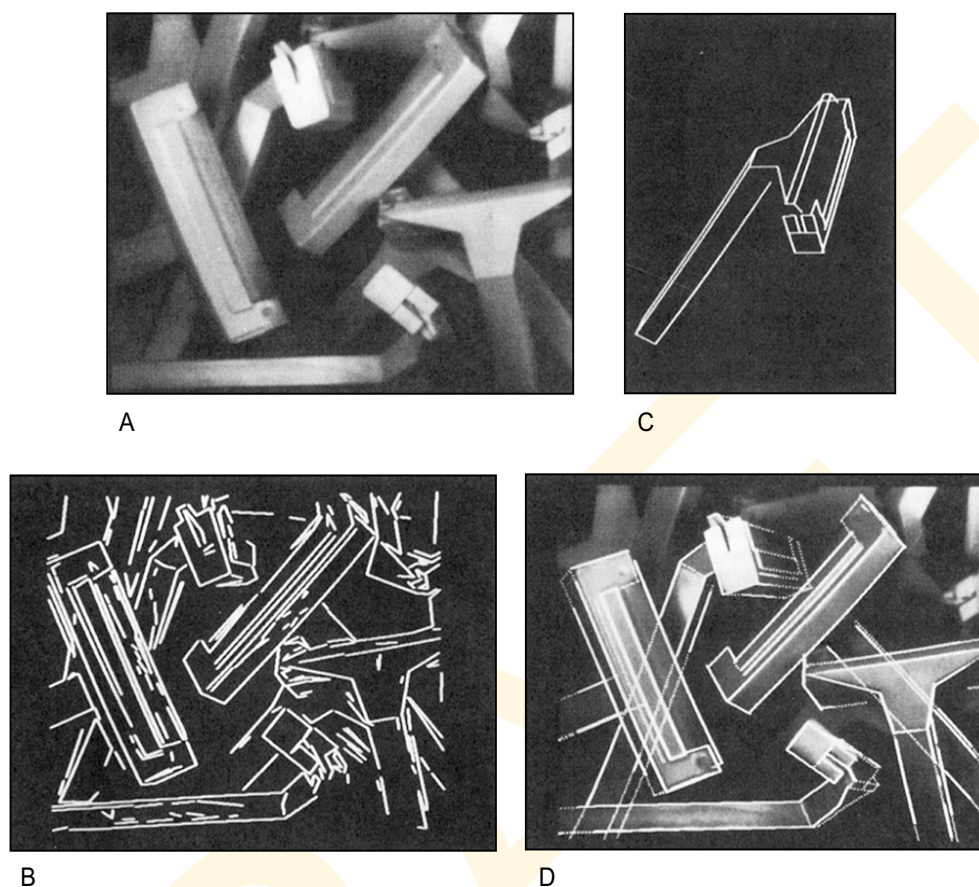
4. Определение уровня, которому проекция эта соответствует в трехмерной модели.

Модель с наилучшим соответствием на шаге 4, т.е. повышающая минимальный порог, идентифицируется затем как наблюдаемый объект.

Все теории — от двухмерных до трехмерных — исходят из того, что все модели известных ранее объектов хранятся в памяти в виде трехмерных представлений пространственных координат, что позволяет им использовать всю структуру проективных трансформаций. Следовательно, мэппинг очевидных признаков двухмерной ретинальной проекции в соответствующие признаки трехмерной модели, а также обратный мэппинг, ограничиваются правилами проективной геометрии. Сии ограничения позволяют этим теориям обойти сложности, проиллюстрированные рис. 9.3.11: четыре изображения будут скорее стыковаться с четырьмя разными трехмерными моделями, соответствующими различным трехмерным формам, нежели с тем же аспектным представлением. Данное усовершенствование — прямое следствие применения проективных ограничений в процессах приводки и трансформации.

Машинные имитации, выполненные канадцем Дэвидом Ловэ (1985), а также Д. Хуттенлочером и С. Ульманом (1987), демонстрируют то, что такой подход может решить проблему распознавания объектов в различных ракурсах (хотя бы в упрощенных условиях). На рис. 9.3.12 дан пример интерпретации изображения моделью Ловэ. В части А дано входное изображение: ведерко, наполненное бритвами. Задача программы в том, чтобы идентифицировать отдельные бритвы и определить их физическую ориентацию.<sup>1</sup> В части В даны исходные линейные сегменты, идентифицированные с применением алгоритма, разработанного на основе метода Марра — Хилдрет (1980) по поиску яркостных краев. Сие используется затем для определения наилучшего ракурса, с которого это изображение стыкуется с трехмерной моделью. Часть С демонстрирует трехмерную проволочную модель *категории* бритв (показано как двухмерное изображение с определенного ракурса), что соответствует линейным сегментам из части В. Часть D — финальное соответствие входного изображения и внутренней модели бритвы в различных ракурсах. Пунктиры в части D указывают на края, чье наличие хоть и было спрогнозировано моделью, но которых не обнаружил поисковик краев низкого уровня. Скорее, они были найдены процессами сравнения «вверх-вниз», когда трехмерная модель проектировалась на двухмерное изображение.

<sup>1</sup> Задача такого плана представляет особый интерес в вопросах машинного зрения, поскольку локализация объектов в промышленных задачах, в частности захват предметов из корзины роботом — это мотив многих и многих исследований.



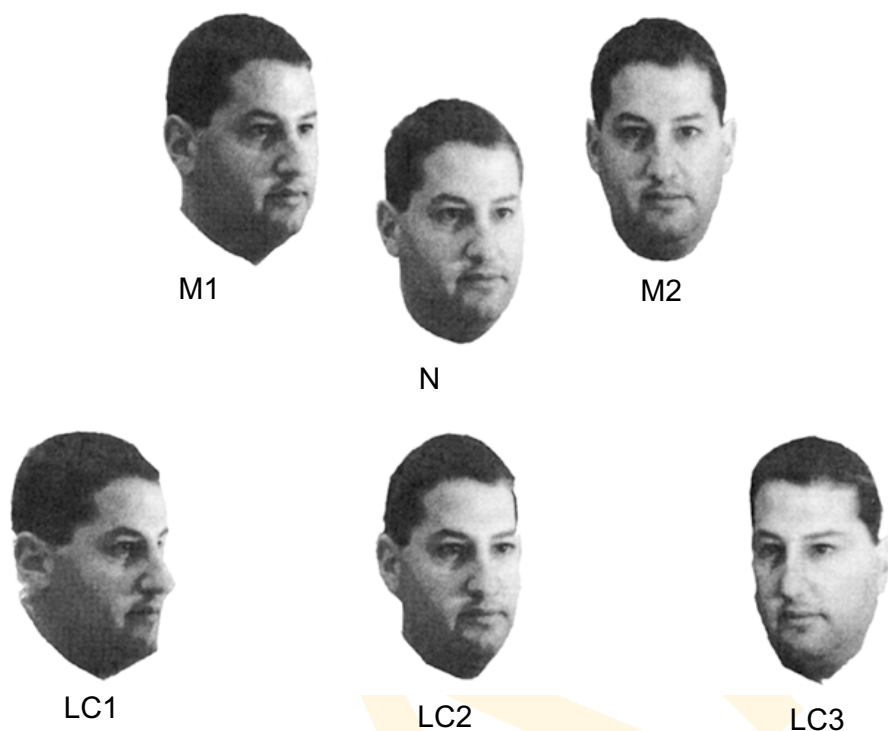
**Рис. 9.3.12.** Работа SCERPO-модели, разработанной Ловэ. В части А дано изображение ведерка с бритвами. В части В показаны края, извлеченные SCERPO из изображения А в строго восходящем режиме. В части С дана трехмерная проволочная модель одноразовой бритвы. В части D — успешный поиск соответствий между краями и различными проекциями внутренней модели бритвы.

По материалам D. Lowe, 1985.

#### *9.3.3.4 Приводка к двумерным ракурсным комбинациям*

Когда вопрос согласования двумерных ракурсов решен, вопрос распознавания двумерных объектов редуцируется до замены функции трехмерной модели малым числом двумерных ракурсов. Ряд недавних количественных исследований позволил существенно продвинуться в этой области (Poggio & Edelman, 1990; Ullman, 1996; Ullman & Basri, 1991). Целью являлся поиск метода, который может произвести новые, уникальные двумерные ракурсы объекта из нескольких сохраненных в памяти, минуя необходимость обращения к трехмерной модели в памяти.

Израильские исследователи Симон Ульман и Ронен Басри (1991) продемонстрировали, что (как минимум в определенных условиях) эта цель достижима, через доказательство того, что все возможные ракурсы объекта можно реконструировать как линейную комбинацию лишь из трех правильно выбранных ортогографических проекций одного и того же трехмерного объекта. На рис. 9.3.13 даны несколько наиболее явных примеров, основанных на этом методе. Два фактических двумерных ракурса лица человека (модель M1 и M2) соединены так, что получались другие двумерные ракур-



**Рис. 9.3.13.** Новые ракурсы, полученные комбинацией исходных изображений. M1 и M2 демонстрируют два ракурса одного и того же лица. N — промежуточный вариант, рассчитанный моделью. LC1, LC2 и LC3 — новые ракурсы, полученные линейной комбинацией M1 и M2 без какой-либо специфической трехмерной модели лица. LC2 соответствует фактическому ракурсу с той же точки наблюдения.

По материалам Ulmann, 1996.

сы того же лица. Один из них — промежуточный ракурс, *интерполированный* между двумя моделями (линейная комбинация LC2), а два других *экстраполированы* за их пределы (линейные комбинации LC1 и LC3). Отметим близкое сходство между интерполированным ракурсом (LC2) и ракурсом фактическим, взятым с соответствующей точки (новый вид N). Однако сие скорее исключительный результат, достигаемый лишь в определенных условиях, и экологический нереальный.

Три ключевых условия Ульмана — Басри (1991) таковы:

1. Все точки объекта должны быть видимыми в любом его ракурсе.
2. Должно быть заранее известным корректное соответствие всех точек по каждой паре ракурсов.
3. Ракурсы должны отличаться только через т.н. *строгую трансформацию* и/или масштабом.

**Первое условие** предполагает, что ни в одном из трех ракурсов у объекта нет закрытых точек. Этому условию вполне удовлетворяют проволочные фигурки, полностью просматриваемые под любым углом. Но обычные объекты удовлетворить ему не могут. Причина очевидно в том, что большинство объектов состоят из непрозрачных материалов, закрывающих все точки вдоль зрительной оси позади себя. Согласимся с тем, что линейные комбинации лица на рис. 9.3.13 оказываются на поверку не столь эффектны, поскольку оказалось, что метод работает только по тем точкам, что видны в обоих ракурсах. То, что получилось — это лишь маска поверхности лица, но не



головы. Разницу можно найти, если тщательно взглянуть в края — голова завершается грубо и неестественно (LC1 и LC3). Отметим, что метод линейной комбинации непригоден для получения портрета в профиль, поскольку задняя часть головы не представлена ни в одной из исходных моделей (M1 и M2). Чтобы линейная комбинация позволила распознавать объекты при всех возможных углах наблюдения, самокроющие объекты должны обладать тремя различными двухмерными моделями по каждому набору ракурсов, в которых видны разные точки. В сравнении с теорией аспект-граф, это значит, что полное представление трехмерного объекта через линейную комбинацию потребует втрое большего числа двухмерных ракурсов, чем число аспектов у объекта. То есть, ежели число аспектов в сложном объекте и без того велико, то число двухмерных ракурсов окажется втрое большим.

**Второе условие** предполагает, что соответствие между точками в сохраненных памяти двухмерных ракурсах известно до объединения ракурсов фактических. Хотя решение проблемы соответствия — это нетривиальное вычисление по сложным объектам, его можно заполучить скорее «off-line», нежели в процессе распознавания.

**Третье условие** означает, что процесс объединения ракурсов не способен дать точной комбинации, если отличия в двухмерных ракурсах включают пластические деформации объекта. К примеру, когда мы видим одного человека стоящим, а других сидящими, их комбинация вовсе необязательно окажется потенциальным ракурсом стоящего человека. Это ограничение затрудняет работу с телом и лицом движущихся существ, равно как и с неподвижными объектами, выполненными из пластичных материалов (к примеру, одежда) или имеющих составную структуру (например, ножницы). Математики, конечно, ищут пути решения этих проблем (см. Ullman, 1996), но всегда следует помнить о жестких ограничениях в логике линейных комбинаций.

Отметим также, что результаты, полученные Ульманом и Басри (1991) указывают на то, что двухмерные ракурсы можно объединить, дабы получить новые в точно определенных условиях, но они ничего не говорят нам о том, как эти ракурсы можно использовать для распознавания объекта из его ретинальной проекции. Дальнейшее развитие методологии линейных комбинаций потребовало разработки механизма наилучшего соответствия между входящим ракурсом и линейной комбинацией ракурса модельного как составляющей процесса распознавания объекта. Один из подходов — это использование небольшого числа признаков для поиска наилучшей комбинации модельных ракурсов. При этом те же признаки можно использовать для привошки ракурса к трехмерной модели. Прочие методы также имеют право на существование, но они слишком сложны для описания на страницах нашей книги.<sup>1</sup>

Т. Поджио и С. Эдельман (1990) предложили некий нелинейный метод, названный ими *генерализованными радиально-базисными функциями (GRBF)* и предназначенный для создания новых двухмерных ракурсов из множества ракурсов, сохраненных памятью. Метод позволяет приблизительно определять неизвестные ракурсы, применяя при этом известные модельные в качестве ограничителя. Основная идея аналогична прочим методам приближения, таким как фитинг гладкой кривой к набору точек: если нам известен набор точек от неизвестной функции, существуют количественные методы поиска самой гладкой кривой, проходящей через них (или хотя бы рядом). В методе Поджио — Эдельмана, ракурсы, сохраненные в памяти, аналогичны точечным данным, и набор неизвестных новых ракурсов аналогичен точкам, расположенным на кривой, фитингующей их. Новые ракурсы, полученные от известных предметов, не

<sup>1</sup> См. Ullman, 1996.



могут быть идеально правильными, но при этом вполне разумной аппроксимацией, лежащей очень близко к известным ракурсам и много более точной, чем кажется на первый взгляд. GRBF-метод можно рассматривать как расширение метода Ульмана — Басри (1991), но который обычно требует более трех ракурсов. С другой стороны, он также позволяет восстанавливать позицию объекта относительно наблюдателя, на что метод линейных комбинаций не способен.

### 9.3.3.5 Минусы

Несмотря на всю эффектность результатов, достигаемых в строго заданных лабораторных условиях, у описанных выше концепций в общем описании ими зрительного распознавания объектов и их классификации имеется букет серьезных изъянов. Пять основных таковы:

1. *Трехмерность структуры.* Аспект-графы и привodka к двумерным комбинациям ракурсов не в силах учесть нашу способность к восприятию трехмерной структуры объектов. В самом деле, лишь бросив беглый взгляд на предмет, мы уже чувствуем, что почти все знаем о его трехмерной структуре, включая то, как нужно протянуть руку, чтобы взять его, и что мы почувствуем, если захотим потрогать его. Как сие возможно, если всё, что есть в нашем распоряжении — это структурированный набор двумерных проекций? Мы можем апеллировать к карте глубин, построенной на стереопсисе или параллаксе движения (как в марровом 2.5D-скетче), но в реальности мы вполне способны к устойчивому восприятию трехмерной структуры объектов, даже если смотрим на них одним, и при этом неподвижным глазом. В каком-то смысле сие не зависит от того, как много различных двумерных интерполяций или аппроксимаций можно воспроизвести из сохраненных двумерных ракурсов или действительно ли они достаточны для распознавания: проблема в том, что они все еще — суть двумерные представления и, следовательно, неадекватны представлению трехмерной структуры.

2. *Новые объекты.* Сравнение двумерных ракурсов с трехмерными моделями в памяти не страдает от трехмерности, поскольку эксплицитные трехмерные модели известных объектов уже хранятся в ней (Huttenlocher & Ullman, 1987; Lowe, 1985; Ullman, 1989). Когда объект распознан, его полная трехмерная структура становится понятной. Но что происходит, ежели объект новый, скажем, такой, как фрагмент абстрактной скульптуры или нереальный объект на рис. 9.1.3? Несомненно, мы воспринимаем такие объекты со всей ясностью и, разумеется, трехмерными, даже если отсутствует предшествующая внутренняя трехмерная модель, соответствующая им. Такие модели должны строиться сугубо из входной зрительной информации, но напроць неясно, как сего достичь.

3. *Пластичные объекты.* Все три ракурсоспецифические теории, с которыми мы давеча познакомились (аспект-графы, согласование двумерных ракурсов с трехмерными моделями и сравнение двумерных ракурсов с двумерными ракурс-комбинациями) нацелены на распознавание ригидных (твердых) объектов. Но что произойдет, если мы, пользуясь этой схемой, попытаемся зрительно распознать человеческое тело? Когда мы начнем сравнивать двумерные ракурсы с внутренними трехмерными моделями, то нам понадобятся сотни различных моделей. Для танцора, который представляется во множестве поз, число взлетает до тысяч. Мы можем аппроксимировать множество таких деформаций через локальные ригидные трансформации подмножеств объекта, как при восприятии человеческой фигуры, но сие неизбежно вводит в игру функциональный эквивалент частей — сложность, от которой все исследователи как раз и пытаются уйти.



Рис. 9.3.14. Дом на водопаде Фрэнка Райта.

4. *Дольчатость структуры.* В конце седьмой главы мы рассуждали о том, что большинство сложных объектов обладают легко читаемой долеговой структурой. Ракурсоспецифические представления мы рассматривали в целом как не содержащие какого-либо эксплицитного представления долей, поскольку они состоят из множеств неотчетливых точек или признаков низкого уровня, таких как края и вертексы. Ульман (1996) показал, что части (равно как и цельные объекты) можно представлять в памяти по отдельности. Сия вполне здравая идея служит нам напоминанием, что долеговая схема распознавания, такая, в частности, как RBC, и схемы ракурсные — отнюдь не исключают друг друга, но вполне могут объединиться в некие гибридные подходы.

5. *Вариации экземпляров.* Ситуации, к которым успешно приложимы ракурсоспецифические теории до сих пор ограничиваются идентичными объектами, у которых меняется лишь угол наблюдения. К примеру, корзина полная бритв (рис. 9.3.12) или распознавание различных ракурсов одного и того же лица. Однако же как обстоит дело с классификацией объектов более вариабельной категории начального уровня, таких как стулья, собаки, дома и пр.? Большое число метрических вариаций различных предметов в большинстве категорий полностью дискредитирует идею ракурсоспецифических теорий в решении вопросов зрительной классификации. Чтобы применять эти методы к категориям начального уровня, последние должны быть представлены в логике конкретных прототипов (Ullman, 1996). Но даже в этом случае маловероятно, что система начнет нормально работать. Насколько разные здания будут соответствовать единственному представлению, вне зависимости от своей типичности или репрезентативности? Давайте попробуем согласовать изображение знаменитого домика

на водопаде Фрэнка Ллойда Райта<sup>1</sup> (рис. 9.3.14) с прототипом дома вообще. Не представляется возможным, чтобы процесс классификации обоих домов прошел на уровне сравнения изображений.

Яростный спор между сторонниками идеи долевого структурного описания и приверженцами концепции ракурсно-пикториальной приводки, усугубился изменением во взглядах у Тарра — Бюльтхоффа (1995) и Бидермана — Герхардштайна (1995) на интерпретацию данных Бидермана — Герхардштайна (1993) о предшествующем прайминге при различных ракурсах. Контраст между двумя подходами часто толкуют как «либо-либо». Однако вполне возможно, что задействуются оба процесса — долевыми и ракурсными — но для решения разных задач (Farah, 1992; Tarr & Bülthof, 1995).

Ракурсоспецифические представления смотрятся вполне подходящими для распознавания одного и того же объекта в различных ракурсах, поскольку в этой ситуации нет вариаций структуры объекта (все отличия между ретинальными проекциями можно истолковать через вариации ракурса). Распознавание специфических объектов представляет трудности для теорий структурного описания, поскольку их представления редко бывают достаточно специфичными для различения отдельных экземпляров. В противоположность сему теории структурных описаний (например, RBC), выглядят вполне подходящими для классификации на начальном уровне, поскольку оперируют более абстрактными представлениями, которые лучше обращаются с вариациями формы различных экземпляров одной категории. Это то самое место, где ракурсоспецифические теории сдают позиции.

Еще один существенный момент заключен в том, что ракурсные и долевыми схемы можно очень эффективно объединить — они не исключают друг друга и могут исполняться параллельно (Hummel & Stankiewicz, 1996). То есть, когда данный ракурс согласуется с ракурсным представлением в памяти, распознавание идет быстрее и точнее; когда нет, классификация опирается на более медленный и более сложный процесс сравнения противоположных структурных описаний. Какое из возможных решений данного конфликта наиболее эффективно, пока еще неизвестно. Остается надеяться на дальнейший научный прогресс.

## 9.4 ИДЕНТИФИКАЦИЯ БУКВ И СЛОВ

До сих пор мы говорили о зрительной классификации обычных трехмерных объектов. Теперь же настало время обратиться к некоей весьма специфической теме — вопросу о том, каким образом наше зрение идентифицирует письменный текст, состоящий из букв и слов. Сие частный случай зрительной классификации, но он невероятно важен в поле современной культуры. Если мы поймем механизм зрительных процессов, лежащих в основе чтения, мы сможем решить массу проблем, например: наконец-то научить студентов скорочтению, помочь больным с нарушениями восприятия письменной речи, научить компьютеры «читать» и т.д. Уже в ходу новейшие технологии, например, системы распознавания текста, но по функциональности им еще очень далеко до зрения человека, да и строятся такие системы на принципах, мало схожих с естественным чтением.

Итак, разговор мы должны начать с очень серьезного замечания: идентификация

<sup>1</sup> Фрэнк Ллойд Райт (Frank Lloyd Wright) — американский архитектор-новатор. Оказал огромное влияние на развитие западной архитектуры в первой половине XX века. Создал «органическую архитектуру» и пропагандировал открытый план в ней. — *Прим. пер.*



букв и слов в письменной речи — это не чтение в обычном смысле. Обычное чтение требует от нас не только идентификации слов, но также понимания их значений в определенном контексте. Восприятие речи — это весьма сложная ментальная функция, требующая как знания значений слов, так и способности когерентно интерпретировать их в рамках контекста. Сие не вопрос перцепции как таковой, но сложная концептуальная проблема. Нам по силам разобрать лишь внешнюю сторону вопроса: процессы идентификации литер и слов, составляющих текст. Как мы увидим далее, даже сие есть незаурядная задача.

Исследования в сфере зрительной классификации букв и слов начались раньше работ в сфере зрительной классификации объектов, отчасти вероятно благодаря ограниченности самого вопроса. Основные упрощения проистекают из следующих свойств текстов:

1. *Двухмерность.* Обычный письменный текст состоит исключительно из знаков, на плоской поверхности. Сие значит, что структура проксимального стимула плотно согласуется со структурой стимула дистального (разумеется, когда поверхность с текстом располагается во фронтальной плоскости или возле нее). Большинство проблем, связанных с восприятием глубины пространства, достижением константности восприятия и завершения частично закрытых объектов — исчезают автоматически.

2. *Комбинаторная структура.* Все индоевропейские тексты состоят из последовательности знаков, принадлежащих одному небольшому множеству. В английском языке 52 буквы (26 прописных и 26 строчных), 10 цифр, плюс некоторое число знаков пунктуации и прочих символов.<sup>1</sup> Сие значит, что знаки эти образуют «натуральное» множество атомарных примитивов, из которых конструируются все слова, фразы и тексты.

Важно отметить, что перечисленные факторы хоть и упрощают вопрос восприятия слов и литер, до окончательного решения вопроса еще очень далеко, и ряд проблем остается, например, — вопрос идентификации букв в различных положениях, при различных размерах и ориентациях. Исходный статус гарнитур как натуральных примитивов, очевидно, способен решить проблему подобно тому, как геоны решают задачи в рамках RBC: действительно, Бидерман говорил об использовании геонов как некоего «алфавита» зримых объектов. Даже если и так, тот факт, что буквы сами по себе — суть сложные объекты, означает, что проблема не решается лишь через собственно алфавит.

### 9.4.1 Идентификация букв

Самоочевидно, что разговор следует начать с вопроса идентификации букв. И чтобы упростить его, будем говорить о восприятии лишь 26 прописных (т.е. заглавных) литер алфавита в их печатном (не рукописном) варианте.

Вспомним, что базовая сложность в представлении формы двухмерных паттернов уже обсуждалась нами в гл. 8. Здесь же мы просто повторим некоторые положения из нее:

1. *Шаблоны.* Стандартные шаблоны крайне неудобны, поскольку для распознавания букв во всех возможных расположениях, размерах и ориентациях понадобится огромное число этих шаблонов. Если гарнитура шрифта известна заранее (как скажем в системах чтения банковских чеков) можно использовать предобработанные шабло-

---

<sup>1</sup> В языках наподобие японского и, в особенности, китайского набор символов радикально больше и исчисляется тысячами.

ны текста, нормированные по размеру, расположению и ориентации. Однако шаблоны нельзя рассматривать всерьез в роли базиса общей теории классификации букв зрением человека, поскольку мы можем читать великое множество гарнитур и их начертаний, включая те, которых ранее никогда не видели.

2. *Признаки.* Главный недостаток признаковых представлений — в нехватке тех признаков, что способны обеспечивать классификацию в разных гарнитурах, а также в сложностях количественной идентификации этих признаков.

3. *Структурные описания.* Как и в теориях объектной классификации структурные описания — суть опция. Их серьезным недостатком является то, что они весьма сложны как сами по себе, так и в сравнении друг с другом.

Концепция признаков предлагалась разными исследователями. Обычно речь шла о долевых и собственных признаках, хотя некоторые теории работали только с долевыми (McClelland & Rumelhart, 1981). В таблице на рис. 9.4.1 дано представление возможного множества признаков прописных букв латинского алфавита. В ячейках таблицы значится встречаемость каждого из признаков каждой из 26 букв алфавита. Отметим, что в нем нет ни одной пары букв с одинаковым набором признаков. Сие необходимо для корректной классификации всех 26 литер. Однако у многих литер имеются общие признаки, например у букв «В» и «D», или «Е» и «F». Нахлест множеств признаков у той или иной пары букв — есть теоретическая мера их прогнозируемого перцептивного подобия, которая, следовательно, должна плотно коррелировать с эмпирической мерой потенциальной спутываемости двух литер при их классификации. Например, как часто наблюдатель будет принимать фактическую «В» за «D» и наоборот? Мера спутываемости может использоваться (и использовалась) при тестировании гипотез представления букв в зрительной системе, хотя результаты так и не позволили выявить оптимальный (правильный) набор признаков.

Прочие теории представления литер допускают непрерывные вариации параметров, что разнит их друг с другом. Теории эти выводятся из процедуры многомерного шкалирования, описанной в гл. 8. Эмпирическая мера подобия литер в паре, такая, в частности, как их спутываемость, проектируется в некое многомерное пространство, в котором каждая литера представлена точкой. Расстояние между точками соответствует неподобию букв.

Один из вариантов непрерывного пространственного представления именуется *Нечеткологической моделью восприятия* — *Fuzzy Logical Model of Perception* — *FLMP* (Massacre, 1989; Massaro & Hary, 1986; Oden & Massaro, 1978). Одним из достоинств модели является ее способность к учету эффектов непрерывного изменения в структуре литер при их классификации. Подобно нечеткологической модели цветовой классификации (см. гл. 3) непрерывный диапазон признаковых значений в FLMP соответствует уровню, на котором данная буква имеет такой-то признак. К примеру, на рис. 9.4.2 дано сколько-то буквоподобных фигур, построенных непрерывным изменением ориентации короткого сегмента и открытости овальной части. Когда сегмент почти горизонтален, а овал открыт (слева вверху или рядом), мы идентифицируем эту фигуру как литеру «G». Когда сегмент однозначно наклонный, а овал полностью закрыт (справа внизу), мы идентифицируем фигуру как букву «Q». Прочие фигуры в представленном массиве неоднозначны и могут восприниматься по-разному.

На рис. 9.4.3 представлены данные эксперимента по классификации, в котором оценивалась доля ответов, когда испытуемые принимали литеру «Q» за «G» (Massaro & Hary, 1986). Данные демонстрируют порядковую прогрессию от G- к Q-откликам.

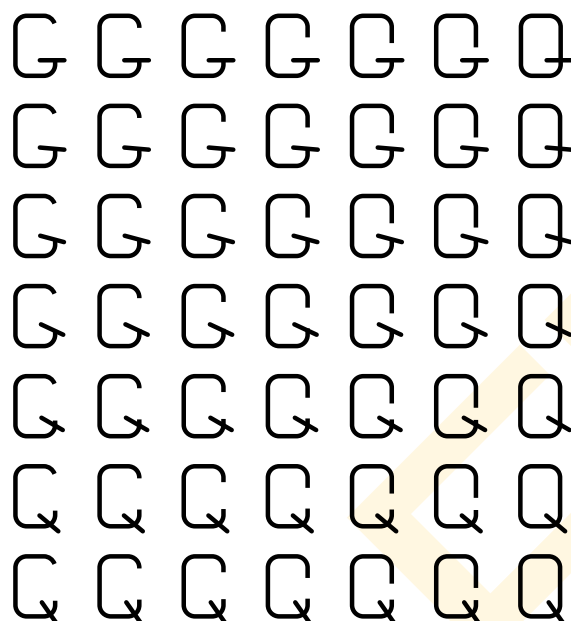
Результат сложно истолковать в рамках сугубо бинарного подхода. Один из способов — это предположить, что величины промежуточных признаков зафиксированы



	Вертикальные линии	Горизонтальные линии	Косые линии	L-связи	T-связи	X-связи	Свободные концы	Искривленные порции	Закрытые порции	Зеркальная симметрия
A	0	1	2	1	2	0	2	0	1	1
B	1	3	0	2	1	0	0	2	2	1
C	0	0	0	0	0	0	2	1	0	1
D	1	2	0	2	0	0	0	1	1	1
E	1	3	0	2	1	0	3	0	0	1
F	1	2	0	1	1	0	3	0	0	0
G	1	1	0	1	1	0	3	1	0	0
H	2	1	0	0	2	0	4	0	0	2
I	1	0	0	0	0	0	2	0	0	2
J	1	0	0	0	0	0	2	1	0	0
K	1	0	2	0	2	0	4	0	0	0
L	1	1	0	1	0	0	2	0	0	0
M	2	0	2	3	0	0	2	0	0	1
N	2	0	1	2	0	0	2	0	0	0
O	0	0	0	0	0	0	0	1	1	2
P	1	2	0	1	1	0	1	1	1	0
Q	0	0	1	0	0	1	2	1	1	0
R	1	1	1	1	2	0	2	1	1	0
S	0	0	0	0	0	0	2	2	0	0
T	1	1	0	0	1	0	3	0	0	1
U	2	0	0	0	0	0	2	1	0	1
V	0	0	2	1	0	0	2	0	0	1
W	0	0	4	3	0	0	2	0	0	1
X	0	0	2	0	0	1	4	0	0	2
Y	1	0	2	0	0	0	3	0	0	1
Z	0	2	1	2	0	0	2	0	0	0

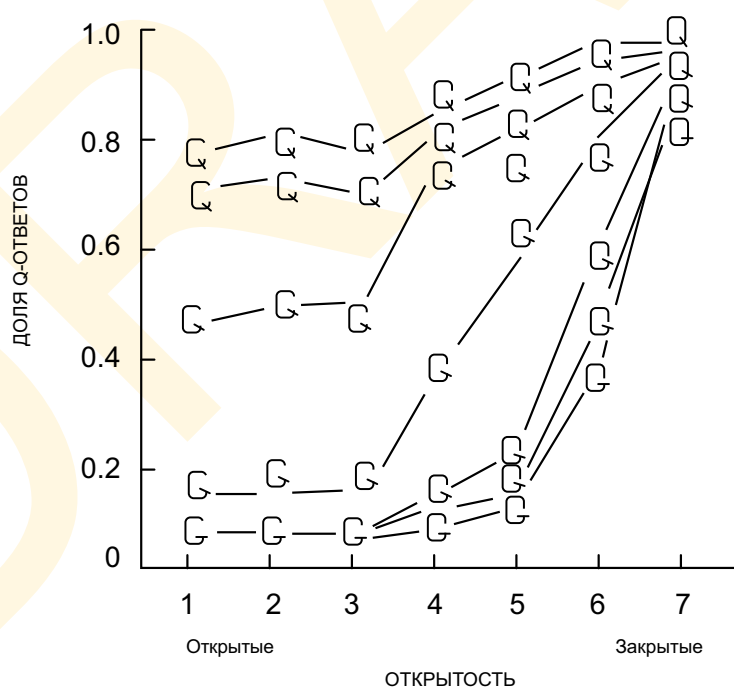
**Рис. 9.4.1.** Возможный набор признаков для распознавания прописных букв английского алфавита. Встречаемость каждого признака в данной букве представлено в соответствующих ряду и колонке. Специфические признаки, показанные здесь, не столь иллюстративны. Отметим, что не существует пары литер с одним и тем же набором признаков в каждой.

однозначно в каждой из попыток как некая величина, и что вероятность каждого значения меняется как функция от параметров стимуляции. В то время как сие прогнозирует ту же общую картину результатов, что и теория нечетких множеств, не удается поймать неопределенные фигуры. Если каждый овал воспринимается как открытый или закрытый и каждый сегмент либо как горизонтальный, либо как косой, то каждая из фигур в данный конкретный момент времени должна восприниматься однозначно как «G» или «Q», или же не как буква (рис. 9.4.2). Но всё иначе, по крайней



**Рис. 9.4.2.** Неопределенные G/Q-стимулы. Семь уровней наклона (меняется по рядам) линейного сегмента объединены с семью уровнями зазора (меняется по колонкам), образуя 49 литеро-подобных стимулов, варьирующих от прототипической G (слева вверху) до прототипической Q (справа внизу).

По материалам Massaro & Hary, 1986.



**Рис. 9.4.3.** Результаты G/Q-эксперимента. Доля испытуемых, идентифицировавших 49 стимулов с рис. 9.4.2. как «Q» дана как функция от наклона линейного сегмента и величины зазора. Линии демонстрируют усредненный прогноз нечеткологической модели восприятия (FLMP).

По материалам Massaro & Hary, 1986.

HWO NMYA RSTELTE NCA OYU RPTERO NWO?

HOW MANY LETTERS CAN YOU REPORT NOW?

**Рис. 9.4.4.** Неформальная демонстрация эффекта словесной доминанты. Англоязычный читатель легко найдет большее число букв в нижней строчке, несмотря на то, что число их то же самое.

мере в нормальных условиях просмотра. Сие роняет ценность бинарного признакового представления как адекватного восприятию этих фигур.

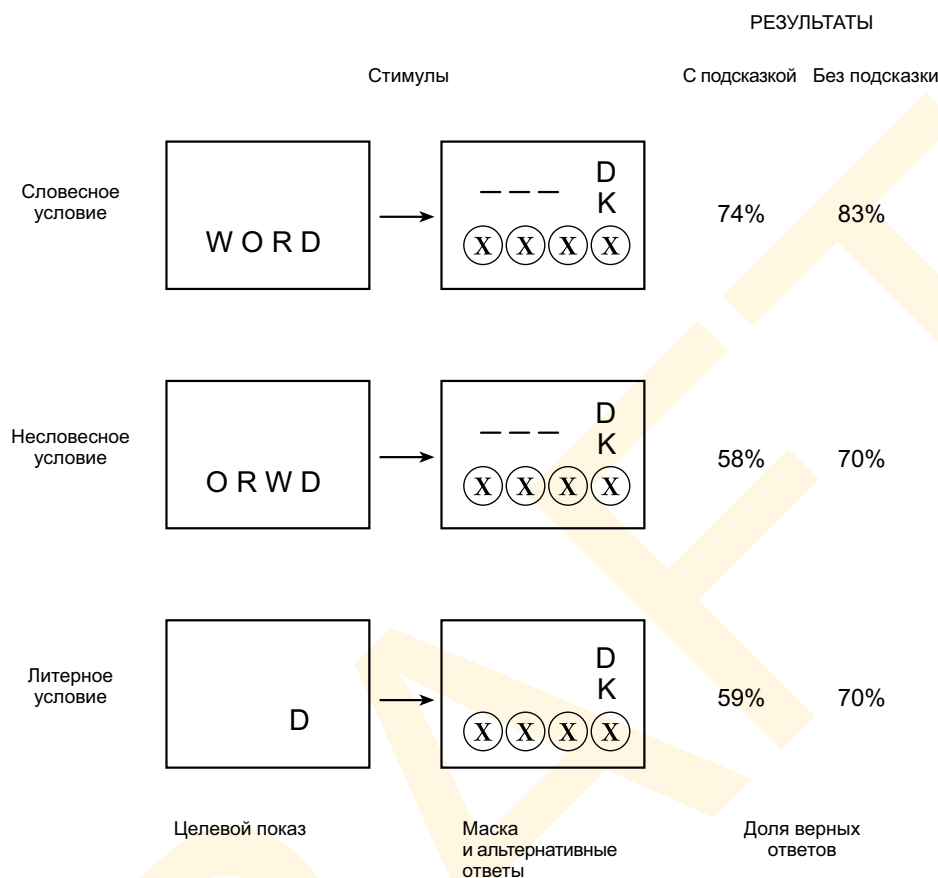
Недостаток же «нечеткого» подхода в том, что, как и многие другие концепции, он не полагает множества четко определенных признаков, адекватного восприятию литер нашим зрением. Скорее он определяет абстрактный репер представления литер в логике каких-то идентифицируемых признаков, с деталями, которые нужно дополнить в дальнейшем. Сие неприемлемая теоретическая ситуация в науке.

#### 9.4.2 Идентификация слов и букв внутри слов

Комбинаторная структура текстового материала указывает на то, что буквы — это естественные примитивы более крупных единиц — слов. Логически сие значит, что при классификации текстов буквы могут работать как независимые единицы, каждая из которых идентифицируется отдельно от остальных. Однако психологически сие не так. Классификация букв фактически зависит от контекста, т.е. от других букв, окружающих данную литеру, подобно тому, как классификация объекта зависит от окружения его другими объектами. Пример, данный на рис. 9.2.12 («THE CAT») — великолепная демонстрация сказанному.

Тот факт, что литеры легче и быстрее распознаются, когда находятся в составе реального слова (т.е. сущности, нагруженной смыслом), нежели в составе некоей бессмысленной последовательности — легко продемонстрировать. Давайте просто посмотрим на верхнюю строчку рис. 9.4.4 в течение 5 с и попытаемся назвать столько букв, сколько сможем. А затем сделаем то же с нижней строчкой. Мы сможем назвать намного больше литер из нижней строки, поскольку состоит она из смыслогруженных слов, образующих фразу, тоже несущую некий смысл. Верхняя же строчка — это перемешанный вариант нижней. Отличие в результатах опыта принято толковать как проявление *эффекта словесной доминанты*, о котором еще более века назад говорил известный американский психолог Джеймс Кэттелл (1886), сравнивший число литер, которые смогли увидеть испытуемые в течение 10 мс в случайном и словесном показах.

К сожалению, имеются сложности в интерпретации результатов экспериментов, доказывающих то, что мы лучше *идентифицируем* буквы, когда они в составе знакомых слов — эффект может заключаться в нашей способности *помнить* буквы. Разумеется, мы способны идентифицировать все литеры в обоих случаях, способны к группировке букв в словах, но оказываемся неспособными запомнить их в случайной строке. Вторая сложность состоит в том, что эффект словесной доминанты реализуется тактикой догадок, которая, скорее, плод неполного восприятия, нежели эффективной классификации. У нас есть очевидная склонность именно *догадываться* о том, какие слова были предъявлены в течение очень короткого времени (в противовес случайным строкам).



**Рис. 9.4.5.** Формализованный опыт, демонстрирующий эффект словесной доминанты. За коротким показом слова, анаграммы или единичной литеры следовала маска и двухальтернативный принудительный выбор. Испытуемые давали много более точный ответ в первом случае (слово).

Действительно, испытуемые угадывают слова намного чаще, чем случайные строки, даже при очень коротком показе (Goldiamond & Hawkins, 1958).

Однако от перечисленных выше проблем можно уйти. Одно из основополагающих исследований в этой сфере было выполнено американским психологом Геральдом Рейчером (1969) и с тех пор неоднократно повторялось и развивалось множеством ученых (например, Wheeler, 1970). В эксперименте Рейчера испытуемым предъявлялось либо слово (например, WORD), либо сравнимый с ним случайный набор букв (например, ORWD) — см. рис. 9.4.5. Сей целевой стимул предъявлялся на очень короткое время (50 мс), а затем следовал маскирующий паттерн, занимавший всё поле зрения и содержащий критичные литеры. Сразу над маской располагались литеры «D» и «K», одна над другой. В каждой попытке испытуемые должны были указать, какую из двух литер фактически предъявляли в той или иной позиции. В половине попыток перед тестовым стимулом испытуемым предлагалось два варианта ответа. Во второй половине варианты предлагались только в тестовом показе, после предъявления целевого стимула.

Отметим, что в условиях предпоказа к содержанию памяти формально не предъявлялось никаких требований. У испытуемых в памяти условно не было ничего, кроме двух вариантов литер. Даже в ситуации постопоказа загрузка памяти была минималь-

ной. Отметим, что в опыте контролировалась тактика угадывания, основанная на дилемме «слово/не слово», поскольку обе альтернативы формировали схожие слова с тремя другими литерами (WORD и WORK). Но даже в этом случае Рейчер продемонстрировал мощное превосходство словесного контекста над несловесным (см. рис. 9.4.5). Итак, неизбежный вывод таков, что на классификацию литер действительно оказывает влияние контекст: литеры в составе слов идентифицируются точнее, чем литеры в случайной комбинации. Сей феномен часто называют *словесно-псевдословесным*.

Ко всему прочему Рейчер включил в опыт еще одно условие, в котором в качестве целевой предъявлялась отдельная литера (рис. 9.4.5). Очевидно, что единичная литера воспринималась бы точнее, чем любая строчка из четырех букв, слов, либо псевдослов, просто потому, что обрабатывается только одна фигура, а не четыре. Удивительным явилось то, что единичная целевая буква воспринимается менее точно, чем в контексте слова. Сие открытие часто звучит в литературе как *эффект слово-буква*. Явление весьма странно, особенно учитывая то, что результаты многих опытов указывают на то, что, чем больше литер присутствует в случайном множестве, тем слабее результат. Короче говоря, мы опишем коннекционистскую теорию распознавания букв и слов, учитывающую сей эффект наравне с эффектом словесно-псевдословесным (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982).<sup>1</sup>

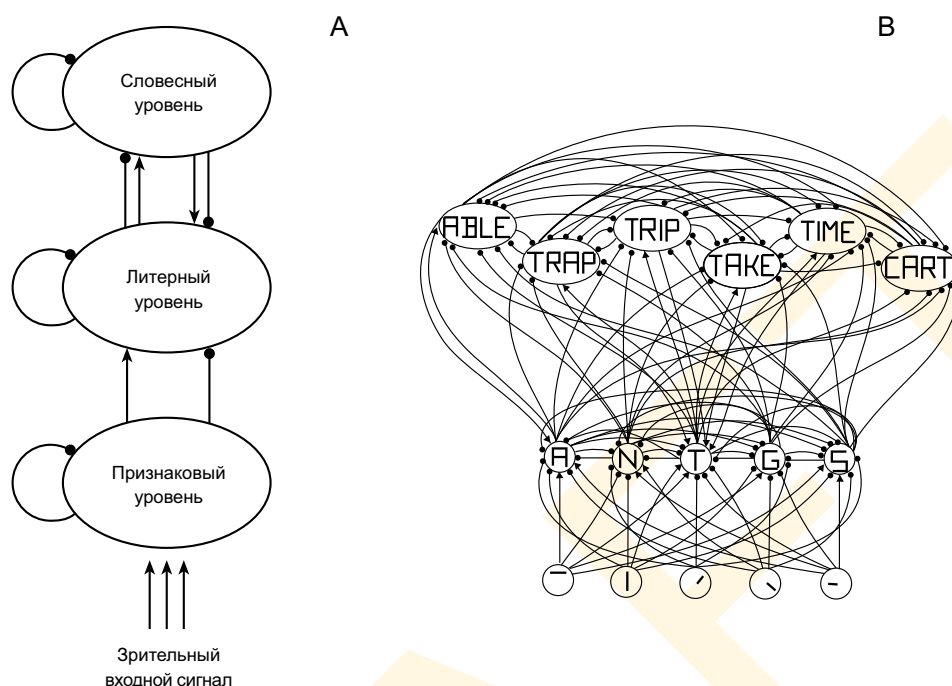
Учитывая результаты Рейчера, мы задаемся вопросом о том, действительно ли восприятие букв в словах до определенной степени сходно с восприятием их в словоподобных комбинациях? Действительно, некоторые псевдослова легко произносятся (например, GRAP), тогда как другие — наоборот (например, RPGA), следовательно, буквы в легко читаемых псевдословах идентифицируются точнее, чем те же буквы в читаемых трудно (McClelland & Jonson, 1977). Сей факт можно объяснить силами коннекционистской модели восприятия букв и слов, которую мы только что разобрали.

### 9.4.3 Модель интерактивной активации

Одной из наиболее претенциозных и известных теорий восприятия литер и слов является теория Мак-Клелланда — Румерхальта (1981), получившая название *модели интерактивной активации* — IA (Interactive Activation). Модель позиционирует многослойную нейронную сеть, состоящую из нейроноподобных узелков и синапсopодобных связей между этими узелками (см. приложение В). Входной зрительный сигнал активирует первый слой узелков (входные «признаковые» узелки), и активация эта распространяется по сети через связи между этими узелками. Активация идет до уровня, зависящего как от уровня активации во всех узелках, с которыми есть связь, так и от силы (веса) этих связей, которые могут быть либо возбуждающими, либо тормозящими. Возбуждающие связи повышают активацию в своем узелке, ингибиторные, соответственно, понижают. Поскольку модель несет в себе петлю обратной связи, паттерн активации динамически меняется во времени, приходя, в конечном итоге, к стабильной картине, в которой часть узелков высокоактивна, а остальные «молчат». Цель IA-модели в имитации перцептивных процессов, лежащих в основе классификации литер и слов, когда аналогичная картина активации в своем стабильном состоянии соответствует реальному восприятию в сходных условиях.

<sup>1</sup> Имеются некоторые аргументы в пользу того, что эффекты «слово-буква» и «словесно-псевдословесный» возникают как следствие работы одного и того же механизма, хотя, по идее, должно быть наоборот. Принцметал и Сильверс (1994) показали, к примеру, что «слово-буква» требует наличия маски, тогда как «словесно-псевдословесный» эффект — нет.





**Рис. 9.4.6.** Архитектура модели интерактивной активации. В части А дана общая схема, в которой единицы, представляющие признаки, литеры и слова соединяются через возбуждение и торможение. В части В показано небольшое подмножество отдельных единиц и связей сети (подробнее в тексте).

По материалам McClelland & Rumelhart, 1981 и 1986.

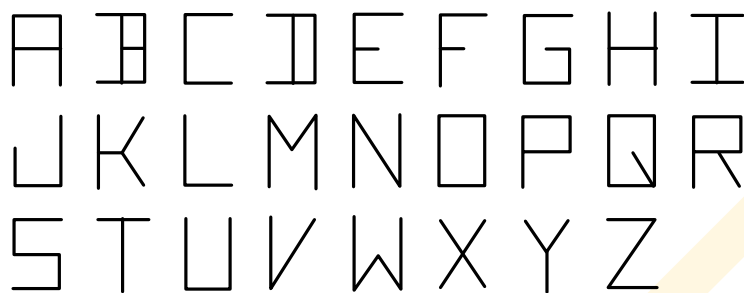
Для имитации классификации литер и четырехлитерных слов сеть структурирована на три отличных друг от друга слоя узелков: *признаковый*, *литерный* и *словесный* (см. рис. 9.4.6). Итак, рассмотрим архитектуру этой сети.

#### 9.4.3.1 Признаковый уровень

Первый слой несет в себе частепризнаковые узелки, каждый из которых представляет определенный отрезок в определенной позиции. Множество возможных отрезков в определенном шрифте, имитированное IA-моделью, дано на рис. 9.4.7. Все литеры собраны в некое подмножество из 12 возможных отрезков. Таким образом, в признаковом слое для каждой позиции, в которой может появиться буква, имеются 12 узелков. 48 результирующих признаковых узелков, активируются входным зрительным сигналом, когда соответствующий им отрезок представлен в целевом стимуле. Они передают свою активацию узелкам во втором слое через возбуждающие, либо тормозящие связи. Отметим, что существует обратная связь к признаковому слою от узелков слоя вышестоящего: признаки обрабатываются строго по вертикальному направлению.

#### 9.4.3.2 Уровень литер

Во втором слое находятся узелки литер, каждый из которых представляет наличие одной из 26 букв в одной из четырех возможных позиций. 104 ( $26 \times 4$ ) узелка принимают возбуждающие связи от всех признаковых узелков, что представляет отрезки, которые есть у данной буквы. К примеру, узелок, отвечающий за верхний горизонталь-



**Рис. 9.4.7.** Шрифт, использованный в модели интерактивной активации. Каждая литера собрана как подмножество из 12 возможных сегментов внизу рисунка.

По материалам McClelland & Rumelhart, 1981.

ный отрезок, возбуждает узелки по всем литерам, у которых имеется данный признак в гарнитуре данного типа: А, В, С, D, Е, F, G, I, О, Q, R, S, Т и Z. Литерные узелки также имеют ингибиторные связи от всех признаковых узелков, представляющих отрезки, которых нет в букве. То есть, узелок, отвечающий за верхний горизонтальный отрезок подавляет узелки по всем остальным литерам данного шрифта — Н, J, K, L, М, N, U, V, W, X и Y. Сетевой эффект этих связей в том, что когда только отрезки в данной букве (скажем, А) представлены в стимуле, соответствующий А-узелок на литерном уровне окажется наиболее активным. Прочие узелки тоже окажутся активированными до уровня, зависящего от того, как много отрезков наличествуют в литере А. Узелок «Н», к примеру, будет весьма активен, поскольку отличается от А (в данной конкретной гарнитуре) только отсутствием вершинного отрезка (см. рис. 9.4.7).

Для усиления активации на литерном уровне до той степени, когда один лишь узелок правильной литеры остается наиболее активным (а вся сеть выходит в стабильное состояние), все узелки, представляющие литеры в одной и той же пространственной позиции, взаимно подавляют друг друга. Взаимное подавление образует сеть типа «Победителю достается всё» внутри каждой из четырех двадцатишестиузелковых множеств. Сие гарантирует то, что если паттерн активации единожды «угомонился», корректная литера в каждой позиции будет единственной при достаточной активации, подавляя оставшиеся 25 конкурентов сильнее, чем они подавляют ее.<sup>1</sup>

На самом деле активация литерных узелков проходит намного сложнее, чем мы только что показали, поскольку они также получают обратный сигнал с уровня словесного. Мы рассмотрим сей источник активации после того, как опишем собственно структуру слова.

<sup>1</sup> Однако модель может давать сбои, особенно, если представление литер коротко и/или зашумленно.

### 9.4.3.3 Словесный уровень

Высший уровень модели представляет четырехлитерные слова и несет в себе более 1000 узелков, каждый из которых соответствует различным английским словам. Узелки словесного уровня получают возбуждающий сигнал от узелков, представляющих каждую из четырех литер: одну в каждой из пространственных позиций. Узелок, отвечающий за литеру «А» в первой позиции, к примеру, активирует словесные узелки слов «ABLE» и «ACTS» (среди прочих), тогда как узелок, отвечающий за «А» во второй позиции возбуждает «BACK» и «GAVE» (среди прочих). Словесные узелки получают также и тормозящий сигнал от узелков, представляющих оставшиеся 25 литер во всех позициях. Таким образом, узелок, отвечающий за «А» в первой позиции подавляет «BACK» и «GAVE», тогда как аналогичный узелок во второй позиции тормозит «ABLE» и «ACTS». Сии связи литерных узелков с узелками словесными гарантируют, что при полноте информации, правильный узелок окажется активнее других. Наивысшую активацию в финальном устойчивом состоянии сети даст только верно идентифицированное слово, подавив при этом другие словесные узелки сильнее, чем они ее.

### 9.4.3.4 Обратная связь

Поскольку трехслойная сеть, рассматривавшаяся нами до сих пор, классифицирует литеры в разных позициях независимо, она не в силах прогнозировать контекстные феномены, такие как эффект словесной доминанты или эффект «слово-буква». Чтобы научить этому IA-модель, Мак-Клелланд и Румельхарт (1981) предположили, что существует некая обратная связь, идущая от словесного уровня, назад к уровню литер. В частности, они добавили возбуждающие связи от узелков слов к узелкам букв, то есть, к примеру, узелок слова «ABLE» возбуждает узелки литер А, В, L и Е, соответственно, в первой, второй, третьей и четвертой позициях. При этом не существует тормозной обратной связи от слов к литерам.

Благодаря возбуждающим обратным связям IA-модель становится способна имитировать несколько важных контекстных эффектов. Вначале рассмотрим феномен словесной доминанты, суть которого в том, что литеры воспринимаются вернее в составе слова, нежели внутри случайной комбинации букв. Модель имитирует сие за счет того, что узелок, представляющий слово, жестко активируется, а затем посылает обратный активирующий сигнал на правильный узелок литеры, усиливая тем его активацию. В случайной комбинации букв не имеется соответствующего словесного узелка и, соответственно, литерный узелок не получает обратного сигнала со словесного уровня. Сие отличие в активации на литерном уровне позволяет модели идентифицировать буквы в словах быстрее, чем в случайной комбинации литер.

Поскольку IA-модель несет в себе узелки только по реальным четырехбуквенным словам, мы можем подумать, что сеть не может прогнозировать псевдоэффект словесной доминанты, при котором в произносимых псевдословах (например, «MAVE») могут восприниматься быстрее и точнее, чем те же буквы в их случайной комбинации (например, «AEMV»). Но раз не существует узелка ни для «MAVE», ни для «AEMV», то за счет чего же возникает отличие? Однако было показано, что модель дает отличие, очень сходное с тем, что мы наблюдаем в нашем восприятии: работа с произносимыми псевдословами находится где-то по середине между словами настоящими и случайной комбинацией букв. Причина в том, что даже несмотря на отсутствие узелка, отвечающего за «MAVE», эта строчка литер частично активирует многие словесные узелки, содержащие три из четырех букв (например, CAVE, HAVE, MATE, MAKE), и

сия частичная активация также дает обратный сигнал на литерный уровень. Ни одна из активаций не оказывается столь же сильной, как у реальных слов, скажем, *SAVE*, но сходный эффект имеет место. Таким образом, с тем же успехом мы имитируем и псевдословесный эффект.

Обратная связь со словесного уровня на литерный позволяет также имитировать эффект «слово-буква». Литеру можно идентифицировать точнее в реальном слове, нежели когда она представлена самостоятельно, поскольку единичная буква не получает подкрепляющей обратной связи со словесного уровня. Больше того, имитационный потенциал модели при работе с единичной литерой сопоставим с таковым при работе с псевдословом, поскольку в обоих случаях не существует очевидной обратной связи со словесного уровня, влияющей на активацию литерных узелков. Сей результат четко согласуется с реальным восприятием.

#### 9.4.3.5 Минусы

Несмотря на успех в прогнозировании контекстных эффектов, работа модели не лишена недостатков. К примеру, Массаро (1989) показал, что имеет место несоответствие некоторым экспериментальным данным. Однако Мак-Клелланд (1991) фактически оспорил сей тезис, сосредоточившись на той дилемме, что действительно ли контекст фактически влияет на сенсорное подтверждение наличия данной литеры (как то исполняет IA-модель через петлю обратной связи от словесного уровня к литерному), или же контекстная информация объединяется с информацией сенсорной по данной литере на некоем более высоком уровне? В нечеткологической модели Массаро (FLMP — Fuzzy Logical Model of Perception), которую мы коротко описали прежде, сенсорная информация о целевой литере исходно отделена от информации контекстной. Сии два независимых источника информации объединяются затем на более поздней интегрирующей стадии, индицируя то, насколько данная литера представлена в основе сенсорной и контекстной информации. Ключевое отличие в том, что в FLMP сенсорная информация не меняется под действием контекста, в отличие от модели IA.

Описанная проблема весьма сложна, поскольку мы не можем заглянуть в голову наблюдателю, проверить уровень активации в литерном представлении и увидеть, оказал контекст влияние или нет. Скорее мы пытаемся лишь угадать то, что происходит в этой самой голове посредством непрямых психологических измерений (такого рода предположения всегда опираются на допущения сомнительной точности). Текущая ситуация в науке такова, что из-за принципиальных разногласий между двумя типами теорий, остается непонятным, как интерпретировать экспериментальные данные. Возможно, новые опыты прольют свет на истинность обеих концепций.

Однако безотносительно к исходу теоретических споров факт существования контекстных эффектов в зрительной классификации литер внутри слов и объектов внутри сцен, не вызывает сомнений. Эти эффекты позволяют нашему восприятию использовать ограничения, налагаемые общей структурой входного зрительного сигнала. Не все комбинации объектов равновелико вероятно могут возникнуть в данной сцене, и не все комбинации литер равновелико вероятны в словах. Зрительная система человека чувствительна к контекстной информации и использует ее при классификации входного зрительного сигнала. Результат быстрее и эффективнее по тем сценам и словам, что часто встречаются в повседневной жизни, но ценой более медленной и менее эффективной обработки случайных наборов объектов и букв. Однако сие очень выгодное эволюционное достижение, поскольку нормальные сцены и слова, по определению, встречаются нам гораздо чаще, чем случайные или необычные сочетания.