

УДК 004.922, 004.932

О СТАБИЛЬНОСТИ ЗРИТЕЛЬНОГО ПОЛЯ И СФЕРОИДАЛЬНЫХ РЕТИНОИДАХ

А. М. Ковалев

*Институт автоматизации и электрометрии СО РАН,
630090, г. Новосибирск, просп. Академика Коптюга, 1
Конструкторско-технологический институт научного приборостроения СО РАН,
630058, г. Новосибирск, ул. Русская, 41
E-mail: amkov@ngs.ru*

Предлагается структура кратковременной памяти — сфероидальный трёхмерный ретиноид — для стабилизации широкоугольного зрительного поля, реализации процесса стереопсиса, формирования и хранения бинокулярного визуального пространства в сознании наблюдателя.

Ключевые слова: ретиноид, ретинотопия, стереопсис, визуальное пространство.

Введение. Одной из важнейших функций мозга является зрительное восприятие целостной, предметной и константной окружающей среды. В психологии проблема константности зрительного поля традиционно делится на две: пространственную (где?) и предметную, или фокусную (что?) [1, с. 185–187]. Первая проблема связана с движением самого наблюдателя, его глаз, головы и тела. Решение её призвано объяснить, почему зрительное поле остаётся стабильным при постоянном изменении сетчаточных изображений, вызванном, например, движением глаз (саккады), наклонами головы (торсии) и т. д. Геометрия константного пространства показывает, где расположены предметы, их размеры и удалённость. Вторая проблема связана с движением самих предметов в пространстве. Её решение помогает ответить на вопрос, как человек узнаёт предметы несмотря на то, что их изображение на сетчатке меняется. Это инвариантное распознавание: что за предмет?

В работе [2] показано, что механизмы константности зрительного поля, называемые также «непредметными механизмами инвариантности» (НМИ), являются компенсационными. Они реализуют «обратные преобразования» и осуществляют отображения сигналов с сетчаток, полученных после локального рецепторного анализа, на трёхмерный «константный зрительный экран» (КЗЭ) в строгом соответствии с движениями наблюдателя. Информация о движениях поступает от глазодвигательной и вестибулярной систем. Константный зрительный экран обладает памятью. Включение вестибулярного аппарата стирает экран, и новое состояние определяется на основе сигналов от глазодвигательной системы.

Проблему деятельности НМИ и КЗЭ в нейрофизиологическом и психофизическом плане попытался разрешить А. Трегуб [3, 4]. Предложена нейроструктура кратковременной памяти, её базовым элементом является специальный тип автаптических нейронов с положительной обратной связью на собственный аксон. С помощью таких нейронов и промежуточных нейронов возбуждения/торможения создаются автаптические ячейки, которые образуют динамическую память, названную ретиноидом. Построены схемы двумерных и трёхмерных ретиноидов, способные решать когнитивные задачи, в том числе трёхмерное представление зрительного поля, размещение и представление «себя» в среде, своего движения и селективного внимания.

В [5] на основе анализа функций ретиноидов и определения геометрии визуального пространства (ВП) отмечен ряд недостатков предложенной модели. Трёхмерное изображение

формируется на множестве фронтопараллельных плоскостей. Удаление от наблюдателя пропорционально линейному параллаксу плоскости. Поскольку оптическая система глаза не отличается изопланатизмом, модель может работать при малых полях зрения ($\sim 30^\circ$). Характерно большое время заполнения ретиноидов, связанное с тем, что при обновлении памяти учитывается только фовеальная часть поля ($2-3^\circ$). Нарушен психофизический закон Эммерта из-за вырожденного проективного преобразования.

Целью данной работы является устранение недостатков модели Трегуба путём её модернизации, направленной на обеспечение стабильности широкоугольного зрительного поля внутри сфероидального визуального пространства.

Сферическая координатная система. Естественное требование к сенсорной поверхности — сохранение формы при перемещении и/или вращении проекций предметов или самой поверхности, т. е. соблюдение геометрических принципов транзитивности движения и вращения, подобных «постулатам движения» в евклидовой плоскости. Набор поверхностей, на которых возможны движения фигур без деформации с тремя степенями свободы, в евклидовом пространстве ограничен — это только плоскости и сферы; им соответствуют два типа двумерной геометрии: плоская и сферическая, как говорят, евклидова и риманова. И лишь в геометрии Лобачевского для «свободного» движения фигур к сферам добавляются эквидистантные и предельные поверхности [6].

Будем считать, что поверхность сетчатки является, скорее, сферой, чем плоскостью. Из известных систем координат [7] наиболее удобна для организации «параллельных» вычислений диспаратностей и расстояний эпиполярная система координат Гельмгольца (рис. 1). Ось x соединяет центры вращения глаз и полюсы полусфер, имитирующих сетчатки. Оси z параллельны носозатылочной оси, лежащей в медиальной плоскости головы. Оси y перпендикулярны первым двум. Эпиполярные плоскости, или плоскости визирования, проходящие через ось x , определяются двугранным углом θ . Считается, что $\theta = 0$ для эпиполярной плоскости, содержащей оси z . Множество эпиполярных плоскостей θ пересекает «сферическую» сетчатку вдоль множества эпиполярных меридианов θ . Круглые конусы определяются углом φ между образующей конуса и координатной плоскостью yz . Множество круглых конусов φ пересекает сетчатку, образуя множество широт φ .

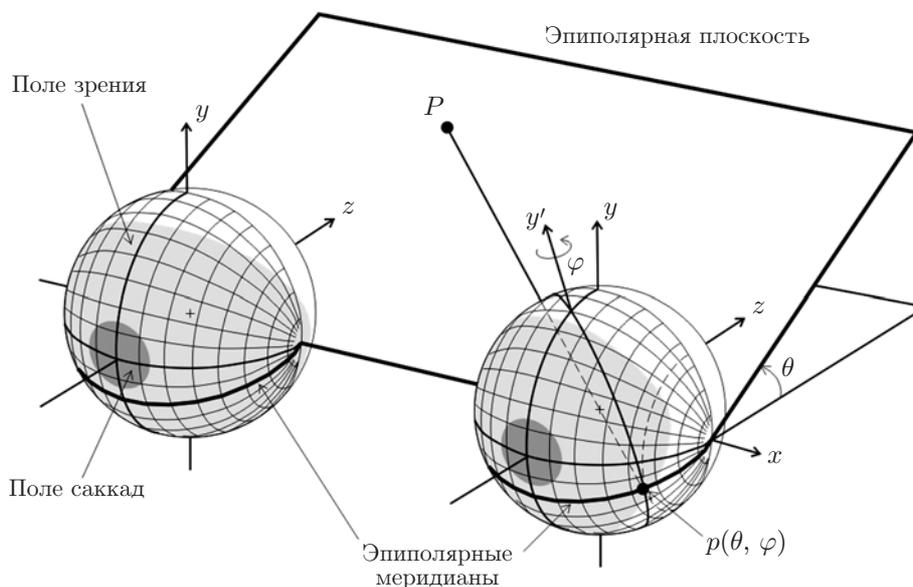


Рис. 1. Система координат Гельмгольца

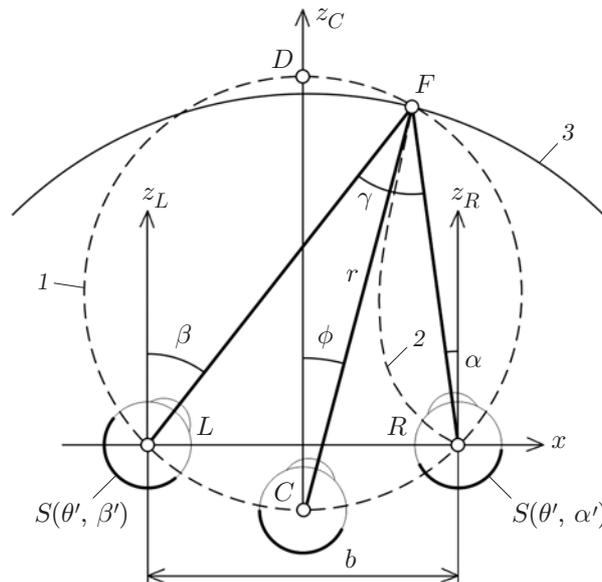


Рис. 2. Триангуляционный метод: 1 — изовергентный круг ($\gamma = \text{const}$), 2 — гипербола Гиллебранда ($\phi = \text{const}$), 3 — окружность ($r = \text{const}$)

На рис. 1 показана проекция точки P в точку $p(\theta, \varphi)$ для исходного положения сетчатки (область фовеа на оси z). Если необходимо совместить фовеальную часть сетчатки с точкой $p(\theta, \varphi)$, то достаточно выполнить два последовательных поворота глаза: вокруг оси x на угол θ (при этом ось y займёт позицию y') и вокруг оси y' на угол φ . Также на рисунке показаны поля зрения и поля саккад, или скачкообразных вращений глаза. При углах поворота не более $15\text{--}20^\circ$ типичны повороты глаз, состоящие из двух и более скачков с промежуточными фиксациями, которые не осознаются субъектом. При больших углах поворота до 60° скачки, как правило, сопровождаются поворотом головы. Минимальная амплитуда произвольных скачков $1\text{--}30'$ [8].

Триангуляция. Основным методом определения абсолютных диспаратностей и расстояний в бинокулярной зрительной системе является триангуляционный метод. На рис. 2 в эпиполярной плоскости θ показан треугольник LFR , где L и R — центры вращения левого и правого глаза; $LR = b$ — глазная база; F — точка рефлекторной фиксации взгляда. Условие пересечения зрительных осей в точке фиксации F выполняется с помощью двух вергентных вращений: суправергенции, выравнивающей зрительные оси по вертикали, или долготы $\theta_L = \theta_R = \theta$, и конвергенции — по горизонтали, или широте на угол $\gamma = \varphi_R - \varphi_L$. Углы по широте φ_R и φ_L отсчитываются от осей z_R и z_L соответственно.

Как известно [7, 9], locusом, или геометрическим местом точек плоскости с $\gamma = \text{const}$, является круг. Через центры проекции глаз (задние узловые точки) проходит круг Вит-Мюллера, через центры вращения глаз — изовергентный круг. Направление бинокулярного взгляда ϕ совпадает со зрительной осью циклопического глаза, центр вращения C которого лежит на пересечении изовергентного круга 1 и медиальной плоскости головы (ось z_C). Положение центра изменяется при фиксации точек, удалённых на разные расстояния. Locusами точек плоскости с $\phi = \text{const}$ являются известные гиперболы Гиллебранда 2 [9]. Пусть $\varphi_R = \alpha$ и $\varphi_L = \beta$, а r — расстояние от центра C до точки F , или радиус окружности 3, тогда

$$\gamma = \alpha - \beta; \quad \phi = \frac{\alpha + \beta}{2}; \quad r = \frac{b \cos \phi}{\sin \gamma}. \quad (1)$$

Эксперименты Фоли [10] показали, что в экиполярной плоскости геометрическое место точек, равноудалённых от наблюдателя, — не изовергентный круг, а кривая, близкая к окружности радиуса r .

Ретинотопия. Угловые размеры поля зрения индивидуальны, но в среднем равны 60° к носу, 55° вверх, $90\text{--}100^\circ$ к виску и $70\text{--}75^\circ$ вниз, что составляет телесный угол $\sim 1,27\pi$. Бинокулярное поле $\sim 120^\circ$ по горизонтали и вертикали, что приблизительно равно $0,94\pi$. Наконец, общее поле зрения $\sim 1,6\pi$, или $\sim 80\%$ площади полусферы [11]. Сетчатка каждого глаза содержит 120–130 млн палочек и 6–8 млн колбочек, а зрительный нерв, который соединяет сетчатку и мозг, имеет лишь 0,8–1 млн аксонов от ганглиозных клеток. Таким образом, визуальная информация сжимается примерно на два порядка.

Ретинотопия, или визуотопия, есть топологическое представление сетчатки на коре головного мозга. Оно нелинейно, но сохраняет соответствие рецептивных полей. То есть стимуляция двух соседних областей на сетчатке вызывает возбуждение двух смежных областей коры головного мозга. Это представление первого рода, непрерывное. Роса в [12] показал, что существуют и представления второго рода, разрывные.

Нелинейность преобразования изображений сетчатки объясняется рядом причин: во-первых, абберациями физиологической оптики глаза, которые увеличиваются на периферии зрительного поля (особенно астигматизм); во-вторых, неравномерным распределением по сетчатке рецептивных полей и ганглиозных клеток. С этими двумя факторами согласована разрешающая способность глаза, которая максимальна для фовеальной ямки (~ 1 угл. мин) и быстро уменьшается к периферии (со скоростью $\sim (4\text{--}5)$ угл. мин на каждые 10° отклонения от фовеа). Таким образом, создаётся основа для кодирования информации с высокой степенью сжатия.

Пусть острота разрешающей способности глаза вдоль горизонтального меридиана согласно [11] имеет вид

$$\Sigma(\varphi) = (1 + k\varphi)^{-1}, \quad (2)$$

где Σ — острота (в обратных угловых минутах); φ — широта на интервале $[0\text{--}90^\circ]$; k — коэффициент уменьшения разрешающей способности (в угл. мин/град). Тогда число аксонов, или пикселей с максимальным разрешением в заданном интервале φ , запишем как

$$N(\varphi) = 60 \int_0^\varphi \frac{d\varphi}{1 + k\varphi} = \frac{60}{k} \ln(1 + k\varphi). \quad (3)$$

При $k = 0,4$, $\varphi = 90^\circ$ число пикселей $N = 542$. Общее число пикселей порядка $\pi N^2 = 0,92 \cdot 10^6$.

С учётом (2) на рис. 3 представлены зрительное поле (а) и его ретинотопическая карта на коре головного мозга (б). Форма сжатого поля зрения не расходится с современными представлениями нейрофизиологов [12, 13]. В работе [13] такую форму называют «грушей». Отчётливо видны полевые абберации, кривизна поля. В районе фовеа ($0\text{--}3^\circ$) и зоне ясного видения (до 15°) наблюдаются большие угловые, так называемые кортикальные увеличения, которые уменьшаются к периферии. Известные функциональные элементы зрительной системы: латеральное колленчатое тело, первичная зрительная кора $V1$, экстрастриарные зоны коры ($V2\text{--}V6$), а также глазодвигательная система (верхние бугорки четверохолмия) — используют ретинотопические карты.

Двумерный сфероидальный ретиноид. Проблема реализации и локализации механизмов стабилизации зрительного поля сводится к тому, как отыскать элементы зрительной системы, «рецептивные поля» которых будут «оторваны» от сетчатки и привяза-

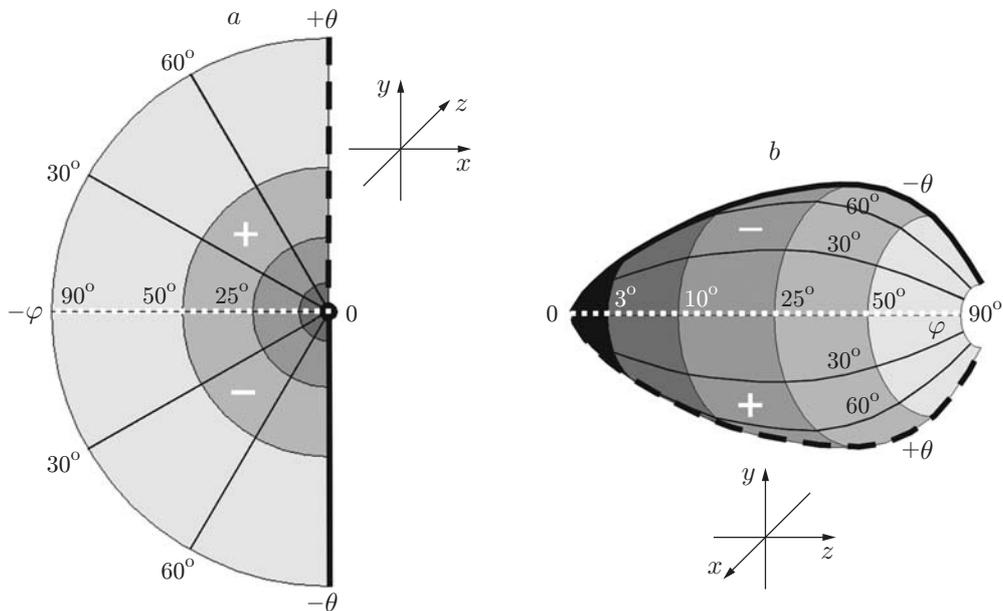


Рис. 3. Ретинотопия: *a* — левая половина зрительного поля — четверть сферы, *b* — ретинотопическая карта этого участка в правом полушарии — половина «груши» (0 — центр фовеа, полюс сферической системы координат; белая пунктирная линия — горизонтальный меридиан; штриховая и сплошная жирные линии — вертикальный меридиан)

ны к внешнему пространству [2]. Действительно свободное движение фигур без искажения формы по «грушевидной» поверхности, представляющей сетчатку, невозможно. Трегуб попытался решить вопрос с помощью двумерных ретиноидов [3], которые при каждой саккаде получают мозаику пикселей малого углового размера (2–3°) из фовеальной части «груши» и привязывают её к внешнему «плоскопараллельному» пространству. Недостатки такого способа рассмотрены в [5]. Это, прежде всего, замедленная реакция на вестибулярные сигналы, требующие быстрой смены зрительного контента.

На рис. 4 показано другое решение. Приведено сечение двумерного сфероидального ретиноида (двумерный *S*-ретиноид), содержащего декодер и два регистра. Полномасштабный ретиноид — это объёмная многослойная структура, где каждый слой является

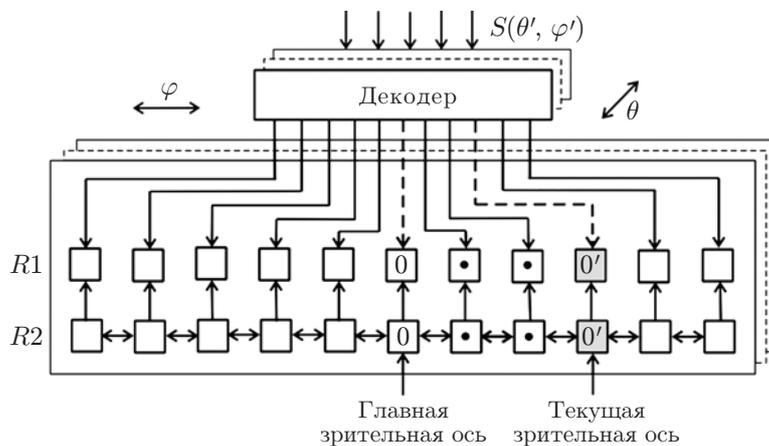


Рис. 4. Двумерный сфероидальный ретиноид

репликой изображённого сечения, снабжённой дополнительными межслойными соединениями. Декодер выполняет три основные функции: во-первых, «обратное» преобразование сигналов от зрительного нерва $S(\theta', \varphi')$ в изображение на сфере; во-вторых, вращение сферического поля для совмещения с текущим положением зрительной оси θ, φ (см. рис. 1); в-третьих, преобразование координат в систему Гельмгольца.

Регистр $R1$ необходим для хранения изображения поля на эпиполярных меридианах θ . Каждая ячейка регистра $R1$ на широте φ помнит матрицу пикселей, соизмеримую с мозаикой $\sim(12 \times 15)$ в колонке стриарной коры $V1$ [13]. Возможны два режима записи информации в регистр $R1$: полная перезапись со стиранием прежнего содержания и запись комбинации изображений, старого и нового.

Сдвиговой регистр $R2$ предназначен для размещения «себя» в среде, своего движения и селективного внимания. Код 0 — главная зрительная ось, код $0'$ — текущая зрительная ось. Регистр $R2$ соответствует регистру $R5$ ретиноидов Трегуба [3].

В предложенной модели памяти двумерное изображение внешнего пространства хранится на сфере в неподвижной относительно головы системе координат Гельмгольца. Память обеспечивает константность зрительного поля, в котором постоянные изменения сетчаточных изображений, вызываемые движениями глаз, не представлены в восприятии. При движении только глаз не может произойти изменений в относительном расположении предметов среды и наблюдателя, поэтому содержание кратковременной памяти сохраняется. Движения глаз лишь способствуют сокращению слепых зон [5] и улучшают разрешающую способность поля. При адекватном возбуждении вестибулярного аппарата, когда голова (и/или тело) приходит в движение, относительное расположение предметов и наблюдателя изменяется. Возникают новые окклюзии и параллаксы, происходит смещение и поворот главной зрительной оси. В этих условиях по сигналу от вестибулярного аппарата осуществляют полную перезапись двумерного S -ретиноида со стиранием прежнего изображения.

Трёхмерный сфероидальный ретиноид. На рис. 5 в эпиполярной плоскости θ показано сечение бинокулярной модели кратковременной памяти. Полномасштабный трёхмерный S -ретиноид — это объёмная многослойная структура, где каждый слой обслуживает отдельный эпиполярный меридиан θ . Для простоты иллюстрации показаны регистры с размером ячейки 10° , которые обеспечивают работу трёхмерного ретиноида в бинокулярном поле $\pm 60^\circ$.

Корреляционные кластеры (элементы круглой формы) матрицы 3 иннервируются аксонами от ячеек левого и правого ретиноида. Задача кластера — найти взаимную корреляцию мозаик, приходящих от левого и правого глаза. На меридиональном уровне находится взаимная корреляция между мозаиками каждой ячейки левого и правого ретиноида. Большой коэффициент корреляции свидетельствует об успешной фузии (слиянии, сплаве) монокулярных изображений предметов в процессе стереопсиса. Взаимная корреляция между изображениями левого и правого глаза на уровне мозаик выполняется всеми корреляционными кластерами трёхмерной матрицы одновременно. И в этом основное достоинство параллельного «нейрокомпьютинга».

Индексация элементов a_{ij} матрицы 3 такова, что строки i соответствуют углу конвергенции глаз $\gamma = \alpha - \beta$, или абсолютной диспарантности относительно $\gamma = 0$, а столбцы j — направлению бинокулярного взора $\phi = (\alpha + \beta)/2$, или взгляда циклопического глаза. Таким образом, строки матрицы представляют семейство изовергентных кругов, а фронтопараллельные плоскости трёхмерной матрицы с учётом угла θ — семейство тороидальных поверхностей. Итак, мы реализовали биполярную систему координат, в которой точка внешнего физического пространства задаётся с помощью трёх углов γ, ϕ, θ .

Рудольф Лунбург — создатель теории бинокулярного визуального пространства [14] — считал, что тороидальные поверхности $\gamma = \text{const}$ воспринимаются как концент-

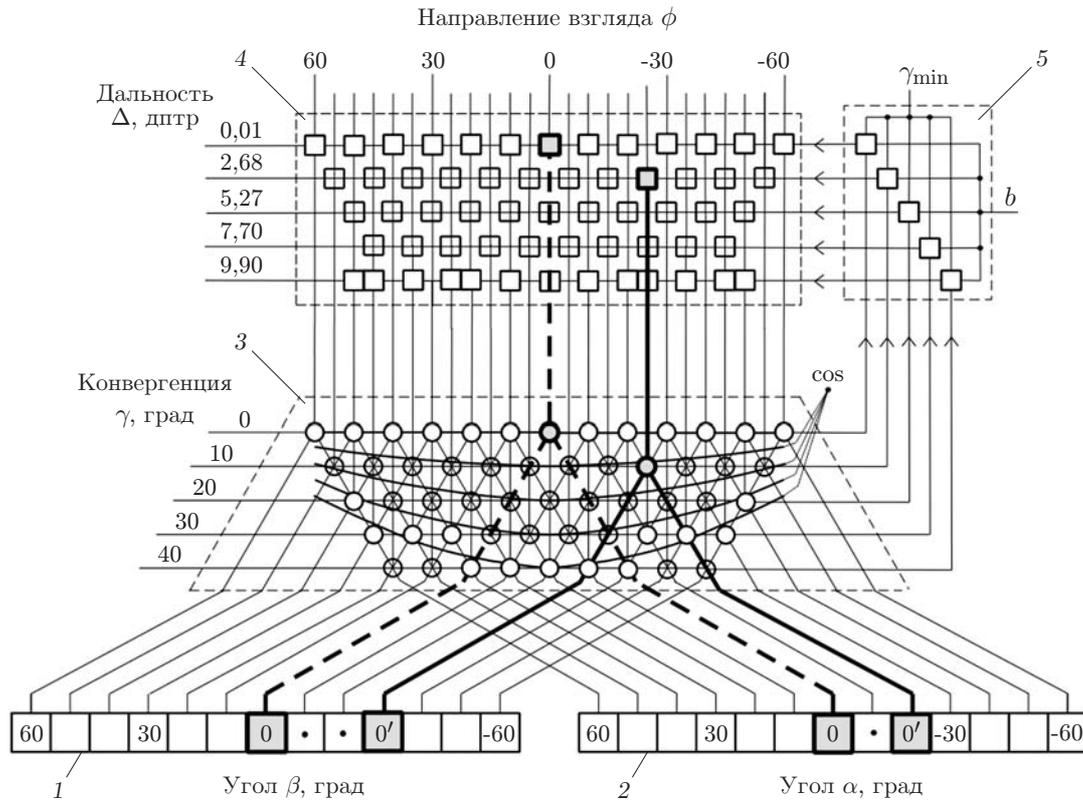


Рис. 5. Трёхмерный сфероидальный ретиноид: 1, 2 — левый и правый двумерные S -ретиноиды; 3 — матрица корреляционных кластеров; 4 — память зрительного экрана; 5 — преобразователь индексов

рические сферы вокруг центра наблюдения. Поэтому ВП можно задать в сферической системе радиусом и двумя углами: $\rho = 2 \exp(-\sigma\gamma)$, ϕ, θ , где $\sigma = \text{const}$, а сфера с радиусом $\rho = 2$ представляет физическую бесконечность. Лунбург предложил гомоморфное отображение физического пространства и оказался неправ. Уже упоминались эксперименты Фоли [10], показавшие, что равноудалёнными воспринимаются лишь точки окружности, а не изовергентного круга. Таро Индо в [9, с. 5] доказывает, что физическое и визуальное пространства изоморфны и даже «изометричны», особенно на малых расстояниях [15]. Иначе невозможна визуальная локомоция без ошибок и катастроф.

Поставим такой вопрос: как изображается сфера на матрице 3? Из (1) радиус сферы $r = b \cos \phi / \sin \gamma$. Пусть $\gamma = \gamma_0$ и $\sin \gamma = \sin \gamma_0 \cdot \cos \phi$ при $\phi = 0$, тогда

$$r = b / \sin(\gamma_0) = \text{const} \quad \text{при} \quad \gamma = \arcsin(\sin \gamma_0 \cdot \cos \phi) \approx \gamma_0 \cos \phi. \quad (4)$$

Сфера задаётся также тремя углами γ, ϕ, θ , но с учётом (4) появляется новая «косинусная» индексация элементов матрицы 3. На рис. 5 показана виртуальная «прошивка» таких элементов (cos), имеющая смысл абсолютной диспаратности точек сферы. Заметим, что индекс столбцов никак не изменяется.

Матрица элементов памяти 4 представляет упоминавшийся ранее КЗЭ, который отображает ВП на концентрических сферах. Элементы памяти хранят фузированные микроизображения, полученные корреляционными кластерами в процессе стереопсиса. Каждый элемент матрицы 3 с косинусной индексацией (4) пересылает данные вдоль направления ϕ соответствующему элементу строки матрицы 4. Иначе говоря, косинусная «строка»

матрицы 3 преобразуется в «горизонтальную» строку матрицы 4. При помощи преобразователя индексов 5 «конвергенция» трансформируется в соответствующую «аккомодацию»:

$$\Delta = r^{-1} + d^{-1} = b^{-1}(\sin \gamma_0 + \gamma_{\min}), \quad (5)$$

где Δ — дальность (в диоптриях); r — расстояние (в метрах); b — глазная база (в метрах); γ_{\min} — разрешающая способность глаз (в радианах); $d = b/\gamma_{\min}$ — предельное расстояние. Выражение (5) легко объяснить с психофизической точки зрения [9]. Визуальное пространство ограничено и замкнуто. Человек никогда не видит «бесконечность» или «пустоту». Конец линии взгляда упирается либо в предмет, либо в фон. Фон воспринимается как предельно удалённая поверхность (небо как свод). Из (5) сфера ВП оказывается заданной в системе координат

$$\rho = \frac{1}{\Delta} = \frac{rd}{r+d}, \quad \phi, \theta, \quad (6)$$

где ρ, r — радиусы сфер визуального и физического пространств; d — предельное расстояние в любом из направлений ϕ, θ . Выражение (6) представляет проективное преобразование, которое выполняется с помощью дробно-линейной функции. В [15, 16] показано, что такое преобразование отображает евклидово пространство на модель Клейна пространства Лобачевского. Поэтому ВП на экране трёхмерного S -ретиноида оказывается моделью гиперболического пространства с радиусом кривизны, равным d .

Предложенная модель допускает решение разнообразных зрительных задач. Структура ретиноида легко вписывает положение личного «Я», своего движения и селективного внимания, обращённого к предметам внутри ВП (см. рис. 5, жирные линии). При работе с последовательными образами ретиноид производит обратные преобразования по сравнению с оптическими преобразованиями сетчаточных изображений, что соответствует опытным данным [2]. В частности, это относится и к закону Эммерта. При приближении к точке фиксации последовательные образы уменьшаются в размерах, а реальные изображения увеличиваются. Вращения, независимые от сетчатки, и комбинированная запись допускают ортоскопическое восприятие. Возможно отображение монокулярных изображений на любой сфере экрана в бикулярном виде и т. д.

Заключение. В данной работе предложена модель кратковременной памяти системы зрения — трёхмерный сфероидальный ретиноид, обеспечивающий стабильность зрительного поля при движении глаз путём преобразования и хранения изображений поля в неподвижной относительно головы системе координат. Выполняется процесс стереопсиса на матрице корреляционных кластеров. Формируется и сохраняется визуальное пространство в сознании наблюдателя. Модель соответствует требованиям бикулярной системы зрительного восприятия.

К сожалению, предложенная модель не учитывает параллельного и согласованного функционирования бикулярных и монокулярных механизмов, аккомодации глаз. Более 40 % общего поля зрения является монокулярным. Причём оно также объёмно, как и бикулярное. Проблема состоит в отсутствии признанной теории аккомодации на «сетчаточном» уровне.

Ещё одна проблема связана с «расходом» большого объёма «нейронного» материала при реализации ретиноида. Расход можно сократить с помощью ретинотопических карт, снабжённых метками пространственного положения клеток (или мозаик). Работы в этой области могут быть предметом последующих исследований. Возможно, что такой подход разрешит и проблему локализации ретиноидов в нейрофизиологическом плане.

СПИСОК ЛИТЕРАТУРЫ

1. **Шиффман Х. Р.** Ощущение и восприятие. С.-Пб.: Питер, 2003. 924 с.
2. **Зенкин Г. М., Петров А. П.** О механизмах константности зрительного восприятия пространства // Сенсорные системы. Л.: Наука, 1979. С. 25–39.
3. **Trehub А.** The Cognitive Brain. Cambridge: MIT Press, 1991. 329 p.
4. **Trehub А.** Space, self, and theater of consciousness // Consciousness and Cognition. 2007. **16**, N 2. P. 310–330.
5. **Ковалев А. М.** О визуальном пространстве и ретиноидах Трегуба // Автометрия. 2011. **47**, № 1. С. 102–109.
6. **Ковалев А. М.** О центральных проекциях трехмерного пространства // Автометрия. 1996. № 6. С. 4–15.
7. **Howard I. P., Rogers B. J.** Binocular Vision and Stereopsis. N. Y.: Oxford University Press, 1995. 736 p.
8. **Гиппенрейтер Ю. Б.** Движения человеческого глаза. М.: Изд-во МГУ, 1978. 256 с.
9. **Indow Т.** The global structure of visual space // Advanced Series of Mathematical Psychology. World Scientific, 2004. Vol. 1. 228 p.
10. **Foley J. M.** Locus of perceived equidistance as a function of viewing distance // JOSA. 1966. **56**, N 6. P. 822–827.
11. **Ковалев А. М.** О числе элементов изображения в поле зрения наблюдателя // Автометрия. 1997. № 3. С. 30–34.
12. **Rosa M. G. P.** Visual maps in the adult primate cerebral cortex: some implications for brain development and evolution // Brazillian Journ. Med. and Biol. Res. 2002. **35**, N 12. P. 1485–1498.
13. **Schmolesky М.** The Primary Visual Cortex // Online textbook.
URL: <http://webvision.med.utah.edu/book/part-ix-psychophysics-of-vision/the-primary-visual-cortex/> (дата обращения: 22.03.2012).
14. **Luneburg R. K.** The metric of binocular visual space // JOSA. 1950. **40**, N 10. P. 627–642.
15. **Ковалев А. М.** О моделях визуального пространства // Оптика и спектроскопия. 2006. **100**, № 1. С. 134–141.
16. **Ковалев А. М.** Об анизотропной модели визуального пространства // Автометрия. 2006. **42**, № 6. С. 53–62.

Поступила в редакцию 22 марта 2012 г.
