

К математической теории зрительного восприятия

В. Н. Козлов

Зрение занимает особое место среди органов чувств человека. По некоторым оценкам биологов до 90% сенсорной (то есть от органов чувств) информации составляет зрительная информация. Столь же очевидно значима визуальная информация и для работы технических устройств: роботов, компьютеров, систем наведения на цели, и пр.

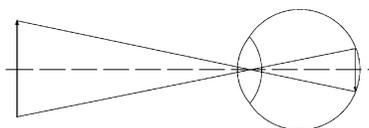


Рис. 1.

Глаз и получение изображения на сетчатке обычно представляют так, как это сделано на рис. 1. Хрусталик выполняет в этой схеме роль простой линзы, обеспечивающей формирование перевернутого действительного изображения. Поместим перед глазом некоторый объект. На сетчатке сформируется совокупность возбужденных рецепторов. Эта совокупность — и ничего более — есть тот первичный материал, который будет далее анализироваться мозгом в процессе распознавания. Можно сказать, что для последующих этапов распознавания эта совокупность возбужденных рецепторов и есть объект. Мы можем, конечно, предполагать в процессе распознавания выделение каких-то характеристик изображения, константных признаков и пр., — но все это будет позже и может базироваться только на анализе первичного материала.

Вопрос, который нас здесь интересует, состоит в следующем: по каким принципам, или по каким правилам, происходит работа с информацией от рецепторов сетчатки, приводящая, в конечном счете, к распознаванию объекта, помещенного перед глазом? Эти правила можно также назвать функциональными механизмами зрительного восприятия и, соответственно, ставить вопрос об изучении этих механизмов. К сожалению, нет методов, которые позволяли бы прямо и непосредственно изучать функциональные механизмы в живом организме. Единственное, что можно делать — это изобретать правила, которые обеспечивали бы в той или иной степени требуемый конечный результат (распознавание), и рассматривать эти правила как гипотезу о соответствующем функциональном механизме. При этом, конечно, важно, чтобы изобретаемые правила в как можно большей степени учитывали бы известные в биологии данные об устройстве глаза, его функционировании и об условиях, в которых происходит работа глаза.

Выяснение функциональных механизмов зрительного восприятия, очевидно, важно и имеет самостоятельную ценность. При этом, однако, может создаться впечатление, что эта задача все же обособлена от вопросов, традиционно изучаемых в биологии при исследованиях зрительной системы, и непосредственно с ними не связана. Можно задаться вопросом о том, какое место задача о функциональных механизмах зрения занимает в ряду других исследований зрительной системы. Ответом, отчасти, может служить то обстоятельство, что, хотим мы того или нет, но как правило при исследованиях зрительного восприятия явно или неявно (чаще неявно) имеется ввиду некоторая гипотеза о правилах переработки информации в зрительной системе. Так, например, когда рассуждают о нейронах, регистрирующих появление на сетчатке световых пятен определенной формы или световых полос определенного направления (такие исследования очень распространены), предполагают, тем самым, что имеются «детекторы» признаков в виде тех же пятен или полос, и дальнейшая работа происходит с информацией от этих «детекторов». Нетрудно видеть, что это и есть, в сущности, гипотеза о некотором функциональном механизме. Неявно же руководствуясь некоторой гипотезой,

легко увидеть в эксперименте то, что ей соответствует, и не принять во внимание то, что в нее не укладывается.

Говоря об изображениях и их распознавании с позиций математики, имеет смысл сразу как можно яснее определиться, во-первых, с тем, что понимается под изображением, во-вторых, с тем, как сравниваются изображения, как оценивается то, что называется их «похожестью», «близостью по форме», совпадением «в основных чертах», и пр.

В этой работе под изображением понимается конечное (непустое) множество точек на плоскости. Содержательным обоснованием этому может служить то, что любое реальное (нецветное) изображение можно аппроксимировать изображением из точек, причем в нужной мере можно передать все градации «серого цвета» разной плотностью точек в разных частях изображения. Такое представление не закрывает дорогу и к рассмотрению цветных изображений, поскольку, как известно, цветное изображение можно представить тремя нецветными. Наконец, все, что мы видим, мы видим посредством глаз. Изображение из среды проецируется на сетчатку глаз, что приводит к возбуждению части рецепторных клеток, то есть, в конечном счете, к формированию на сетчатке аналога составленного из точек изображения.

Сравнение «на похожесть» двух изображений предлагается осуществлять таким их наложением одного на другое, при котором соответствующие точки на двух изображениях оказываются как можно более близко друг от друга. Рассмотрим следующий пример. Изображения A и B , например, цифры «2» состоят каждое из n точек. Фигуру B движениями, то есть параллельными переносами, поворотами и «переворотами» (преобразованиями симметрии относительно прямой) специальным образом расположили по отношению к A . Это расположение характеризуется тем, что каждая точка из B оказывается от ровно одной соответствующей ей точки из A на расстоянии, не большем некоторого R . Положим, что это R уже никакими движениями фигуры B уменьшить нельзя. В этом случае можно считать, что фигура B своими контурами повторяет фигуру A , а величина R — характеристика рассогласования в контурах. Если на изображение A

пытаться наложить движениями не «двойку», а другую фигуру, например, «четверку», то, можно ожидать, характеристика рассогласования будет существенно больше. На этом и предлагается основывать процедуры различения фигур по форме, и в целом распознавания изображений.

Отметим, что мы не можем считать заранее известным, какие точки на A и B являются соответствующими. Поэтому схема решения задачи предполагает проверку всех $n!$ возможных вариантов соответствия точек на A и B друг другу. Пусть ψ — некоторое одно из таких соответствий. Обозначим через B^p множество всех изображений, получаемых из B параллельными переносами и поворотами. Полагаем, что на изображениях из B^p сохраняется нумерация точек, порожденная изображением B , то есть если $B' \in B^p$, то через b'_j в B' обозначена точка, в которую при соответствующем преобразовании перешла точка b_j из B ($j = 1, \dots, n$). Пусть $b_{\psi(i)}$ — точка в изображении B' , сопоставленная отображением ψ точке a_i . Обозначим через $l(B')$ длину наибольшего из отрезков $(a_i b'_{\psi(i)})$ ($i = 1, \dots, n$). Пусть минимум величин $l(B')$, рассматриваемых для всех $B' \in B^p$, достигается на изображении B'_0 . Обозначим этот минимум через $r_\psi(A, B)$. Рассмотрим изображение \bar{B} , полученное из B преобразованием симметрии относительно какой-либо прямой, и пусть минимум $r_\psi(A, \bar{B})$ достигается на изображении B''_0 из \bar{B}^p . Обозначим через $R_\psi(A, B)$ меньшую из величин $r_\psi(A, B)$ и $r_\psi(A, \bar{B})$, через B_0^ψ — изображение из $B^p \cup \bar{B}^p$ (то есть B'_0 или B''_0), на котором этот минимум достигается.

Множество $B^p \cup \bar{B}^p$ обозначим через B^* . Ясно, что B^* — множество всех изображений, которые можно получить из B изометрическими преобразованиями.

Далее остается выбрать из величин $R_\psi(A, B)$, полученных для всех возможных ψ , наименьшую. Этот минимум обозначаем через $R(A, B)$, биекцию, на которой он достигается, обозначаем через ψ_0 и называем искомым соответствием между точками изображений A и B . Изображение $B_0^{\psi_0}$ обозначаем через B_0 и называем искомым, взаиморасположение A и B_0 — тоже искомое. Величину $R(A, B)$ называем также расстоянием между A и B .

В описанной схеме нераскрытой пока осталась процедура определения величин $r_\psi(A, B)$. Без ограничения общности можем полагать, что биекцией ψ точке a_i изображения A сопоставляется точка b'_i изображения B' из B^* . Точки a_i и b'_i называем соответствующими, соответствующими называем и отрезки $(a_j a_j)$ и $(b'_i b'_j)$ ($i \neq j$, $i, j = 1, \dots, n$).

Введем понятие угла между изображениями. Зафиксируем некоторые p и q и в качестве угла между изображениями в паре (A, B') возьмем угол φ , образованный отрезками $(a_p a_q)$ и $(b'_p b'_q)$ или их продолжениями. Без ограничения общности можно полагать, что для исходных изображений A и B отрезки $(a_p a_q)$ и $(b'_p b'_q)$ параллельны, однонаправлены (то есть если в отрезке $(a_p a_q)$ слева направо идет сначала точка a_p , затем a_q , то и в отрезке $(b'_p b'_q)$ слева направо идут последовательно точки b'_p и b'_q), и угол φ равен нулю. Для произвольного B' из B^p полагаем угол φ между B' и A находящимся в промежутке от 0 до 2π .

Обозначим через $\{B\}_\varphi$ множество всех тех изображений из B^p , которые имеют данный угол φ с изображением A . Ясно, что во множестве $\{B\}_\varphi$ изображения переводимы друг в друга параллельными переносами.

Ранее [2, 3] было доказано утверждение, которое применительно к обозначениям настоящей статьи выглядит следующим образом.

Теорема 1. *В $\{B\}_\varphi$ существует и единственно изображение, на котором достигается минимум величин $l(B')$, рассматриваемых для всех $B' \in \{B\}_\varphi$.*

Этот минимум — обозначим его через $r(\varphi)$ — представляется тем самым как функция от угла φ . Далее вопрос сводится к тому, чтобы определить минимум величин $r(\varphi)$ при углах $0 \leq \varphi < 2\pi$.

В этих построениях используется понятие характеристического изображения C для пары (A, B') , где $B' \in B^*$. Оно состоит из точки O и точек c_1, \dots, c_n , называемых соответственно центром и точками ядра. В качестве O берется произвольная точка плоскости. Затем параллельным переносом отрезка $(a_i b'_i)$ ($i = 1, \dots, n$) совмещаем точку

a_i с точкой O . Точку, в которую переходит при этом b'_i , обозначаем через c_i . Обратной процедурой при заданных A и C получается, нетрудно видеть, изображение B' . Отметим, что некоторые из точек характеристического изображения могут совпадать (сливаться), и в этом смысле C является особым изображением.

Показано, что ядра характеристических изображений для всех изображений из $\{B\}_\varphi$ переводимы друг в друга параллельными переносами. Это значит, что эти характеристические изображения различаются только положением центра относительно точек ядра. Окружность минимального по радиусу круга, включающего все точки ядра, называем ключевой. Такая окружность при заданном ядре существует и единственна. Показано, что для изображения B' из $\{B\}_\varphi$, на котором достигается минимум величин $l(B')$, центр характеристического изображения должен совпадать с центром ключевой окружности. Это и позволяет построить такое изображение B' .

Пусть c_i и c_j — точки из ядра характеристического изображения ($i \neq j$, $i, j = 1, \dots, n$). Из рассмотрения в [2, 3, 4] следует, что если параллельными переносами отрезков $(a_i a_j)$ и $(b'_i b'_j)$ совместить точку a_j с c_i , и точку b'_j с c_j , то точки a_i и b'_i совместятся в одну точку, обозначим ее через c_{ij} . При заданных изображениях A и B тем самым стороны $(c_i c_{ij})$, $(c_j c_{ij})$ и угол между этими сторонами в треугольнике $c_i c_j c_{ij}$ (обозначим его через α_{ij}) можно считать известными. Следовательно, может быть получена и длина отрезка $(c_i c_j)$: $(c_i c_j)^2 = (a_i a_j)^2 + (b'_i b'_j)^2 - 2(a_i a_j)(b'_i b'_j) \cos \alpha_{ij}$. Поскольку угол α_{ij} зависит от угла φ между изображениями, то длина отрезка $(c_i c_j)$ есть функция от угла φ .

Из формулы для отрезка $(c_i c_j)$ следует, что он равен нулю, то есть точки c_i и c_j слиты в одну, только при равенстве длин отрезков $(a_i a_j)$ и $(b'_i b'_j)$, их параллельности и однонаправленности. Угол α_{ij} при этом равен нулю. Существует лишь конечное число углов φ , при которых это может иметь место. Для каждого такого φ из множества $\{B\}_\varphi$ выбираем то изображение B' , для которого центр характеристического изображения пары (A, B') совпадает с центром ключевой окружности, и при этом слившиеся точки c_i и c_j лежат на ключевой окружности. Множество всех таких изображений обозначаем через

U_0^ψ . При дальнейших рассмотрениях полагаем, что уже никакие две точки ядра на ключевой окружности не слиты в одну.

Положим, найдено изображение B'_0 , на котором достигается минимум величин $l(B')$, полученных при всех $B' \in B^p$. Далее априори возможны три случая: на ключевой окружности находятся соответственно две, три и более трех точек ядра.

Первый случай. На ключевой окружности находятся две точки c_{i_1} и c_{i_2} ($i_1, i_2 \in (1, \dots, n)$), причем отрезок $(c_{i_1} c_{i_2})$ должен быть диаметром ключевой окружности. В этом случае соответствующие отрезки $(a_{i_1} a_{i_2})$ и $(b_{i_1}^0 b_{i_2}^0)$ должны быть параллельны, однонаправлены и середины их должны совпадать. Если поочередно в качестве порождающих эту пару отрезков рассматривать все отрезки из A и соответствующие отрезки из B , то этим условием определяется конечное множество изображений из B^p , обозначим его через U_1^ψ .

Второй случай. На ключевой окружности находятся три точки $c_{i_1}, c_{i_2}, c_{i_3}$ ($i_1, i_2, i_3 \in (1, \dots, n)$). Треугольник с вершинами в этих точках должен быть остроугольным. Длину отрезков $(a_{i_1} b_{i_1}^0)$, $(a_{i_2} b_{i_2}^0)$, $(a_{i_3} b_{i_3}^0)$ обозначим через R , прямые L_1, L_2, L_3 , на которых лежат эти отрезки, будем считать направленными, с направлением от a_{i_j} к $b_{i_j}^0$ ($j = 1, 2, 3$).

Доказано, что прямые L_1, L_2, L_3 должны пересекаться в одной точке.

Точку пересечения обозначим через O_L и назовем центром трехосника, состоящего из осей L_1, L_2, L_3 . На каждой оси часть ее от центра в направлении оси называем положительной, оставшуюся часть — отрицательной. Отрезки $(O_L a_{i_1})$, $(O_L a_{i_2})$, $(O_L a_{i_3})$ представляем величинами соответственно x, y, z . Длину отрезка $(O_L a_{i_1})$ полагаем равной $|x|$, x полагаем положительным, если a_{i_1} находится в положительной части оси L_1 , и отрицательным в противном случае. Аналогичное предпологаем и про y и z . Углы (положительные, меньшие π , в сумме составляющие 2π) между осями L_1 и L_2 , L_1 и L_3 , L_2 и L_3 обозначаем через соответственно α, β, γ .

Ясно, что задание конкретных значений для $x, y, z, \alpha, \beta, \gamma, R$ определяет конкретное положение точек $b_{i_1}^0, b_{i_2}^0, b_{i_3}^0$ относительно точек $a_{i_1}, a_{i_2}, a_{i_3}$, а значит и конкретное изображение из B^p .

Показано, что все варианты взаиморасположения точек $a_{i_1}, a_{i_2}, a_{i_3}$ и $b_{i_1}, b_{i_2}, b_{i_3}$ на трехоснике находятся среди решений следующей системы уравнений:

$$\left\{ \begin{array}{l} \alpha + \beta + \gamma = 2\pi, \\ (a_{i_1} a_{i_2})^2 = x^2 + y^2 - 2xy \cos \alpha, \\ (b_{i_1} b_{i_2})^2 = (x + R)^2 + (y + R)^2 - 2(x + R)(y + R) \cos \alpha, \\ (a_{i_1} a_{i_3})^2 = x^2 + z^2 - 2xz \cos \beta, \\ (b_{i_1} b_{i_3})^2 = (x + R)^2 + (z + R)^2 - 2(x + R)(z + R) \cos \beta, \\ (a_{i_2} a_{i_3})^2 = y^2 + z^2 - 2yz \cos \gamma, \\ (b_{i_2} b_{i_3})^2 = (y + R)^2 + (z + R)^2 - 2(y + R)(z + R) \cos \gamma. \end{array} \right. \quad (1)$$

Обозначим через U_2^ψ множество изображений из B^p , определяемых решениями систем (1) для всех троек точек $a_{i_1}, a_{i_2}, a_{i_3}$ и соответствующих точек $b_{i_1}, b_{i_2}, b_{i_3}$ ($i_1, i_2, i_3 \in (1, \dots, n)$).

Третий случай. На ключевой окружности находятся четыре и более точек. Рассмотрим четыре из них: $c_{i_1}, c_{i_2}, c_{i_3}, c_{i_4}$. Из теоремы Птолемея следует, что для того, чтобы эти четыре точки находились на одной окружности, необходимо и достаточно, чтобы выполнялось одно из трех равенств:

$$\pm(c_{i_1} c_{i_2})(c_{i_3} c_{i_4}) \pm (c_{i_1} c_{i_3})(c_{i_2} c_{i_4}) \pm (c_{i_1} c_{i_4})(c_{i_2} c_{i_3}) = 0.$$

Длина каждого из отрезков в этих уравнениях есть функция от угла φ . Для каждого угла φ , являющегося решением, определено (единственное) изображение из $\{B\}_\varphi$, для которого центр характеристического изображения совпадает с центром ключевой окружности. Такие уравнения составляем для каждой четверки точек $a_{i_1}, a_{i_2}, a_{i_3}, a_{i_4}$ и соответствующих точек $b_{i_1}, b_{i_2}, b_{i_3}, b_{i_4}$. Множество таким образом получаемых изображений из B^p обозначаем через U_3^ψ .

Далее меняем изображение B преобразованием симметрии относительно прямой на изображение \bar{B} , и проделываем все аналогичное описанному выше для изображения \bar{B} и множества \bar{B}^p . Получаем множества $\bar{U}_0^\psi, \bar{U}_1^\psi, \bar{U}_2^\psi, \bar{U}_3^\psi$.

Объединением всех множеств $U_0^\psi, U_1^\psi, U_2^\psi, U_3^\psi, \bar{U}_0^\psi, \bar{U}_1^\psi, \bar{U}_2^\psi, \bar{U}_3^\psi$ и для всех биекций ψ получаем множество U .

Следующее утверждение можно считать некоторой перефразировкой теоремы из [2, 3].

Теорема 2. *Искомое изображение B_0 может находиться только среди изображений множества U .*

Самое трудоемкое в построении множества U — необходимость в переборе $n!$ биекций ψ . Вместе с тем, как это следует из описанного, взаиморасположение A и B' для всех изображений B' из U определяется фактически искомым взаиморасположением их фрагментов из двух, трех или четырех точек, соответствующих друг другу в рамках каждой из возможных биекций (эти фрагменты можно назвать ключевыми). Поэтому вместо перебора всех биекций ψ , можно, при построении множества U , перебирать все фрагменты из двух, трех и четырех точек на изображениях A и B все варианты сопоставления этих фрагментов (из одинакового числа точек) друг другу, и все варианты сопоставления друг другу точек во фрагментах. В этом состоит идея, которая ниже описывается подробнее.

Частью или фрагментом изображения A назовем любое (непустое) подмножество \tilde{a} точек изображения A .

Для каждого фрагмента \tilde{a} из двух точек изображения A и каждого фрагмента \tilde{b} из двух точек изображения B ищем их искомое взаиморасположение (см. описание первого случая выше). С учетом возможности по разному сопоставлять друг другу точки во фрагментах \tilde{a} и \tilde{b} общее число вариантов сопоставления фрагментов друг другу будет $(2!)(C_n^2)^2$. Каждый вариант сопоставления определяет два (с учетом преобразования симметрии) изображения из B^* . Множество всех таких изображений обозначаем через U_1^ψ .

Аналогично рассматриваем $(3!)(C_n^3)^2$ вариантов сопоставления друг другу точек во всех фрагментах \tilde{a} из трех точек и всех фрагментах \tilde{b} из трех точек. Определяемые для каждого из таких вариантов искомые взаиморасположения фрагментов (описание второго случая) дают некоторые изображения из B^* . Множество всех этих изображений обозначим через \tilde{U}_2 .

Наконец, рассмотрение аналогичным образом $(4!)(C_n^4)^2$ вариантов сопоставления фрагментов из четырех точек позволит сформировать множество изображений из B^* , которое обозначим через \tilde{U}_3 .

Труднее сформировать \tilde{U}_0 , аналог множества U_0^ψ . Изображениям из \tilde{U}_0 соответствуют «вырожденные» характеристические изображения, со слившимися точками c_{i_1} и c_{i_2} ядра. Однако необходимое условие для «вырожденности» — параллельность, однонаправленность и равенство по длине отрезков $(a_{i_1} a_{i_2})$ и $(b'_{i_1} b'_{i_2})$ — определяет не конкретное изображение из B^* , а только угол между таким изображением B' и изображением A . Такая же проблема возникала и выше, при формировании множества U_0^ψ , но там была задана в целом биекция ψ , это позволяло отобразить во множество U_0^ψ только те из изображений из B^* с заданным углом φ , у которых центр характеристического изображения совпадает с центром ключевой окружности.

Воспользуемся, однако, тем, что слившиеся точки c_{i_1} и c_{i_2} не могут быть единственными точками на ключевой окружности. Должна быть либо еще одна точка c_{i_3} , либо еще две точки c_{i_3} и c_{i_4} , не слившиеся с c_{i_1} и c_{i_2} . В качестве a_{i_3} и b_{i_3} (в варианте с тремя точками) могут быть взяты каждая из $(n - 2)$ точек изображения A , и каждая из $(n - 2)$ точек изображения B , то есть $(n - 2)^2$ вариантов. Аналогично вариантов с четырьмя точками может быть не более $2(C_{n-2}^2)^2$. Для каждого из вариантов берем из B^* для включения в \tilde{U}_0 те изображения, для которых центр характеристического изображения совпадает с центром окружности, на которой лежат точки $c_{i_1}, c_{i_2}, c_{i_3}$ или $c_{i_1}, c_{i_2}, c_{i_3}, c_{i_4}$. Отметим, что формируемое таким образом множество \tilde{U}_0 включает, очевидно, каждое из множеств U_0^ψ, \bar{U}_0^ψ для всех ψ как подмножество.

Обозначим через \tilde{U} объединение множеств $\tilde{U}_0, \tilde{U}_1, \tilde{U}_2, \tilde{U}_3$. Из описания процедуры формирования множества \tilde{U} следует, нетрудно видеть, следующее утверждение (аналог теоремы 2).

Теорема 3. *Изображение B_0 принадлежит множеству \tilde{U} .*

Формирование \tilde{U} , в отличие от U , не требует перебора $n!$ взаимно-однозначных соответствий между точками изображений A и B . При построении \tilde{U} сопоставляются друг другу только фрагменты \tilde{a} и \tilde{b}

и точки в них. Число же вариантов такого сопоставления, как это видно из рассмотренного выше, зависит от n полиномиально.

Однако формирование множества \tilde{U} не есть окончание задачи. Надо определить, каким образом из \tilde{U} будет выделяться искомое изображение B_0 и как будут сопоставляться друг другу точки на A и B_0 .

Пусть включение изображения B' в \tilde{U} определялось фрагментами \tilde{a} и \tilde{b} изображений A и B' при некотором соответствии $\psi_{\tilde{a}\tilde{b}}$ между точками этих фрагментов. По построению, соответствующие точки во фрагментах находятся друг от друга на одном расстоянии, которое обозначим через $R_{\tilde{a}\tilde{b}}$. Если B' — искомое изображение, то и все остальные точки из A должны находиться от соответствующих точек из B' на расстоянии меньшем или равном $R_{\tilde{a}\tilde{b}}$. Отсюда следует процедура определения того, что можно назвать приемлемым изображением из \tilde{U} . Для каждой точки a_i из A определим множество Q_{a_i} всех тех точек из B' , расстояние до которых от точки a_i не больше $R_{\tilde{a}\tilde{b}}$. Аналогично через $Q_{b'_j}$ обозначим множество тех точек из B' , расстояние до которых от b'_j не превышает $R_{\tilde{a}\tilde{b}}$. Если все множества Q_{a_i} и $Q_{b'_j}$ ($i, j = 1, \dots, n$) непусты, то изображение B' — приемлемое.

Назовем приемлемое изображение B' узкоприемлемым, если можно из каждого множества Q_{a_i} выбрать ровно по одной точке изображения B' так, чтобы все выбранные точки были разными. Таким выбором, нетрудно видеть, определяется некоторое взаимнооднозначное соответствие между точками двух изображений.

Среди узкоприемлемых выбираем изображение с наименьшей величиной $R_{\tilde{a}\tilde{b}}$. Оно и будет искомым изображением, а соответствующая биекция — искомым соответствием между точками изображений A и B .

Изменим теперь определение искомого изображения и искомого соответствия между точками изображений. Это изменение будет состоять в том, что в качестве искомого будем выбирать изображение B_0 с наименьшей величиной $R_{\tilde{a}\tilde{b}}$, но среди приемлемых изображений, а не узкоприемлемых. Точке a_i , будем теперь полагать, соответствует не одна точка, а все точки из множества Q_{a_i} , точке b_j^0 соответству-

ют все точки из множества $Q_{b_j^0}$. Кроме того, теперь можно полагать, что изображения A и B состоят не обязательно из одинакового числа точек, то есть $i = 1, \dots, n$, $j = 1, \dots, m$.

К изменениям в определении искомого изображения B_0 и искомого соответствия между точками A и B_0 мы пришли последовательностью шагов. Однако эти изменения можно сформулировать и независимо от этой последовательности. Пару изображений (A, B') охарактеризуем величиной $l(B')$, которая есть наименьшее такое число, что каждая точка одного изображения имеет на расстоянии, не большем $l(B')$, хотя бы одну точку другого изображения. Расстоянием $R(A, B)$ между изображениями A и B назовем наименьшую из величин $l(B')$, полученных для всех B' из B^* . Изображение B_0 , на котором этот минимум достигается, называем искомым. Взаиморасположение A и B_0 — искомое. Каждой точке a_i из A (каждой точке b_j^0 из B_0) сопоставляем все те точки Q_{a_i} из B_0 (все те точки $Q_{b_j^0}$ из A), которые находятся от нее на расстоянии не большем $R(A, B)$.

Отметим, что если следовать описанной выше процедуре построения искомого взаиморасположения, то для ключевых фрагментов \tilde{a} и \tilde{b} соответствие между их точками и при новом определении должно быть строго биективным. Однако это ограничение нетрудно снять рассмотрением довольно очевидных вырожденных подслучаев в рамках рассмотренных случаев первого, второго и третьего.

Однако в случае, когда точек в изображениях много (а реальное изображение, например, изображение на телевизионном экране — это сотни тысяч точек), даже полиномиальная зависимость от числа точек неприемлема, или, как минимум, нежелательна. В еще большей степени это обстоятельство присутствует применительно к глазу. Одна из особенностей зрительной информации — ее огромные объемы. Глаз человека содержит около 130 миллионов светочувствительных элементов (палочек и колбочек). Вместе с тем распознавание изображений осуществляется иногда лишь за доли секунды. Вряд ли это можно сделать, используя всю информацию на сетчатке глаза. Должно существовать некоторое «сито» и для изображений из среды, и для визуальной информации из памяти при оперировании с ними. Пара-

докс, однако, состоит в том, что попытки до собственно распознавания выделить, как это иногда делается, на изображениях «опорные точки», «важные детали», и пр. есть тоже распознавание, то есть возникает до некоторой степени замкнутый круг. Ясно, что нужно избегать этого. Поэтому сравнение изображений как таковых заменяется рассмотрением того, что можно назвать их набросками, эскизами, или приближениями. Эти эскизы могут состоять из гораздо меньшего числа точек, чем исходные изображения, и, вместе с тем, оценка схожести между эскизами определенным образом связывается с оценкой схожести между исходными изображениями. Распознавание при таком подходе перестает прямо зависеть от количества точек в изображениях и появляется возможность рассматривать изображения, бесконечные по числу составляющих их точек («непрерывные» изображения).

Показывается, что для функции, характеризующей похожесть изображений, выполняется неравенство треугольника. Это позволяет сократить перебор при сравнении распознаваемого изображения с изображениями, хранящимися в памяти.

Изображения могут быть весьма схожими и даже тождественными по форме, но существенно разными по размерам. За счет этого расстояние между ними может оказаться большим. Изометрические преобразования не дают возможность «уравнивать» изображения по размерам. Сделать это, однако, можно, рассматривая класс преобразований более широкий, чем изометрические — преобразования подобия [4]. И хотя в этом случае есть свои особенности, в целом предшествующие определения и построения не слишком меняются при расширении на преобразования подобия.

Положим теперь целью расширить рассматриваемые преобразования до аффинных. У этого случая есть специфика, которая состоит содержательно в том, что, в отличие от подобных и, тем более, изометрических преобразований, в целом аффинные преобразования не сохраняют форму в изображениях. Очевидно, что, например, чрезмерными сжатиями и растяжениями фигуру можно сделать неузнаваемой в сравнении с оригиналом. Кроме того, априори неясно, что

понимать под наиболее точным совмещением двух фигур, когда каждая из них рассматривается с точностью до аффинных преобразований.

Пусть изображение A состоит из точек a_1, \dots, a_n , изображение B — из точек b_1, \dots, b_n , ψ — взаимнооднозначное соответствие между точками изображений A и B , которым точке a_i из A ставится в соответствие точка $b_{\psi(i)}$ из B ($i = 1, \dots, n$). Обозначим через B^* множество всех изображений, получаемых из B аффинными преобразованиями. Полагаем, что на произвольном B' из B^* сохраняется нумерация, порожденная изображением B , то есть через b'_i на B' обозначена точка, в которую переходит при соответствующем преобразовании точка b_i из B . Точки a_i и $b'_{\psi(i)}$ называем соответствующими, соответствующими называем и отрезки $(a_i a_j)$ и $(b'_{\psi(i)} b'_{\psi(j)})$.

Зададимся некоторым положительным числом ε . Обозначим через $\{B\}^\varepsilon$ множество всех таких изображений B' из B^* , для которых длина каждого отрезка $(b_i b'_i)$ ($i = 1, \dots, n$) не больше ε . Преобразования, переводящие изображения из $\{B\}^\varepsilon$ друг в друга, назовем ε -аффинными. Содержательно их можно интерпретировать как некоторые ограниченные, локальные аффинные преобразования для B .

Через $l_A^\psi(B')$ обозначим длину наибольшего из отрезков $a_i b'_{\psi(i)}$ ($i = 1, \dots, n$). Рассмотрим B_0 — некоторое изображение из B^* , и ψ_0 — одно из взаимнооднозначных соответствий между точками изображений A и B . Пусть существует такое ε_1 , что для всех B' из $\{B_0\}^{\varepsilon_1}$ и при всех биекциях ψ минимум величин $l_A^\psi(B')$ достигается на изображении B_0 и при биекции ψ_0 . Пусть существует такое ε_2 , что для всякой пары изображений (A', B'_0) , получаемой ε_2 -аффинным преобразованием пары (A, B_0) как целого, выполняется аналогичное свойство: для всех B'' из $\{B'_0\}^{\varepsilon_1}$ и при всех биекциях ψ минимум величин $l_{A'}^\psi(B'')$ достигается на изображении B'_0 и при биекции ψ_0 . Тогда B_0 называем искомым изображением для изображения A , биекцию ψ_0 — искомым соответствием между точками в A и B . Величину $l_A^{\psi_0}(B_0)$ обозначаем через $R_A(B)$ и называем расстоянием от исходного изображения A до B .

Если, с одной стороны, ограничить ε -аффинные преобразования только преобразованиями изометрическими или подобными, а с дру-

гой — полагать ε_1 и ε_2 сколь угодно большими, то предшествующие определения искомого изображения B_0 и искомой биекции ψ_0 переходят в соответствующие понятия для преобразований изометрических или подобия.

Введем некоторое ограничение на рассматриваемые изображения. Будем полагать, что в них никакие два отрезка между точками изображения не параллельны друг другу (и, в частности, никакие три точки изображения не лежат на одной прямой). С содержательной точки зрения это не очень существенное ограничение. Действительно, пусть в изображении A есть параллельные отрезки. Рассмотрим круги радиуса δ , где δ — некоторое положительное число, с центрами в точках a_i изображения A ($i = 1, \dots, n$). Каким бы малым не было δ , всегда можно выбрать по одной точке a'_i в каждом круге так, что в изображении A' из точек a'_i ($i = 1, \dots, n$) уже не будет параллельных отрезков. Ясно, что в содержательном плане при достаточно малом δ изображение A' практически неотлично от A .

Назовем изображение B' из B^* согласованным с A , если существуют в B' два отрезка $(b'_{j_1} b'_{j_2})$ и $(b'_{j_3} b'_{j_4})$, равные, параллельные и однонаправленные с соответствующими отрезками $(a_{i_1} a_{i_2})$ и $(a_{i_3} a_{i_4})$ в A . Параллельные отрезки, например, $(a_{i_1} a_{i_2})$ и $(b'_{j_1} b'_{j_2})$, называем однонаправленными, если, при условии, что в $(a_{i_1} a_{i_2})$ слева направо сначала идет точка a_{i_1} , а затем a_{i_2} , то и в отрезке $(b'_{j_1} b'_{j_2})$ слева направо идет сначала точка b'_{j_1} , затем b'_{j_2} .

Пару произвольных отрезков $(a_{i_1} a_{i_2})$ и $(a_{i_3} a_{i_4})$ в A можно рассматривать как задающую «внутреннюю» систему координат: прямые, на которых лежат отрезки, определяют оси этой системы, сами отрезки — масштабные единицы по этим осям. Координаты точек изображения A в такой системе не зависят от аффинных преобразований изображения. В такой интерпретации согласованность B' с A означает «уравнивание» двух внутренних систем координат изображений A и B' , и, в этом смысле, как бы приведение их к общей системе.

Теорема 4. *Если B_0 — искомое изображение для A , то B_0 согласовано с A .*

Сочетание определяемого теоремой 4 условия с другим необходимым условием — совпадением центра характеристического изображения с центром ключевой окружности — и позволяет далее вычленишь из B^* конечное подмножество U изображений, среди которых только и может находиться искомое изображение.

Эти построения легко распространяются на трехмерный случай, когда трехмерное изображение, или тело, представлено конечным множеством точек в трехмерном евклидовом пространстве. Аналогичным двумерному случаю образом вводится понятие искомого изображения B_0 для заданного изображения A . И в трехмерном случае справедлив аналог теоремы 4, только согласование изображений A и B , в отличие от плоского случая, проводится не по двум, а по трем отрезкам.

Зрительную систему можно представлять как канал, связывающий внешнюю трехмерную среду с ее трехмерной внутренней моделью, заложенной в виде некоторого кода в нервной системе. Внутренняя модель играет существенную роль в распознавании, прогнозировании и организации поведения. Отличительной чертой канала связи является то, что входная информация о трехмерных объектах дана в виде двух их проекций на сетчатку. Отсюда возникает задача о кодировании плоских и объемных изображений, о характере математических процедур, сопровождающих зрительный механизм восстановления трехмерных объектов по их проекциям на сетчатку.

Изображение на сетчатке непрерывно смещается, поворачивается, изменяется в размерах, сжимается, растягивается за счет движений объекта перед глазом и самого глаза, постоянных микроизменений позы и из-за тремора. Аналогичные моменты присутствуют и в проблемах, связанных с машинным зрением и робототехникой. В целом это приводит к задаче такого описания изображений, которое было бы инвариантно к аффинным преобразованиям его на плоскости.

Назовем двумерным изображением конечное множество точек на плоскости. Перенумеруем некоторым образом точки изображения A так, чтобы номера были попарно различны. Обозначим через M_A множество этих номеров. Пусть S_{mnu} и S_{ksp} — площади

треугольников с вершинами в тройках точек с номерами m, n, u и k, s, p и пусть $\rho_{mnu, ksp} = S_{mnu}/S_{ksp}$. Полагаем, что порядок номеров в тройках не важен, сами тройки различны и при $S_{ksp} = 0$ значение $\rho_{mnu, ksp}$ не определено. Множество индексированных чисел $\rho_{mnu, ksp}$ для всех таких пар троек обозначим через T_A . Код изображения A — пара $\langle M_A, T_A \rangle$. Изображения, все точки которых расположены на одной прямой, не рассматриваем, поскольку код для них не определен. Изображения A и B с кодами $\langle M_A, T_A \rangle$ и $\langle M_B, T_B \rangle$ назовем эквивалентными, если существует такая биекция $\psi : M_A \rightarrow M_B$, что для любых m, n, u, k, s, p из M_A выполнено $\rho_{mnu, ksp} = \rho_{\psi(m)\psi(n)\psi(u), \psi(k)\psi(s)\psi(p)}$. Ясно, что эквивалентность изображений содержательно означает одинаковость их кодов с точностью до перенумерации точек. Два изображения называем аффинно эквивалентными, если они переводимы друг в друга аффинными преобразованиями. Изображение называем плоским, если все его точки не лежат на одной прямой или на двух параллельных прямых.

Теорема 5. *Два плоских изображения эквивалентны тогда и только тогда, когда они аффинно эквивалентны.*

Из теоремы 5 следуют простые и эффективные способы распознавания изображений, отличающихся друг от друга произвольными аффинными преобразованиями.

Пусть i_1, \dots, i_k — номера k ($k \geq 3$) точек изображения A . Рассмотрим выпуклый многоугольник, включающий все эти точки, часть из которых образует его множество вершин. Ясно, что такой многоугольник определяется однозначно. Пусть S_{i_1, \dots, i_k} — площадь этого многоугольника, которую называем площадью k -точечника. Если точки i_1, \dots, i_k расположены на одной прямой, то полагаем $S_{i_1, \dots, i_k} = 0$. Пусть j_1, \dots, j_k — другой k -точечник из A . Полагаем $\rho_{i_1, \dots, i_k, j_1, \dots, j_k} = S_{i_1, \dots, i_k}/S_{j_1, \dots, j_k}$. При этом, если $S_{j_1, \dots, j_k} = 0$, то $\rho_{i_1, \dots, i_k, j_1, \dots, j_k}$ считаем не определенным. Множество всех таких индексированных чисел $\rho_{i_1, \dots, i_k, j_1, \dots, j_k}$ обозначаем через T_A . Кодом изображения A называем пару $\langle M_A, T_A \rangle$. Изображения A и B с кодами $\langle M_A, T_A \rangle$ и $\langle M_B, T_B \rangle$ называем k -эквивалентными, если существует

такая биекция $\psi : M_A \rightarrow M_B$, что для любых $i_1, \dots, i_k, j_1, \dots, j_k$ из M_A выполнено $\rho_{i_1, \dots, i_k, j_1, \dots, j_k} = \rho_{\psi(i_1), \dots, \psi(i_k), \psi(j_1), \dots, \psi(j_k)}$.

Можно продолжить определение k -эквивалентности и на случай $k=2$, понимая под $S_{i_1 i_2}$ расстояние между точками с номерами i_1 и i_2 .

Отношения k -эквивалентности и аффинной эквивалентности являются отношениями эквивалентности. Порождаемые ими на множестве всех плоских изображений разбиения на классы эквивалентности обозначим через соответственно P^k и P^a .

Теорема 6. *Разбиения P^k ($k \geq 2$) и P^a совпадают только при $k=3$.*

Назовем трехмерным изображением A или телом конечное множество точек в трехмерном евклидовом пространстве. Занумеруем попарно различными номерами точки изображения A . Пусть \bar{M}_A — множество этих номеров, V_{mnuv} и V_{kspq} — объемы тетраэдров с вершинами в четверках точек с номерами m, n, u, v и k, s, p, q , и пусть $\rho_{mnuv, kspq} = V_{mnuv}/V_{kspq}$. Полагаем, что порядок номеров в четверках не важен, сами четверки различны, и для случая, когда $V_{kspq} = 0$, $\rho_{mnuv, kspq}$ не определено. Множество индексированных чисел $\rho_{mnuv, kspq}$ для всех таких пар четверок обозначим через \bar{T}_A . Кодом тела A назовем пару $\langle \bar{M}_A, \bar{T}_A \rangle$. Тела A и B с кодами $\langle \bar{M}_A, \bar{T}_A \rangle$ и $\langle \bar{M}_B, \bar{T}_B \rangle$ назовем эквивалентными, если существует такая биекция $\psi : \bar{M}_A \rightarrow \bar{M}_B$, что для любых m, n, u, v, k, s, p, q из \bar{M}_A выполнено $\rho_{mnuv, kspq} = \rho_{\psi(m)\psi(n)\psi(u)\psi(v), \psi(k)\psi(s)\psi(p)\psi(q)}$. Тела, все точки которых расположены в одной плоскости, называем двумерными и для них рассматриваемый код не определен. Тела называем аффинно эквивалентными, если они переводятся друг в друга аффинными преобразованиями. Трехмерное изображение назовем объемным, если все его точки не лежат в одной плоскости или в двух параллельных плоскостях.

Теорема 7. *Два объемных изображения эквивалентны тогда и только тогда, когда они аффинно эквивалентны.*

Восстановление трехмерного изображения по плоским проекциям служит, с одной стороны, предположительной основой механизмов

стереоскопического зрения в живых организмах, с другой — является важной задачей в рамках машинного зрения для робототехники. В описанных схемах такого восстановления точка m проецируется на две плоские сетчатки (рис. 2), проекции ее есть точки соответственно m_1 и m_2 . Если известно положение этих точек на сетчатках, то, используя геометрические соображения и построения, можно восстановить положение точки m . Если тело T состоит из конечного множества точек, то, восстанавливая положение каждой точки, можно восстановить поточечно все тело. Проекции S_1 и S_2 тела на две сетчатки несколько разные за счет того, что каждый глаз «видит» тело под своим углом зрения, в своем ракурсе. Именно этой разностью и обеспечивается возникновение стереоскопического эффекта.

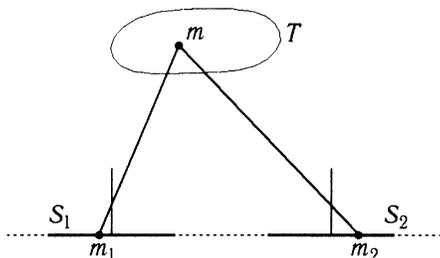


Рис. 2.

Главная проблема в рамках машинного стереозрения — это проблема идентификации соответствующих друг другу точек на двух проекциях. На рис. 2 изображены по одной точке на каждой сетчатке. Когда таких точек много, то неясно, какую из них на одной проекции сопоставлять данной точке на другой. Это нельзя сделать, например, простым наложением изображений на сетчатках друг на друга, поскольку эти изображения разные за счет разных ракурсов. Предполагать, что уже «распознано», какие части изображений соответствуют друг другу, и на этой основе сопоставлять точки проекций тоже нельзя, так как задача распознавания предполагается решаемой на более поздних этапах и ее решение отчасти должно основываться на результатах восстановления трехмерного изображения.

Дополнительные трудности возникают при попытках объяснить этой схемой механизмы стереоскопического зрения в живых организмах. Действительно, для того, чтобы знать расстояние между соответствующими точками на двух проекциях, нужно, как минимум, иметь известным расстояние между двумя сетчатками. Неясно, на каком основании мы можем считать это расстояние априори известным. Внутри одного вида расстояния между глазами у разных особей, очевидно, несколько разнятся, не говоря уже об особях разных видов. На протяжении жизни, вследствие изменений в размерах тела, это расстояние тоже меняется. Трудно поэтому считать его жестко генетически обусловленным и потому известным. Если же предполагать, что это расстояние становится известным из индивидуального опыта, то надо каким-то образом объяснить механизм получения такого рода информации. Кроме того, сетчатки на рис. 2 изображены расположенными в одной плоскости. Однако они могут быть «повернуты на объект», то есть развернуты каждая перпендикулярно лучу от точки m . Изображения на сетчатках при этом по сравнению с рис. 2 подвергнутся сжатию. Если рассматриваемый объект расположен сбоку и глаза повернуты в его сторону, то одна сетчатка расположена несколько дальше от объекта, чем другая, что приводит к некоторой разнице в размерах проекций. В целом рассуждения такого рода приводят к необходимости восстанавливать трехмерное изображение не только по данной паре S_1 и S_2 его плоских проекций, но и по любой паре \tilde{S}_1 и \tilde{S}_2 , полученной из соответственно S_1 и S_2 аффинными преобразованиями.

Рассмотрим тело T и прямую, называемую направлением проекции. Направления проекции назовем разными, если они не параллельны. Проведем через каждую точку тела T прямые, параллельные направлению проекции α и называемые лучами. Полагаем α таким, что на каждом луче находится только одна точка тела. Таких направлений проекции бесконечное множество, не таких — только конечное. Назовем плоскость, пересекающую лучи, плоскостью проекции, изображение, образованное точками пересечения лучей с плоскостью проекции — проекцией тела (на данную плоскость и по данному направлению). Рассматриваем проекции тела T по разным направлениям и

на разные плоскости. Оговорим, что если T — двумерное изображение, то α полагаем не параллельным плоскости этого изображения. Взаимнооднозначное соответствие между точками двух изображений назовем их разметкой. Соответствующие друг другу точки будем обозначать одной буквой (с разными индексами). Ясно, что описанным выше устанавливается взаимнооднозначное соответствие между точками тела T и точками проекций S_i ($i = 1, 2, \dots$). Если α — точка тела T , то точку проекции S_i , лежащую с ней на одном луче, обозначим через a_i и будем называть проекцией точки a . Это устанавливает и взаимнооднозначное соответствие между точками проекций S_i и S_j : соответствующие друг другу точки являются проекциями одной и той же точки тела T . Размеченные изображения A и B назовем a' -эквивалентными, если можно перевести их одно в другое аффинными преобразованиями так, что совместятся соответствующие друг другу точки (обозначение: $A \approx B$). В противном случае A и B назовем a' -разными (обозначение: $A \not\approx B$). Часть изображения A , состоящую из его точек a, b, \dots, v будем обозначать как $A(a, b, \dots, v)$. Три точки тела, не лежащие на одной прямой, назовем гранью, определяемую этими точками плоскость — плоскостью грани. Три точки проекции, не лежащие на одной прямой, назовем треугольником.

Имея тело T , заданное направление проекции α и меняя плоскости проекции, можно получить некоторое множество $\{S\}$ проекций. Все проекции из $\{S\}$ будут попарно a' -эквивалентными. С другой стороны, тело T — не единственное, проецированием которого можно получить множество $\{S\}$ проекций. Таким будет, например, тело T' , полученное заменой каждой точки x тела T , находящейся на луче α_x проецирования, на какую-либо другую точку x' на том же луче. В частном и вырожденном случае все точки тела T могут находиться и в одной плоскости. Итак, при заданном направлении проекции получить данное множество $\{S\}$ проекций можно проецированием некоторого множества $\{T\}$ тел. Из этого следует, что имея одну или несколько проекций из множества $\{S\}$, нельзя восстановить тело T . Мало того, нельзя даже распознать, не имеем ли мы дело с вырожденным случаем, когда T — двумерное изображение. Проекция из $\{S\}$ можно интерпретировать как изображения тела T в одном ракурсе.

Следовательно, для того, чтобы восстановить тело, или даже только определить, не двумерное ли оно, нужно иметь более чем одну проекцию, причем в разных ракурсах (то есть при разных направлениях проекции).

Пусть S_1 и S_2 — проекции тела T по двум разным направлениям α_1 и α_2 . Доказано, что S_1 и S_2 a' -разные тогда и только тогда, когда T — не двумерное изображение. Это позволяет ввести следующее определение. Будем говорить, что точка d_1 на S_1 лежит в плоскости треугольника $a_1b_1c_1$, если $S_1(a_1, b_1, c_1, d_1) \approx S_2(a_2, b_2, c_2, d_2)$. В этом случае и d_2 лежит в плоскости треугольника $a_2b_2c_2$. В противном случае говорим, что d_1 лежит вне плоскости треугольника $a_1b_1c_1$ (d_2 лежит вне плоскости треугольника $a_2b_2c_2$).

Рассмотрим на S_1 и S_2 четверку троек точек $\langle (a_1b_1c_1)(a_2b_2c_2)(c_1d_1e_1)(c_2d_2e_2) \rangle$ такую, что $a_1b_1c_1$ и $a_2b_2c_2$ — треугольники, точки d_1 и e_1 лежат вне плоскости треугольника $a_1b_1c_1$ (и, соответственно, точки d_2 и e_2 лежат вне плоскости треугольника $a_2b_2c_2$), и из троек $c_1d_1e_1$ и $c_2d_2e_2$ хотя бы одна — треугольник. Такую четверку назовем правильной. Ее точки на S_1 и S_2 являются проекциями точек a, b, c, d, e тела T , причем abc и cde — грани, и плоскости этих граней разные. Обозначим через L линию пересечения этих плоскостей, через L_1 и L_2 — проекцию этой линии на S_1 и S_2 . Поскольку \tilde{S}_1 и \tilde{S}_2 получены из S_1 и S_2 аффинными преобразованиями, то через \tilde{L}_1 и \tilde{L}_2 обозначим соответственно трансформированные L_1 и L_2 на них. Доказано, что по \tilde{S}_1 и \tilde{S}_2 можно определить положение \tilde{L}_1 и \tilde{L}_2 на них.

Опишем процедуру $\text{Alg } T'$ построения некоторого тела T' по \tilde{S}_1 и \tilde{S}_2 . Пусть $a_1b_1c_1$ и $a_2b_2c_2$ — треугольники и точка e_1 лежит вне плоскости треугольника $a_1b_1c_1$. Выберем некоторую прямую α' (не параллельную плоскости изображения \tilde{S}_1) в качестве направления проекции. Проведем через точки a_1, b_1, c_1, e_1 лучи, параллельные α' , и на каждом луче возьмем по точке соответственно a', b', c', e' так, чтобы они не лежали в одной плоскости. Пусть теперь d_1 — произвольная точка на \tilde{S}_1 . Построим соответствующую ей точку d' тела T' .

1) Пусть d_1 лежит в плоскости треугольника $a_1b_1c_1$. Совместим аффинными преобразованиями изображения \tilde{S}_1 точки a_1, b_1, c_1 с точками соответственно a', b', c' . Точка, в которую перейдет d_1 , есть d' .

2) Пусть d_1 лежит вне плоскости треугольника $a_1b_1c_1$. Пусть хотя бы одна из троек $c_1d_1e_1$ и $c_2d_2e_2$ — треугольник. Тогда $\langle (a_1b_1c_1)(a_2b_2c_2)(c_1d_1e_1)(c_2d_2e_2) \rangle$ — правильная четверка. Построим по \tilde{S}_1 и \tilde{S}_2 прямые \tilde{L}_1 и \tilde{L}_2 на них. Аффинными преобразованиями изображения \tilde{S}_1 совместим точки a_1, b_1, c_1 с точками соответственно a', b', c' . Прямую, в которую преобразуется \tilde{L}_1 , обозначим через L . Пусть, для определенности, $c_2d_2e_2$ — треугольник. Аффинными преобразованиями изображения \tilde{S}_2 (с прямой \tilde{L}_2) совместим \tilde{L}_2 с L и точку e_2 с точкой e' . Точка, в которую перейдет при этом d_2 , есть d' .

Пусть теперь и $c_1d_1e_1$, и $c_2d_2e_2$ — не треугольники. Возможны две ситуации.

2а) Точки c, d, e в теле T лежат на одной прямой. На \tilde{S}_1 и \tilde{S}_2 это проявляется тем, что в каждой из четверок a_1, c_1, e_1, d_1 и b_1, c_1, e_1, d_1 точки лежат в одной плоскости. В свою очередь, точки a_1, c_1, e_1, d_1 лежат в одной плоскости, если, при $a_1c_1e_1$ и $a_2c_2e_2$ — треугольниках, $\tilde{S}_1(a_1, c_1, e_1, d_1) \approx \tilde{S}_2(a_2, c_2, e_2, d_2)$.

Если же, например, $a_1c_1e_1$ — не треугольник (то есть направление проекции α_1 параллельно плоскости грани ace), то все четыре точки a_1, c_1, e_1, d_1 должны лежать на одной прямой. Аналогично определяется принадлежность одной плоскости и точек b_1, c_1, e_1, d_1 .

Аффинным преобразованием изображения \tilde{S}_1 совмещаем точки c_1 и e_1 с точками c' и e' . Точка, в которую переходит d_1 , есть d' .

2б) Точки c, d, e в теле T образуют грань, и направления α_1 и α_2 оба параллельны этой грани. Если точка d лежит на одной прямой с точками a и e , или b и e , то построение d' сводится к пункту 2а. Если aed и bed — грани, то, например, четверка $\langle (a_1b_1c_1)(a_2b_2c_2)(a_1e_1d_1)(a_2e_2d_2) \rangle$ — правильная, и построение точки d' сводится к началу пункта 2.

Теорема 8. Если S_1 и S_2 суть a' -разные проекции тела T , $S_1 \approx \tilde{S}_1$, $S_2 \approx \tilde{S}_2$ и тело T' построено по \tilde{S}_1 и \tilde{S}_2 использованием $\text{Alg } T'$, то тела T и T' a' -эквивалентны.

Содержательно теорема 8 состоит в следующем. В процедуре $\text{Alg } T'$ присутствует много моментов, которые могут варьироваться. Можно, например, выбирать разные \tilde{S}_1 и \tilde{S}_2 , по разному брать исходную четверку точек на \tilde{S}_1 (или на \tilde{S}_2) и т. д. Варьируя такие моменты, можно построить некоторое множество $\{T\}$ тел. Теорема, однако, утверждает, что все они a' -эквивалентны телу T и, значит, a' -эквивалентны между собой. Тело посредством процедуры $\text{Alg } T'$ восстанавливается, тем самым, с точностью до аффинных его преобразований. При этом нет необходимости знать расстояние между проекциями S_1 и S_2 (аналог расстояния между сетчатками глаз), тело T' строится по произвольным образом сдвинутому, повернутому, сжатому или растянутому, уменьшенному или увеличенному проекциям тела T .

Выше уже отмечалось, что по одной проекции восстановить тело нельзя. Из теоремы 8 следует, что для этого достаточно (с точностью до аффинных преобразований) двух проекций. Можно предположить, что это имеет некоторое отношение к тому обстоятельству, что зрительное восприятие повсеместно осуществляется посредством именно двух глаз.

Работа с восстановленным телом может включать в себя построение проекций по произвольным направлениям, в том числе и таким, для которых на каждом луче проекции может быть более чем одна точка. Такие проекции называем неполными. Конечно, имея восстановленное тело, можно построить произвольную его проекцию. Показано, однако, что любую проекцию, в том числе и неполную, можно получить непосредственно по \tilde{S}_1 и \tilde{S}_2 , не прибегая к восстановлению тела в качестве, так сказать, промежуточного этапа. Такие построения могут оказаться полезными, например, в компьютерных играх.

До сих пор предполагалось, что точки на \tilde{S}_1 и \tilde{S}_2 размечены. Положим теперь, что разметка неизвестна. Если каждое из \tilde{S}_1 и \tilde{S}_2 состоит из n точек, то возможны $n!$ вариантов разметки. Пусть при данном варианте разметки точкам x_{i_1}, \dots, x_{i_n} на \tilde{S}_1 ставятся в соответствие точки y_{j_1}, \dots, y_{j_n} на \tilde{S}_2 . Пусть существует тело T' из точек z_1, \dots, z_n и его проекции S'_1 и S'_2 такие, что $S'_1 \approx \tilde{S}_1$, $S'_2 \approx \tilde{S}_2$, и точки x_{i_k} и y_{j_k} являются проекциями точки z_k , $k = 1, \dots, n$. Тогда

данный вариант разметки назовем приемлемым или решением. По условию, одно решение (с исходным телом T) существует, однако в общем случае решение не единственное. Пример нетрудно построить с использованием симметричного объемного тела.

Опишем процедуру проверки данной разметки на приемлемость. Выберем на \tilde{S}_1 и \tilde{S}_2 точки $x_{i_a}, x_{i_b}, x_{i_c}, x_{i_e}$ и $y_{j_a}, y_{j_b}, y_{j_c}, y_{j_e}$ такие, что $x_{i_a}x_{i_b}x_{i_c}$ и $y_{j_a}y_{j_b}y_{j_c}$ — треугольники и $\tilde{S}_1(x_{i_a}, x_{i_b}, x_{i_c}, x_{i_e}) \not\approx \tilde{S}_2(y_{j_a}, y_{j_b}, y_{j_c}, y_{j_e})$. Проведем лучи, параллельные некоторому направлению α' через все точки изображения \tilde{S}_1 . На лучах, проходящих через $x_{i_a}, x_{i_b}, x_{i_c}, x_{i_e}$ возьмем точки z'_a, z'_b, z'_c, z'_e так, чтобы они не лежали в одной плоскости. Если разметка приемлемая, то далее использованием $\text{Alg } T'$ можно было бы построить остальные точки тела T' и они лежали бы на лучах от соответствующих точек на \tilde{S}_1 . Затем аналогичное сделаем для \tilde{S}_2 : проведем лучи по направлению α'' , на лучах через точки $y_{j_a}, y_{j_b}, y_{j_c}, y_{j_e}$ отметим точки $z''_a, z''_b, z''_c, z''_e$, не лежащие в одной плоскости, и далее, если разметка приемлемая, можно построить тело T'' . Возьмем теперь всю конструкцию из точек изображения \tilde{S}_1 и лучей проекции через эти точки, и аффинными преобразованиями ее как целого совместим точки z'_a, z'_b, z'_c, z'_e с точками $z''_a, z''_b, z''_c, z''_e$. Если тела T' и T'' существуют, то, в силу теоремы 8, они a' -эквивалентны, значит, они совместились бы полностью, и в каждой их точке будут сходиться пары лучей от соответствующих точек на \tilde{S}_1 и \tilde{S}_2 (и только они).

Отметим, что не все возможные $n!$ вариантов разметок нужно проверять на приемлемость. Пусть, например, при данных четверках $x_{i_a}, x_{i_b}, x_{i_c}, x_{i_e}$ и $y_{j_a}, y_{j_b}, y_{j_c}, y_{j_e}$ лучи проекции от некоторой точки x на \tilde{S}_1 и сопоставленной ей (при данной разметке) точки y на \tilde{S}_2 не сходятся. Тогда они не будут сходиться и при всех вариантах разметок, в которых сопоставляются друг другу точки данных четверок и точка x — точке y . Все такие разметки могут быть заранее исключены как неприемлемые. С учетом этого оценка числа вариантов разметки для проверки на приемлемость может быть понижена до n^4 .

Отметим, что такие построения могут представлять интерес и в связи с проблемами томографии, поскольку в этом случае тоже происходит восстановление трехмерного тела по проекциям.

Автор выражает глубокую благодарность В. Б. Кудрявцеву, который был научным руководителем этой работы, начиная со студенческих времен ее автора, за внимание и поддержку работы на всем ее протяжении.

Список литературы

- [1] Kozlov V.N. Image coding and recognition and some problems of stereovision // Pattern Recognition and Image Analysis. V. 7. N 4. 1997. P. 448–466.
- [2] Козлов В. Н. О распознавании аффинно разных дискретных изображений // Интеллектуальные системы. 1998. Т. 3. Вып. 3–4. С. 95–122.
- [3] Козлов В. Н. О зрительном образе, математических подходах к определению этого понятия и о распознавании изображений // Журнал вычислительной математики и математической физики. 1999. Т. 39. № 11. С. 1929–1946.
- [4] Козлов В. Н. Элементы математической теории зрительного восприятия. М.: Издательство Центра прикладных исследований при механико-математическом факультете МГУ, 2001.