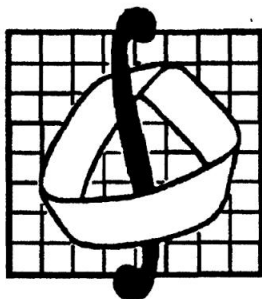


**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА**



Механико-математический факультет

**ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ
ТЕОРИИ ЗРИТЕЛЬНОГО
ВОСПРИЯТИЯ**

В.Н. Козлов

Москва 2001 год

В.Н. Козлов
Элементы математической теории зрительного
восприятия

Книга рассчитана на использование в качестве пособия для специальных курсов, читаемых на кафедре математической теории интеллектуальных систем механико-математического факультета МГУ им. М.В. Ломоносова

Механико-математический факультет МГУ,

2001 г.

Оглавление

Введение	4
Глава I. Кодирование изображений и некоторые проблемы стереозрения	7
§1. Кодирование плоских изображений	8
§2. Кодирование объемных изображений	26
§3. Восстановление объемного изображения по плоским проекциям	36
Глава II. Зрительный образ и геометрические преобразования изображений	53
§1. Преобразования параллельного переноса	57
§2. Изометрические преобразования изображений	66
§3. Преобразования подобия изображений	78
§4. Аффинные преобразования изображений	93
Глава III. Сжатие информации при обработке и распознавании изображений	101
§1. Об определении искомого взаиморасположения двух изображений	102
§2. Приближение изображений их покрытиями	107
§3. Сокращение перебора в памяти	112
§4. Распознавание изображений, заданных с точностью до подобия	114
§5. Распознавание изображений, заданных с точностью до аффинных преобразований	116
Литература	120

Введение

Зрение занимает особое место в ряду других органов чувств человека. По некоторым оценкам биологов до 90% сенсорной (то есть от органов чувств) информации составляет зрительная информация [1]. В немалой мере и собственно мышление, можно предположить, имеет в основе своей оперирование со зрительной информацией из окружающей среды и памяти. Столь же очевидно значима визуальная информация и для работы технических устройств: роботов, компьютеров, систем наведения на цели и пр.

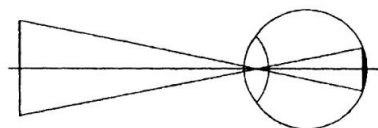


Рис. 1.

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. Сетчатка состоит из рецепторов, возбуждающихся при попадании на них света. От каждого рецептора отходит нервное волокно, позволяющее регистрировать, возбужден соответствующий рецептор или нет.

Вопрос, который нас здесь интересует, состоит в следующем: 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.

Изображение у нас – конечное (непустое) множество точек на плоскости. Такое представление об изображении может показаться искусственным, отвлеченным от того, чем являются реальные изображения. Тем не менее многие реальные изображения – газетные фотографии, распечатки матричного принтера, картинки на экране черно-белого телевизора – представляют собой совокупности точек. Любое реальное черно-белое

изображение может быть «аппроксимировано» изображением из точек, с передачей различной плотностью точек в разных частях изображения всех градаций серого цвета. В какой-то мере черно-белый телевизор можно рассматривать как техническое устройство, которое выполняет такого рода аппроксимацию. Представление об изображении как о совокупности точек не закрывает дорогу и к работе с цветными изображениями: рассмотрение трех монохроматических (в основных цветах) изображений эквивалентно рассмотрению цветного изображения. Наконец, все, что мы видим, мы видим посредством глаз. Изображение из среды проецируется на сетчатку глаз, что приводит к возбуждению части рецепторных клеток сетчатки, то есть, в конечном счете, к формированию на сетчатке составленного из точек изображения. Именно оно и анализируется в последующем мозгом.

В целом работа может быть интерпретирована применительно к трем областям: распознаванию изображений, компьютерному зрению и математическому моделированию зрительного восприятия в живых организмах. Эти три интерпретации имеются ввиду и присутствуют (в разной, может быть, степени) далее в тексте.

Материал работы предполагается к использованию в специальных курсах, читаемых на кафедре математической теории интеллектуальных систем механико-математического факультета МГУ.

Глава I

Кодирование изображений и некоторые проблемы стереозрения

Если перед глазом находится некоторый объект, то изображение его на сетчатке постоянно меняется. Это может происходить по причине движения либо объекта, либо глаза, но, главным образом, вследствие постоянных микродвижений глаз и микроизменений позы. Если объект плоский, то изменения на сетчатке при таких движениях сводятся при определенных допущениях к аффинным преобразованиям. В §1 рассмотрены разные варианты кодирования изображения на сетчатке, на основе которых изображения могут распознаваться при различных преобразованиях, вплоть до аффинных.

Если в §1 рассмотрены коды плоских изображений, то в §2 рассматриваются коды пространственных изображений, инвариантных, в определенном смысле, к различным преобразованиям, вплоть до аффинных.

Можно предполагать, что трехмерная среда порождает трехмерное же ее отражение («модель среды») в нервной системе (конечно, в неявном «закодированном» виде) и что именно наличие «модели среды» позволяет на ее основе организовать распознавание и целенаправленное поведение [6]. Вместе с тем внешний мир воспринимается через посредство плоских проекций на сетчатку. Значит, возникает задача построения трехмерного изображения по его плоским проекциям. Именно это и делается в §3.

Ясно, что математическое моделирование зрительного восприятия прямо соотносится с проблемами анализа изображений и распознавания в рамках компьютерного зрения и робототехники. Основной проблемой в компьютерном стереозрении считается проблема идентификации соответствующих друг другу точек на двух проекциях трехмерного изображения. В

§3 описано точное, в рамках модели, решение этой проблемы. Здесь же рассмотрена задача, возникающая при оперировании с трехмерной внутренней моделью внешней среды. Если считать, что это оперирование основывается на как бы рассмотрении ее «внутренним взором», то это означает получение различных вариантов плоских проекций этой внутренней трехмерной модели. Можно, конечно, считать, что по имеющемуся трехмерному изображению всегда можно построить любую его плоскую проекцию. В §3 показывается, однако, что по паре плоских проекций (исходных, с сетчатки) можно построить произвольную проекцию, не восстанавливая (в качестве, так сказать, промежуточного этапа) собственно трехмерное тело.

§1. О кодировании и распознавании плоских изображений

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.

Если объект перед глазом заменить на другой, то совокупность возбужденных рецепторов изменится. Это и есть основа для того, чтобы в процессе последующего распознавания констатировать изменение объекта. Есть, однако, ситуации, когда объект перед глазом как таковой не меняется, а меняется только его положение. Пример такой ситуации – сдвиг объекта перед глазом, что приводит к параллельному переносу изображения на сетчатке. При этом изображение на сетчатке осталось, в сущности, прежним, изменилось лишь его местоположение, однако первичный материал – совокупность возбужденных рецепторов – может, очевидно, существенно поменяться.

Вопрос: как осуществляется распознавание таких ситуа-

ций? Можно, конечно, предполагать, что распознавание в этих случаях осуществляется точно также, как и в общем случае, когда объект перед глазом действительно поменялся. Мы, однако, исходим здесь из предположения, что существуют отдельные процедуры для распознавания таких простых изменений, как сдвиг, поворот и некоторых других, и что эти процедуры существенно проще, чем процедура распознавания в общем случае. При изменении совокупности возбужденных рецепторов сначала происходит проверка первичного материала на предмет выяснения того, не сводится ли изменение изображения к сдвигу, повороту и пр. И только при отрицательном ответе начинает работать общая процедура распознавания. Эти простые процедуры можно рассматривать и как первый этап в общей процедуре распознавания.

Однако такое сильное предположение про зрительное восприятие – введение дополнительного этапа в распознавание – надо как-то обосновать. Некоторыми доводами в пользу такого предположения могут быть следующие соображения, учитывающие отдельные особенности условий, в которых происходит работа глаза.

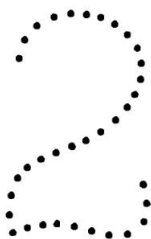


Рис. 2.

В качестве простых преобразований изображения на сетчатке, для которых и существует первый этап в распознавании, мы будем рассматривать сдвиг (параллельный перенос) изображения, поворот, подобное преобразование, сжатие, растяжение и, разумеется, все их сочетания (то есть аффинные преобразования). Этот набор преобразований взят не произвольно, а потому, что эти преобразования получаются при определен-

ных перемещениях объекта перед глазом. Действительно, будем считать, что объект перед глазом – это плоское точечное изображение некоторой фигуры, например, цифры (рис. 2).

В реальном эксперименте такую фигуру можно изобразить точками черной краски на плоской прозрачной стеклянной пластинке. Сетчатку полагаем плоской. Изображение помещается перед глазом на некотором расстоянии так, что его плоскость была параллельна плоскости сетчатки. Если теперь изображение перед глазом сдвинуть (то есть сместить его, сохраняя неизменным расстояние до глаза и сохраняя плоскость изображения параллельной плоскости сетчатки), то и изображение на сетчатке сдвинется (преобразование в виде параллельного переноса).

Если изображение перед глазом повернуть, сохраняя его плоскость параллельной плоскости сетчатки, то и изображение на сетчатке повернется.

Если изображение перед глазом придвинуть к глазу, сохраняя его плоскость параллельной плоскости сетчатки, то изображение на сетчатке увеличится с сохранением подобия. Если изображение перед глазом отодвигать, то изображение на сетчатке уменьшится.

Проведем теперь прямую в плоскости изображения перед глазом и, для наглядности, проходящую через изображение. Повернем на некоторый угол плоскость изображения вокруг этой прямой как вокруг оси. Изображение на сетчатке при этом сожмется в направлении, перпендикулярном проекции этой оси. Обратное вращение вокруг оси изображения перед глазом заставит, очевидно, его проекцию на сетчатке вернуться, растягиваясь, к исходной форме. Надо, однако, отметить, что только к сжатию или к растяжению преобразования на сетчатке сведутся лишь при достаточно малом отношении размеров изображения в направлении, перпендикулярном оси поворота, к расстоянию до глаза. В противном случае возникают искажения на сетчатке, связанные с существенностью разницы в расстояниях до глаза разных частей изображения перед глазом после поворота. Мы полагаем, что условия таковы, что мы можем пренебречь этими искажениями.

Набор перемещений изображения перед глазом, который мы

рассмотрели, полон в том смысле, что любое перемещение плоского изображения перед глазом можно рассматривать как сочетание описанных перемещений. Значит для того, чтобы иметь возможность отслеживать любое перемещение плоского изображения перед глазом, нужно уметь отслеживать на сетчатке сдвиг, поворот, подобное преобразование, сжатие, растяжение и любые их сочетания.

Но это только одна сторона дела. Мы полагали до сих пор, что глаз неподвижен, а положение распознаваемого объекта меняется. Положим теперь, что, наоборот, объект неподвижен, а глаз перемещается. Очевидно, что все сказанное выше о преобразованиях изображения на сетчатке приложимо и к этому случаю. Так, например, если объект неподвижен, а глаз придвигается к нему, причем плоскость объекта и плоскость сетчатки параллельны, то изображение на сетчатке будет увеличиваться в размерах с сохранением подобия. Можно сказать, что при всех таких перемещениях не важно, что в действительности перемещается — глаз или объект. Они могут, в частности, и одновременно менять свое положение. Важно лишь изменение взаимного расположения глаза и объекта. Оно же сводится к тем перемещениям, которые мы рассмотрели выше и, следовательно, к соответствующим преобразованиям изображения на сетчатке.

Следующий вопрос — о том, сколь часто приходится отслеживать такого рода преобразования на сетчатке. И вот здесь можно сказать, что если о неподвижном объекте перед глазом еще говорить можно, то сам глаз неподвижным практически никогда не бывает. Человек, даже если он самому себе кажется неподвижным, на самом деле отнюдь не неподвижен хотя бы из-за непрерывно происходящих микроизменений позы. Сколь бы малы они не были, к некоторому изменению положения изображения на сетчатке они приводят, а значит и к новой совокупности возбужденных рецепторов. К этому можно добавить и изменения, вызванные тремором (микродрожжанием глаз), роль которого в полной мере пока не ясна, но который, несомненно, тоже меняет положение изображения на сетчатке.

Итак, взаимное положение объекта перед глазом и глаза непрерывно меняется, приводя к таким изменениям изображения

на сетчатке, которые мы описали. Новые же объекты, требующие распознавания по существу, появляются, можно считать, гораздо реже. Поэтому задача распознавания описанных выше преобразований изображения на сетчатке должна решаться гораздо чаще, чем задача распознавания нового объекта. Это и служит нам основанием для выделения распознавания сдвига, поворота, сжатия и пр. на сетчатке в отдельный этап. Алгоритмы этого выделенного этапа мы полагаем более простыми, чем общая процедура распознавания, которая, по-видимому, достаточно нетривиальна и требует для своего осуществления каких-либо затрат (что бы под ними не понимать, например, затрат времени). Это, конечно, не исключает и того, что алгоритмы выделенного этапа могут оказаться упрощенным вариантом общей процедуры распознавания, ее частным случаем.

Мы рассмотрим далее разные варианты преобразований изображения на сетчатке и для каждого такого случая будем строить кодировку изображения, инвариантную к рассматриваемым преобразованиям. Распознавание будет основываться на этих инвариантах.

Случай 1. Этот случай соответствует, в сущности, представлению того, что выше было названо первичным материалом. Полагаем, что перед глазом – плоское изображение, составленное из конечного числа точек (пример на рис. 2). Точки, из которых состоит изображение, будем считать, невелики по размерам и потому на проекцию каждой точки на сетчатку приходится один возбужденный рецептор. Изображение представляется совокупностью конкретных возбужденных рецепторов. Положение каждого такого рецептора известно. Его можно задавать по-разному. Будем полагать, что на сетчатке введена декартова система координат и положение каждого рецептора определяется его координатами. Отметим, что можно было бы взять и другие системы координат – суть дела это меняет незначительно.

Пусть дано изображение A . Перенумеруем его точки некоторым образом так, чтобы номера были попарно различны. Кодом K^1 изображения A (обозначение: K_A^1) назовем пару множеств $\langle M_A, T_A \rangle$. Здесь M_A – множество номеров точек изображения A , T_A – множество координат $(x, y)_n$ точек с указанием

их номера, то есть, например, $(5, 3)_n$ означает, что у точки с номером n координатами являются пара $(5, 3)$.

Перенумеровав по иному множество точек изображения A , получим другую пару $\langle M_A, T_A \rangle$. Будем, однако, рассматривать кодировку с точностью до перенумерации и называть изображения, отличающиеся только нумерацией точек, эквивалентными (в смысле случая 1). Таким образом, если имеются изображения A с кодом $\langle M_A, T_A \rangle$ и B с кодом $\langle M_B, T_B \rangle$, то назовем A и B эквивалентными (в смысле случая 1), если существует такая нумерация точек изображения A , при которой его код есть $\langle \tilde{M}_A, \tilde{T}_A \rangle$, и такая нумерация для B , при которой его код есть $\langle \tilde{M}_B, \tilde{T}_B \rangle$, и при этом $\tilde{M}_A = \tilde{M}_B$ и $\tilde{T}_A = \tilde{T}_B$.

Случай 2. Пусть изображение перед глазом некоторым образом смещается и поворачивается. Пусть при этом плоскости изображения и сетчатки остаются параллельными (или, если они не параллельны, то угол между этими плоскостями остается неизменным) и расстояние между плоскостями не меняется. Очевидно, что этот случай можно трактовать и так, что изображение перед глазом неподвижно, а смещается и поворачивается сам глаз. Преобразования изображения на сетчатке в этом случае сводятся, очевидно, к тому, что оно смещается и поворачивается — других изменений нет. Ясно, что при этом меняется и координатный код изображения, поскольку меняются и координаты составляющих изображения точек.

При сдвиге и повороте меняется положение изображения на сетчатке, другими словами, положение изображения по отношению к осям системы координат. Взаиморасположение же точек изображения не меняется, что означает сохранение расстояний между ними. Именно сохранение расстояний между точками изображения при сдвиге и повороте положим в основу нового кода K_A^2 . В коде $\langle M_A, T_A \rangle$, как и ранее, M_A есть множество попарно отличающихся, а в остальном произвольных, номеров точек изображения. Множество T_A составляют все числа $r(t, n)$ (t и n — номера точек из M_A), являющиеся расстояниями между точками изображения, с указанием того, расстоянием между какими точками является данное число. Аналогично тому, как это сделано выше, определим одинаковые с точностью до перенумерации точек коды. Изображения с такими кодами бу-

дем называть эквивалентными (в смысле случая 2). Можно показать, что два изображения эквивалентны тогда и только тогда, когда одно может быть получено из другого комбинацией сдвига, поворота и преобразования симметрии. Это можно рассматривать как частный случай утверждения, доказанного в [3].

Рассмотрим теперь, как соотносятся коды K_A^1 и K_A^2 и чем объясняется различие их свойств. Задавая изображение совокупностью координат всех его точек (код K_A^1), мы, в сущности, задаем нечто большее, чем собственно изображение – неявным образом в таком задании присутствует и внешняя по отношению к изображению система отсчета, то есть система координат с ее осями. Эта система отсчета в реальности должна быть «привязана» к каким-то дополнительным внешним точкам, например, к краям сетчатки. Действительно, посмотрим на код K_A^1 как на некий вариант кода K_A^2 , то есть будем считать, что каждая точка в K_A^1 тоже задается совокупностью расстояний до других точек. В таком случае этими другими точками для произвольной точки a изображения будут точки x_a и y_a на осях координат, являющиеся проекциями точки a на оси. Код K_A^1 как бы предполагает наличие, помимо точек собственно изображения, еще и точек x_a и y_a для каждой точки a . Если присоединить эти точки к изображению, то, очевидно, код K_A^2 уже не будет инвариантным к сдвигу и повороту первоначального изображения, поскольку при этом меняются расстояния от точек первоначального изображения до добавленных точек.

Основываясь на коде K_A^2 , нетрудно осуществить распознавание сдвинутых, повернутых и симметрично преобразованных изображений. Пусть изображения A и B заданы координатами своих точек, то есть заданы кодами K_A^1 и K_B^1 . Мы предполагаем, что B – это сдвинутое, повернутое и симметрично преобразованное изображение A и намереваемся проверить это предположение. Определим коды K_A^2 и K_B^2 – это будут соответственно $\langle M_A, T_A \rangle$ и $\langle M_B, T_B \rangle$. Коды K_A^2 и K_B^2 очевидным образом строятся по кодам K_A^1 и K_B^1 . Если изображения A и B эквивалентны, то коды K_A^2 и K_B^2 должны быть одинаковы с точностью до перенумерации точек. Остается только проверить это для кодов K_A^2 и K_B^2 . Эта процедура может быть определена по-разному

и основывается на очевидных свойствах кодов эквивалентных изображений. Так, мощности множеств T_A и T_B равны, если изображения A и B эквивалентны. Выделим, далее, в T_A подмножества, состоящие из равных по величине элементов и расположим эти подмножества в порядке возрастания этой величины, то есть $T_A = \{t_1^A, \dots, t_m^A\}$. Очевидно, что $t_i^A \cap t_j^A = \emptyset$ для $i, j = 1, \dots, m$ и $i \neq j$, $\bigcup_{i=1}^m t_i^A = T_A$, $\sum_{i=1}^m |t_i^A| = |T_A|$. Здесь $|t_i^A|$ и $|T_A|$ – мощности соответствующих множеств. Такое же представление в виде подмножеств и с такими же свойствами должно иметь место и для изображения B , то есть $T_B = \{t_1^B, \dots, t_m^B\}$. Ясно, что при этом $|t_i^A| = |t_i^B|$ для всех $i = 1, \dots, m$.

Очевидным образом возникающая «привязка» множеств t_i^A и t_i^B друг к другу дает возможность определить соответствие между тройкой точек в A и тройкой точек в B – соответствие между остальными точками определяется этим однозначно.

Сложность кода можно охарактеризовать мощностью множества T_A . Тогда для изображения A , состоящего из n точек, сложность K_A^1 равна n , сложность K_A^2 равна $n(n-1)/2$.

Вместе с тем код K_A^2 явно избыточен. Это видно, например, из того, что если зафиксировать на плоскости положение трех точек изображения, то для определения положения остальных точек все элементы множества T_A , очевидно, не потребуются.

Случай 3. Пусть изображение перед глазом приближается или удаляется с сохранением параллельными плоскости изображения и плоскости сетчатки (или, если они не параллельны, с сохранением неизменным угла между ними). Это же можно трактовать и так, что изображение перед глазом неподвижно, а приближается или удаляется сам глаз. В обоих этих случаях изображение на сетчатке увеличивается или уменьшается в размерах с сохранением подобия. Можно отметить, что такие же изменения на сетчатке возникнут и в том случае, когда изображение не приближается и не удаляется, а увеличивается или уменьшается (с сохранением подобия). Если не располагать дополнительной информацией, то отличить эти два вида преобразования изображения перед глазом – изменение в размерах и изменение расстояния до глаза – нельзя, что и наблюдается в зрительных иллюзиях.

Коды K_A^2 и K_B^2 изображения A при его приближении или

удалении от глаза очевидно меняются, поскольку меняются и координаты точек его проекции на сетчатку, и расстояния между точками этой проекции. Соотносительные же размеры частей изображения при увеличении или уменьшении с сохранением подобия не меняются. Это и положим в основу кода K_A^3 . Так же, как и в предыдущих случаях, перенумеруем множество точек изображения A и обозначим множество номеров через M_A . Далее зададим множество T_A численных значений отношений вида $r(m, n)/r(p, q)$, где $r(m, n)$ и $r(p, q)$ – расстояния между точками с номерами соответственно m и n , p и q . Здесь m, n, p, q – номера из M_A , $m \neq n$, $p \neq q$. Для каждого числа из T_A полагаем известной соответствующую ему четверку номеров m, n, p, q . Код K_A^3 есть пара $\langle M_A, T_A \rangle$. Аналогично тому, как это рассматривалось для предыдущих двух случаев, можно определить одинаковые с точностью до перенумерации точек коды. Изображения с такими кодами будем называть эквивалентными (в смысле случая 3). Можно показать, что два изображения эквивалентны в том и только в том случае, если на плоскости одно может быть получено из другого сдвигом, поворотом, изменением в размерах (с сохранением подобия), преобразованием симметрии или их комбинацией (то есть подобными преобразованиями). Соответствующее утверждение доказано в [6].

Рассмотрим теперь, как соотносятся коды K_A^2 и K_A^3 и чем объясняется различие их свойств. По коду K_A^2 можно, очевидно, построить код K_A^3 . Обратное неверно. Код K_A^2 определяется таким образом, чтобы в описании изображения не участвовала бы «внешняя» по отношению к изображению система координат. Однако нечто от этой внешней системы в определении кода K_A^2 осталось, а именно – единица измерения расстояний между точками изображения. Она, эта единица, предполагается для кода K_A^2 априори заданной. Когда же мы берем отношение $r(m, n)/r(p, q)$ (в коде K_A^3), эта единица измерения устраняется. Действительно, величина отношения $r(m, n)/r(p, q)$ будет одной и той же вне зависимости от того, в каких единицах измеряются расстояния $r(m, n)$ и $r(p, q)$. Отсюда и возникает возможность посредством K_A^3 описывать изображение безотносительно к его размерам.

Можно и несколько по иному интерпретировать код K_A^3 . Возьмем в качестве единицы измерения расстояние между какой-либо парой точек на самом изображении. Пусть это будут точки с номерами p и q . Теперь для того, чтобы для произвольных точек m и n изображения получить расстояние между ними, выраженное в единицах, являющихся расстоянием между точками p и q , нужно, очевидно, проделать следующее. Взять произвольную единицу измерения и с ее помощью получить расстояния $r(m, n)$ и $r(p, q)$. Затем разделить $r(m, n)$ на $r(p, q)$, то есть получить $r(m, n)/r(p, q)$. Это и будет искомым числом. Если теперь поочередно считать единицей измерения расстояние между каждой парой точек в изображении, то мы и придем ко множеству T_A для кода K_A^3 .

При восстановлении по коду достаточно задать конкретной величину расстояния между любой парой точек изображения. Тогда по элементам множества T_A можно получить значения расстояний для всех других пар точек (то есть, в сущности, получить код K_A^2), и затем построить изображение.

Процедуру распознавания изображений, полученных преобразованиями подобия, можно провести аналогично процедуре, описанной для случая 2.

Множество T_A имеет $(n(n-1)/2)(n(n-1)/2 - 1)$ элементов. Вместе с тем код K_A^3 , очевидно, избыточен.

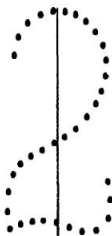


Рис. 3.

Случай 4 (основной). Проведем в плоскости изображения перед глазом прямую (пересекающую изображение). Повернем изображение вокруг прямой на некоторый угол. Изображение на сетчатке сожмется в направлении, перпендикулярном пря-

мой (пример на рис. 3). Этот случай можно трактовать и так, что изображение перед глазом неподвижно, а поворачивается на некоторый угол плоскость сетчатки. Наконец, изображение перед глазом может действительно сжиматься по некоторому направлению. Во всех этих случаях преобразование изображения на сетчатке сведется к сжатию.

Если в исходном положении плоскость изображения непараллельна плоскости сетчатки и затем поворачивается, делаясь параллельной, то, очевидно, изображение на сетчатке при этом будет растягиваться в направлении, перпендикулярном оси поворота.

Еще раз подчеркнем, что только к сжатию или к растяжению преобразование изображения на сетчатке сводится в том случае, когда несущественна разница в расстояниях до глаза от разных частей изображения, возникающая после поворота вокруг оси. В целом эта особенность возникает как часть более широкой проблемы, отмеченной в исследованиях по машинному зрению (см., например, [7,8]). Проекции тела на сетчатку рассматриваются обычно как параллельные проекции (с добавлением возможности преобразований подобия). Глазом же, можно считать с большим основанием, осуществляется центральная проекция. Если, однако, расстояние до объекта велико по сравнению с размерами самого объекта, то различия между центральной и параллельной проекциями невелики и ими можно пренебречь. Мы считаем, что находимся в границах применимости именно этого случая.

Коды K_A^1 , K_A^2 , K_A^3 изображения при сжатии и растяжении, очевидно, меняются. Так, например, код K_A^3 меняется потому, что меняются соотносительные размеры частей изображения (не сохраняется подобие).

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_{kps} — площади треугольников с вершинами в тройках точек с номерами m, n, u и k, p, s и пусть $\rho_{mnu,kps} = S_{mnu}/S_{kps}$. Полагаем, что порядок номеров в тройках не важен, сами тройки различны и при $S_{kps} = 0$ значение $\rho_{mnu,kps}$ не определено. Множе-

ство индексированных чисел $\rho_{tnu,kps}$ 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 в смысле случая 4 (далее просто эквивалентными), если существует такая биекция $\psi : M_A \rightarrow M_B$, что для любых m, n, u и k, p, s из M_A выполнено $\rho_{tnu,kps} = \rho_{\psi(m)\psi(n)\psi(u),\psi(k)\psi(p)\psi(s)}$. Ясно, что эквивалентность изображений содержательно означает одинаковость их кодов с точностью до перенумерации точек. Два изображения называем аффинно эквивалентными (а-эквивалентными), если они переводимы друг в друга аффинными преобразованиями.

Лемма 1. *Если два изображения а-эквивалентны, то они эквивалентны.*

Доказательство. Нужно показать, в сущности, что при аффинных преобразованиях численные значения элементов из множества T кода изображения не меняются. Каждый элемент в T определяется отношением площадей некоторых треугольников. При преобразованиях сдвига, поворота и симметрии эти площади не меняются, а значит не меняется и их отношение. Пусть теперь происходит сжатие (или растяжение) изображения по осям x и y с коэффициентами k_x и k_y . Известно [2], что в этом случае площадь S' фигуры (или любой ее части) будет после преобразования равна $S k_x k_y$ (здесь S – площадь фигуры до преобразования). Отсюда следует, что если $\rho'_{tuv,lpq}$ – соответствующий $\rho_{tuv,lpq}$ элемент множества T после преобразования, то $S_{tuv} k_x k_y / S_{lpq} k_x k_y = S_{tuv} / S_{lpq} = \rho_{tuv,lpq}$. Лемма доказана.

Пусть изображение B состоит из точек, являющихся подмножеством точек изображения A с кодом $\langle M_A, T_A \rangle$. Обозначим через M_B множество номеров тех и только тех точек из A , которые входят в изображение B . Соответственно через T_B обозначим множество всех тех элементов $\rho'_{tuv,lpq}$ из T_A , для которых $(m, u, v, l, p, q) \in M_B$. Изображение B с кодом $\langle M_B, T_B \rangle$ назовем частью изображения A . Любое изображение B' , а-эквивалентное изображению B , назовем подизображением изображения A .

му можно определить длины высот h_{ab} , h_{ac} и h_{bc} , опущенных из точки x на стороны соответственно ab , ac и bc . (Одна из этих высот может оказаться нулевой по длине, две быть нулевыми, очевидно, не могут.) Таким образом, задача сводится к тому, чтобы по заданному треугольнику abc построить точку x такую, что высоты из этой точки на стороны ab , ac и bc треугольника есть соответственно h_{ab} , h_{ac} и h_{bc} . Проводим две прямые, параллельные стороне ab по разные стороны и отстоящие от нее на величину h_{ab} , и две прямые, параллельные стороне ac по разные ее стороны и отстоящие от нее на величину h_{ac} . Возникает параллелограмм, вершины которого обозначены на рис. 4 как Q_1 , Q_2 , Q_3 и Q_4 . По построению, точка x может располагаться только на вершинах этого параллелограмма. Для окончательного определения положения точки x нужно провести две прямые, параллельные оставшейся стороне треугольника abc (стороне bc на рис. 4) и отстоящие от нее на величину h_{bc} . Пересечение этих прямых с вершинами параллелограмма даст возможные положения точки x . Одно такое пересечение, очевидно, гарантированно существует. Вопрос лишь в том, единственным ли может получиться положение восстанавливаемой точки.

Изображение назовем плоским, если все его точки не лежат на одной или двух параллельных прямых.

Лемма 2. *Если по коду плоского изображения выбраны какие-либо три его точки, не лежащие на одной прямой, и задано произвольное (но не на одной прямой) положение этих точек на плоскости, то положение остальных точек изображения определяется по коду однозначно.*

Доказательство. Предположим, что задано положение точек a , b и c и есть отличная от них точка x плоского изображения A , для которой возможны два разных положения x' и x'' . Используя стандартную процедуру и обозначения рис. 4, строим параллелограмм, вершинами которого являются возможные положения точки x .

1. Предположим, что точки x' и x'' совпадают с Q_1 и Q_2 , которые поэтому должны быть равноудалены от стороны bc . Следовательно, отрезок bc должен делить отрезок Q_1Q_2 пополам, то есть точка b должна лежать на отрезке Q_1Q_2 . В этом

случае a, b, c и x (безотносительно к тому, совпадает x с x' или x'') расположены на двух параллельных отрезках: отрезке ac и отрезке Q_1Q_2 . Так как A – плоское изображение, то должна существовать точка – обозначим ее через d – вне этих отрезков и их продолжений. Но тогда точки x' и x'' не равноудалены по крайней мере от одного из отрезков da и dc – пришли к противоречию.

2. Аналогичным образом приходим к противоречию в предположении, что x' и x'' совпадают с Q_1 и Q_4 .

Если x' и x'' совпадают с Q_1 и Q_3 , то отрезок Q_1Q_3 должен делиться стороной bc пополам. Однако серединой диагонали Q_1Q_3 параллелограмма является, по построению, точка a .

Если x' и x'' совпадают с Q_2 и Q_3 , то отрезок Q_2Q_3 должен быть либо параллелен отрезку bc , либо делиться этим отрезком или его продолжением пополам. И то, и другое противоречиво по построению. Аналогичное имеет место и при совпадении x' и x'' с Q_3 и Q_4 .

3. Пусть, наконец, x' и x'' совпадают с Q_2 и Q_4 . Тогда диагональ Q_2Q_4 должна быть параллельной стороне bc . Далее рассуждения сводятся к началу пункта 1 настоящего доказательства. Роли отрезков Q_2Q_4 , ac , bc и ab играют соответственно Q_1Q_2 , ab , ac и bc . Лемма доказана.

Теорема 1. *Два плоских изображения эквивалентны точно тогда, когда они а-эквивалентны.*

Доказательство. В одну сторону теорема 1 следует из леммы 1.

Покажем теперь, что если плоские изображения A и B эквивалентны, то они а-эквивалентны. Выберем на A три точки a, b, c , не лежащие на одной прямой. Пусть следующей из определения эквивалентности биекцией точкам a, b, c сопоставляются точки соответственно a', b', c' изображения B . Аффинным преобразованием изображения B совместим его точки a', b', c' с точками соответственно a, b, c изображения A . Поскольку положение остальных точек изображения B , при заданных точках a', b', c' определяется, согласно лемме 2, однозначно, то, сле-

довательно, они совпадут с точками изображения A . Теорема доказана.

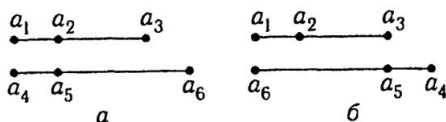


Рис. 5.

Изображения не плоские могут быть эквивалентными, но не a -эквивалентными. Пусть, для примера, изображение U_1 состоит из точек a_1, a_2, a_3 и a_4, a_5, a_6 , расположенных на двух параллельных прямых (рис. 5(а)). Положим, что $r(a_2, a_3) = 2r(a_1, a_2)$, $r(a_5, a_6) = 3r(a_4, a_5)$ (здесь $r(a_i, a_j)$ – расстояние между соответствующими точками). Пусть изображение U_2 (рис. 5(б)) отличается от U_1 только тем, что точки a_4, a_5 и a_6 на нижней прямой «перевернуты» и идут в обратном порядке. Изображения U_1 и U_2 эквивалентны, причем точкам из U_1 ставятся в соответствие точки с теми же номерами из U_2 . Однако их нельзя перевести друг в друга аффинными преобразованиями. На изображении U_1 отрезки прямых, соединяющих a_1 с a_6 и a_3 с a_4 имеют точку пересечения в промежутке между параллельными прямыми, а на изображении U_2 такой точки нет, и при всех аффинных преобразованиях это различие не устранимо.

Пусть плоские изображения A и B заданы своими кодами $\langle M_A, T_A \rangle$ и $\langle M_B, T_B \rangle$. Тривиальный способ проверки их на эквивалентность состоял бы в том, чтобы всеми возможными вариантами однозначно сопоставлять элементы множеств M_A и M_B и проверять при этом, как это следует из определения эквивалентности, равенство соответствующих элементов из множеств T_A и T_B . Если каждое из изображений состоит из n точек, то проверять придется $n!$ вариантов сопоставления.

Процедура проверки может быть упрощена. Построим изображение по его коду. Выберем на A три точки a, b и c , не лежащие на одной прямой. Затем будем поочередно брать всевозможные тройки точек a', b' и c' на B (не лежащие на одной прямой) и считать их совмещенными с точками a, b и c . Если A и B эквивалентны, ψ – биекция из определения эквивалент-

ности, и $a' = \psi(a)$, $b' = \psi(b)$, $c' = \psi(c)$, то при построении по коду и все остальные точки изображения B совместятся с соответствующими точками изображения A . Это и определит полностью биекцию ψ . Число вариантов для проверки при такой процедуре, очевидно, не превышает n^3 .

Можно еще больше упростить процедуру проверки. Назовем точку x изображения A внутренней, если найдутся такие точки a, b и c из A , отличные от x , что образованный ими треугольник содержит x . Пусть V_A – множество всех таких точек из A . Очевидно, что по коду изображения для каждой его точки можно проверить включение $x \in V_A$. Точки из $A \setminus V_A$ назовем внешними или контурными, а часть изображения A , образованную этими точками – контуром. Ясно, что биекция ψ должна порождать взаимно однозначное соответствие точек контуров изображений A и B .

Затем на изображении, образованном точками из V_A , можно выделить его контур, и т.д. В результате изображение однозначно «расслоится» на контуры: первый, второй и т.д. Далее для изображений A и B нужно сопоставлять друг другу только точки из контуров с одинаковыми номерами.

Пусть $i_1, \dots, i_k$ – номера k точек изображения A . Рассмотрим выпуклый многоугольник, включающий все эти точки, часть из которых (и только они) образуют 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 $\rho_{i_1, \dots, i_k, j_1, \dots, j_k} = \rho_{\psi(i_1), \dots, \psi(i_k), \psi(j_1), \dots, \psi(j_k)}$.

Покажем на примере (рис. 6), что продолжить теорему 1 на случай $k \geq 4$ для k -эквивалентных изображений нельзя. Четы-

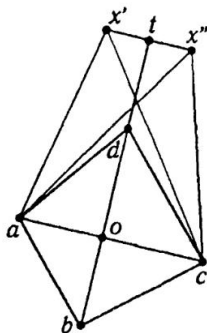


Рис. 6.

рехугольник $abcd$ – выпуклый, с попарно неравными друг другу по площади треугольниками abd , bcd и abc . Площади треугольников abc и adc могут оказаться равными. На прямой, являющейся продолжением диагонали bd симметрично относительно этой прямой откладываем точки x' и x'' так, чтобы отрезок $x'x''$ был параллелен диагонали ac и треугольники $ax's$ и $ax''c$ содержали бы вершину d как внутреннюю точку. Рассмотрим изображение A_1 из точек a, b, c, d и x' и изображение A_2 – из точек a, b, c, d , и x'' . Эти изображения 4-эквивалентны, если площади соответствующих четырехточечников на них равны. Действительно, $abcd$ у A_1 и A_2 – общий, четырехточечники $acdx'$ и $acdx''$ (со внутренней точкой d) имеют одинаковую площадь по построению. Далее имеем $S_{bcdx'} = S_{bcx'x''} - S_{cx'x''}$ и $S_{bcdx''} = S_{bcx'x''} - S_{bt'x'} - S_{dt'x''}$. Отсюда для равенства $S_{bcdx'} = S_{bcdx''}$ нужно, чтобы имело место $S_{cx'x''} = S_{bt'x'} + S_{dt'x''}$. У треугольника $cx'x''$ основание $x'x''$ в два раза больше оснований $x't$ и $x''t$ у треугольников $bt'x'$ и $dt'x''$. Высоты же на это основание у треугольника $cx'x''$ больше, чем у треугольника $dt'x''$, и меньше, чем у треугольника $dt'x'$. Из этого следует, что сохраняя положение прямой, проходящей через точки b, o, d и t и меняя на ней положение точки b , можно подобрать такое ее положение, при котором $S_{bt'x'} = S_{cx'x''} - S_{dt'x''}$. При этом положении точки b четырехточечник $abcd$ останется выпуклым. При найденном положении точки b , нетрудно проверить, будет иметь

место и $S_{bcdx'} = S_{bcdx''}$. Итак, A_1 и A_2 4-эквивалентны, но, очевидно, не переводятся одно в другое аффинными преобразованиями. Продолжение этого примера на случай изображений A_1 и A_2 из любого числа n ($n \geq 5$) точек, и, соответственно, их k -эквивалентности для $k = n - 1$ достигается добавлением нужного числа точек (то есть $n - 5$ точек) на произвольные места на отрезке od на рис 6. 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 . В [6] показано, что изображения A и B 2-эквивалентны тогда и только тогда, когда они переводятся одно в другое преобразованиями подобия. Суммируя все сказанное, приходим к следующему.

Теорема 2. *Для двух плоских изображений условие совпадения их k -эквивалентности ($k \geq 2$) и аффинной эквивалентности справедливо только при $k = 3$.*

§2. О кодировании и распознавании объемных изображений

Назовем трехмерным изображением A или телом конечное множество точек в трехмерном евклидовом пространстве. Занумеруем попарно различными номерами точки изображения A . Пусть M_A множество этих номеров, V_{tnuv} и V_{kspq} — объемы тетраэдров с вершинами в четверках точек с номерами t, n, u, v и k, s, p, q и пусть $\rho_{tnuv, kspq} = V_{tnuv}/V_{kspq}$. Полагаем, что порядок номеров в четверках не важен, сами четверки различны и для случая, когда $V_{kspq} = 0$, $\rho_{tnuv, kspq}$ не определено. Множество индексированных чисел $\rho_{tnuv, kspq}$ для всех таких пар четверок обозначим через T_A . Кодом тела назовем пару $\langle M_A, T_A \rangle$. Тела 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 t, n, u, v и k, s, p, q из M_A выполнено $\rho_{tnuv, kspq} = \rho_{\psi(t)\psi(n)\psi(u)\psi(v), \psi(k)\psi(s)\psi(p)\psi(q)}$.

Тела, все точки которых расположены в одной плоскости, называем двумерными и для них рассматриваемый код не определен. Тела называем аффинно эквивалентными (а-

эквивалентными), если они переводятся друг в друга аффинными преобразованиями.

Лемма 3. *Если два тела α -эквивалентны, то они эквивалентны.*

Доказательство леммы 3 аналогично доказательству леммы 1 и практически его повторяет.

Введенные ранее определения части изображения и подизображения очевидным образом могут быть распространены и на трехмерный случай.

Изображение, состоящее из пяти точек, не лежащих в одной плоскости, будем называть **пятиточечником**. Пусть в пятиточечнике A , состоящем из точек a, b, c, d, e точки a, b, c и e расположены в одной плоскости и точка d вне этой плоскости. Тогда $V_{abce} = 0$ и те отношения из T_A , у которых V_{abce} входит в знаменатель, не определены, а те, у которых V_{abce} входит в числитель (при не равном нулю знаменателе), равны нулю. Очевидно, можно утверждать и обратное: если код $\langle M_A, T_A \rangle$ обладает такими свойствами, то в пятиточечнике A точки a, b, c и e расположены в одной плоскости. В изображении A общего вида точки a, b, c и e , входящие в изображение, расположены в одной плоскости в том и только в том случае, если для любой части изображения A , являющейся пятиточечником и включающей точки a, b, c, e , выполняется указанное выше свойство, определяющее расположенность точек в одной плоскости. В изображении A точки $a_1, \dots, a_n$, ($n \geq 4$) расположены в одной плоскости в том и только в том случае, когда любые четыре точки из $a_1, \dots, a_n$ расположены в одной плоскости. Аналогичным и очевидным образом можно указать необходимые и достаточные свойства кода изображения, при которых три и большее число точек лежат на одной прямой.

Далее будем основываться на следующей процедуре (стандартной) построения по коду пятиточечника положения пятой точки при известных положениях остальных четырех. Обозначим точки с известным положением через a, b, c и e . Полагаем, что эти точки разные, лежат не в одной плоскости и, следовательно, тетраэдр $abce$ имеет ненулевой объем, в остальном же их положение в пространстве можно задавать произвольным

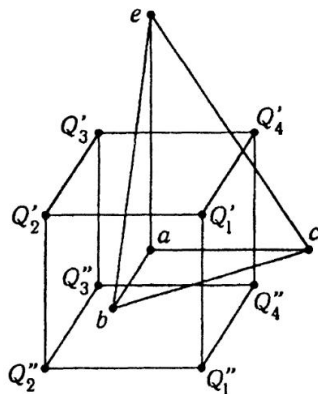


Рис. 7.

образом. Точку, положение которой определяется, обозначим через x . Поскольку объем тетраэдра $abce$ известен, то из соотношений V_{abcx}/V_{abce} , V_{abex}/V_{abce} , V_{acex}/V_{abce} , V_{bcex}/V_{abce} (все это элементы множества T_A), находим объемы V_{abcx} , V_{abex} , V_{acex} , V_{bcex} . Так как площади всех граней тетраэдра $abce$ известны, то можно определить значения длин высот h_{abc} , h_{abe} , h_{ace} и h_{bce} из точки x на грани соответственно abc , abe , ace , bce (две из этих высот могут оказаться равными нулю по длине, три быть нулевыми, очевидно, не могут). Таким образом, задача сводится к тому, чтобы при заданном тетраэдре построить точку x такую, что высоты из этой точки на соответствующие грани тетраэдра есть h_{abc} , h_{abe} , h_{ace} и h_{bce} . При этом известно, что одно решение у этой задачи заведомо существует. Вопрос лишь в том, при каких условиях такая точка единственна. Для построения возможных положений точки x проведем для каждой из граней abc , abe , ace , bce по две плоскости, параллельные грани и по разные ее стороны, и отстоящие от граней на величину высоты из точки x на эту грань. Возникает параллелепипед, вершины которого обозначены на рис. 7 как Q_1' , Q_2' , Q_3' , Q_4' , Q_1'' , Q_2'' , Q_3'' и Q_4'' . По построению, точка x может располагаться только на вершинах параллелепипеда Q . Для окончательного определения положения точки x нужно провести две плоскости,

параллельные оставшейся грани тетраэдра (грани bce на рис. 7) по разные ее стороны и отстоящие от нее на величину h_{bce} . Пересечение этих плоскостей с вершинами параллелепипеда и даст возможные положения точки x .

Трехмерное изображение назовем объемным, если все его точки не лежат в одной плоскости или в двух параллельных плоскостях.

Лемма 4. Если по коду объемного изображения выбраны какие-либо четыре его точки, не лежащие в одной плоскости, и задано произвольное (но не в одной плоскости) положение этих точек в трехмерном пространстве, то положение остальных точек изображения определяется по коду однозначно.

Доказательство. Пусть задано положение точек a, b, c и e объемного изображения A . Предположим, что для некоторой точки x , отличной от a, b, c и e , возможны два разных положения x' и x'' . Используя стандартную процедуру и обозначения рис. 7, построим параллелепипед Q , вершины которого являются возможными положениями точки x .

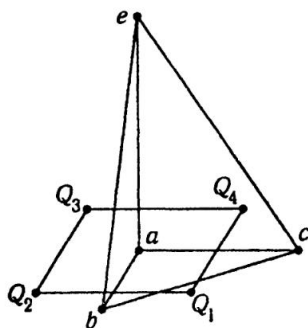


Рис. 8.

1. Предположим, что точка x лежит в плоскости одной из граней тетраэдра $abce$. Пусть, для определенности, это будет грань abc . Параллелепипед Q в этом случае превращается в параллелограмм $Q_1Q_2Q_3Q_4$ (рис. 8).

1а). Положим, что точки x' и x'' совпадают с вершинами Q_1 и Q_2 . Поскольку x' и x'' должны быть равноудалены от грани bce , то отрезок bc должен делить отрезок Q_1Q_2 пополам и, значит, точка b должна лежать на отрезке Q_1Q_2 .

Изображение A – объемное, следовательно, должны существовать точки вне плоскости грани abc и отличные от точки e . Если такая точка d лежит вне плоскости грани ace и вне ребра be и его продолжений, то точки x' и x'' не равноудалены хотя бы от одной из плоскостей граней aed и ced – пришли к противоречию.

Пусть теперь все точки изображения A вне плоскости грани abc и отличные от e лежат либо на ребре be или его продолжениях, либо в плоскости грани ace . В первом случае все точки изображения A можно представить расположенными в плоскости грани bex и параллельной ей плоскости, проходящей через отрезок ac . Во втором – в плоскости грани ace и параллельной ей плоскости, проходящей через отрезок bx . Значит должны существовать одновременно и хотя бы одна точка, обозначим ее через d_1 , на ребре be или его продолжениях, и хотя бы одна точка, обозначим ее через d_2 , в плоскости грани ace (вне отрезка ac и его продолжений). Но тогда точки x' и x'' не равноудалены от плоскости по крайней мере одной из граней d_1d_2c и d_1d_2a .

1б). Аналогичным образом приходим к противоречию, предположив, что точки x' и x'' совпадают с вершинами Q_1 и Q_4 .

На каждой из пар вершин Q_2 и Q_3 , Q_4 и Q_3 точки x' и x'' не могут располагаться в силу того обстоятельства, что они не равноудалены от грани bce .

Наконец, если предположить, что x' и x'' совпадают с вершинами Q_2 и Q_4 , то диагональ Q_2Q_4 параллелограмма должна быть параллельной отрезку bc (только в этом случае точки x' и x'' равноудалены от грани bce). Далее очевидным образом могут быть повторены рассуждения пункта 1а. При этом роль отрезков Q_2Q_4 и bc играют отрезки соответственно Q_1Q_2 и ac , роль граней bce , ace и abe – грани соответственно ace , abe и bce .

Итак, показано, что точка x не может находиться в плоскости грани abc тетраэдра. Далее в доказательстве леммы везде

предполагается, что x находится вне плоскостей граней тетраэдра $abce$.

2. Предположим, что сечение тетраэдра $abce$ или продолжений его ребер верхней гранью параллелепипеда Q есть треугольник, то есть вершина e тетраэдра не лежит на верхней грани.

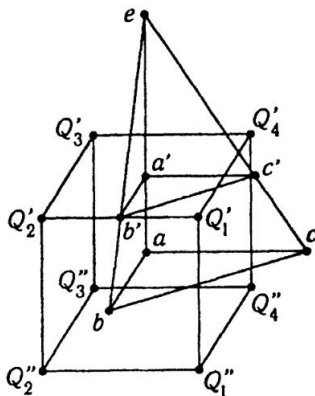


Рис. 9.

2а). Предположим, что точки x' и x'' совпадают с вершинами Q'_1 и Q'_2 (рис. 9). Поскольку x' и x'' должны быть равноудалены от грани bce , то отрезок $b'c'$ должен делить отрезок $Q'_1Q'_2$ пополам, то есть точка b' должна лежать на отрезке $Q'_1Q'_2$. Отсюда следует, что отрезок ac параллелен грани bex (вне зависимости от того, занимает точка x положение x' или x''). Следовательно, точки a, b, c, e и x располагаются в двух параллельных плоскостях: первая из них, обозначим ее через P_1 – плоскость грани bex , вторая, P_2 – ей параллельная и проходящая через отрезок ac .

Поскольку A – объемное изображение, то должны существовать точки вне плоскостей P_1 и P_2 . Здесь возможны два случая. В первом из них существует точка, обозначим ее через d , лежащая вне плоскости грани ase и вне ребра be и его продолжений. Тогда точки x' и x'' не равноудалены от плоскостей хотя бы одной из граней aed и ced – пришли к противоречию.

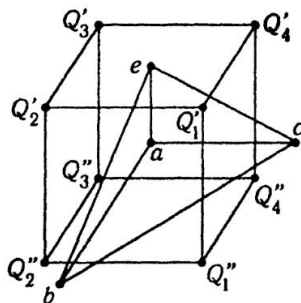


Рис. 10.

Во втором случае все остающиеся точки изображения, отличные от a , b , c , e и x лежат на ребре be и его продолжениях и в плоскости грани ace . Если остающиеся точки расположены только на ребре be и его продолжениях, то это значит, что они расположены в плоскости P_1 . Значит должна существовать точка — обозначим ее через d , лежащая в плоскости грани ace . Если при этом d лежит на отрезке ae или его продолжениях, то она лежит в плоскости P_2 . Если же d — вне ребра ae и его продолжений, то точки x' и x'' не равноудалены от плоскости грани abd (если d не лежит на ребре ae или его продолжениях) и от плоскости грани bcd (если d не лежит на ребре ce или его продолжениях).

26). Аналогичным образом приходим к противоречию, предположив, что точки x' и x'' совпадают с Q'_1 и Q'_4 .

С каждой из пар вершин, Q'_2 и Q'_3 , Q'_4 и Q'_1 точки x' и x'' не могут совпадать в силу того очевидного обстоятельства, что они не равноудалены от грани bce .

Наконец, если предположить, что x' и x'' совпадают с вершинами Q'_2 и Q'_4 , то диагональ $Q'_2Q'_4$ должна быть параллельной отрезку bc (только в этом случае точки x' и x'' равноудалены от грани bce). Далее очевидным образом могут быть повторены рассуждения пункта 2а. При этом роль отрезков $Q'_2Q'_4$ и bc играют отрезки соответственно $Q'_1Q'_2$ и ac , роль граней bce , ace и abe — грани соответственно ace , abe и bce .

3. Предположим теперь, что вершина e тетраэдра $abce$ ле-

жит на верхней грани параллелепипеда Q (рис. 10). При этом из вершин параллелограмма верхней грани параллелепипеда Q для размещения на них точек x' и x'' годятся, очевидно, только концы диагоналей параллелограмма, то есть пары Q'_1 и Q'_3 , Q'_2 и Q'_4 – при совпадении с ними точек x' и x'' они равноудалены от всех граней тетраэдра $abce$. Поскольку A – объемное изображение, то должна существовать точка – обозначим ее через d – вне плоскостей грани abc и верхней грани параллелепипеда Q . Однако, в этом случае точки в парах Q'_1 и Q'_3 , Q'_2 и Q'_4 не равноудалены от плоскостей по крайней мере одной из граней abd , acd и bcd – пришли к противоречию.

4. Аналогичным пунктам 2 и 3 образом могут быть рассмотрены все параллелограммы поверхности параллелепипеда Q и показана невозможность совпадения точек x' и x'' с парами их вершин. Остается рассмотреть случай, когда x' и x'' совпадают с концами диагоналей параллелепипеда Q . При этом сразу могут быть исключены из рассмотрения диагонали $Q''_1Q'_3$ и $Q'_1Q''_3$, поскольку их концы не равноудалены от грани bce . Если же x' и x'' совпадают с Q''_2 и Q'_4 , или с Q'_2 и Q''_4 , то соответствующая диагональ должна быть параллельной грани bce тетраэдра. Далее рассуждения очевидным образом сводятся к пункту 3, причем роль грани bce тетраэдра играет здесь грань abc , роль вершины a – вершина e .

Итак, показано, что для всех возможных случаев предположение о том, что для точки x возможны два разных положения, приводит к противоречию. Лемма доказана.

Теорема 3. *Два объемных изображения эквивалентны точно тогда, когда они аффинно эквивалентны.*

Доказательство основывается на леммах 3 и 4 и повторяет доказательство теоремы 1.

Покажем, что не объемные изображения могут быть эквивалентными, но не аффинно эквивалентными. Отметим сначала два свойства изображений, у которых точки располагаются на скрещивающихся прямых.

Пусть две точки a и b изображения расположены в плоскости P_1 на прямой L_1 , и две точки c и e – в плоскости P_2 ,

параллельной P_1 , на прямой L_2 . Прямые L_1 и L_2 скрещивающиеся, следовательно, тетраэдр с вершинами в точках a, b, c, e имеет ненулевой объем. Очевидно, что если поменять местами точки a и b , то объем тетраэдра $abce$ сохранится. Нетрудно показать, что если сместить отрезок ab по прямой L_1 (сохраняя длину отрезка ab), то объем тетраэдра $abce$ не изменится.

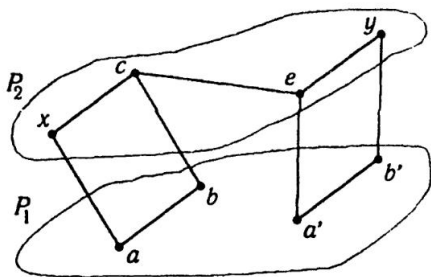


Рис. 11.

Второе свойство связано с возможностью смещать отрезок ab в плоскости P_1 с сохранением его параллельным самому себе, без того, чтобы при этом менялся объем тетраэдра $abce$. Действительно, пусть, как и ранее, в тетраэдре $abce$ ребро ab — в плоскости P_1 , и ребро ce — в плоскости P_2 , параллельной плоскости P_1 . Рассмотрим отрезок $a'b'$ в плоскости P_1 , параллельный отрезку ab и равный ему по длине (рис. 11). Покажем, что $V_{abce} = V_{a'b'ce}$. Построим точки x и y в плоскости P_2 такие, что отрезок ax параллелен отрезку $a'c$, отрезок $b'y$ параллелен отрезку $a'e$. Ясно, что отрезок xc параллелен и равен отрезку ey , следовательно, $xcye$ — параллелограмм, и ce — диагональ в нем. Имеем $V_{abce} = V_{xbce}$, так как отрезок ax параллелен грани bce , и $V_{b'a'ce} = V_{ya'ce}$, так как отрезок $b'y$ параллелен грани $a'ce$. Но $V_{xbce} = V_{ya'ce}$, так как площади S_{xcce} и S_{yccce} равны, и точки b и a' лежат в плоскости P_1 , параллельной плоскости P_2 . Отсюда $V_{abce} = V_{a'b'ce}$.

Пусть теперь задано изображение A , все точки которого располагаются в двух параллельных плоскостях P_1 и P_2 . Ту его часть, которая расположена в плоскости P_1 , обозначим через A_1 , в плоскости P_2 — через A_2 . Все тетраэдры с вершинами в

точках изображения A можно разбить на две группы. В первой из них каждый тетраэдр имеет три точки (назовем их основанием) среди точек одной части (A_1 или A_2), и одну точку (назовем ее вершиной тетраэдра) среди точек другой части. Объем такого тетраэдра определяется произведением площади основания на высоту из вершины на это основание. Эта высота одинакова для всех тетраэдров первой группы. Во второй группе – тетраэдры, две точки которых принадлежат одной части изображения A , и две точки – другой. Такие как A изображения могут быть эквивалентными, но не переводиться одно в другое аффинными преобразованиями. Для доказательства рассмотрим изображение A' , полученное из A преобразованием части A_1 в A'_1 . Это преобразование состоит в преобразовании симметрии точек изображения A_1 относительно какой-либо выбранной в плоскости P_1 точки. При этом преобразовании тетраэдры первой группы не изменяют своих объемов, так как площади оснований и высоты на основания сохраняются. Тетраэдры второй группы тоже сохраняют свои объемы в силу рассмотренных выше двух свойств.

Следовательно, нетрудно видеть, изображения A и A' эквивалентны. Вместе с тем они, в общем случае, не могут быть переведены одно в другое аффинными преобразованиями. Наглядно это представляется, если положить, что A содержит как часть изображение U_1 из примера для плоского случая, а изображение A' – изображение U_2 (рис. 5), причем точки a_1, a_2 и a_3 расположены в плоскости P_1 , точки a_4, a_5, a_6 – в плоскости P_2 .

Как и в двумерном случае, для тел A и B , состоящих из n точек каждое, тривиальный способ проверки их на эквивалентность приводит к $n!$ вариантам сопоставлений точек изображений. Аналогичным плоскому случаю способом, основываясь на лемме 4, число вариантов для проверки можно понизить до n^4 . Можно и дальше понижать эту оценку: ввести понятия внутренних и контурных точек, контура, «расслоить» тело на контуры и сопоставлять для тел A и B только точки соответствующих друг другу контуров.

§3. О восстановлении трехмерного изображения по плоским проекциям

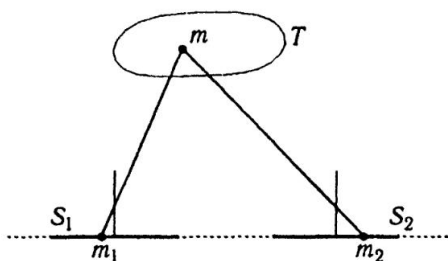


Рис. 12.

Восстановление трехмерного изображения по плоским проекциям служит, с одной стороны, предположительной основой механизмов стереоскопического зрения в живых организмах [9], с другой – является важной задачей в рамках машинного зрения для робототехники [7,8]. В описанных схемах такого восстановления точка m проецируется на две плоские сетчатки (рис. 12), проекции ее есть точки соответственно m_1 и m_2 . Если известно положение этих точек на сетчатках, известно расстояние между сетчатками, 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 поточно все тело. Проекции S_1 и S_2 тела T 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.

Главная проблема [7,8] в рамках машинного стереозрения – это проблема идентификации соответствующих друг другу точек на двух проекциях. На рис. 12 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 расстояния между соответствующими друг другу 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. 12 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. 12 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 рассмотрения 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 $\bar{S}_1$ и $\bar{S}_2$, полученной из соответственно S_1 и S_2 аффинными преобразованиями.

Отметим, что такие построения могут представлять интерес и в связи с проблемами томографии, поскольку в этом слу-

чае тоже происходит восстановление трехмерного тела по проекциям.

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. Взаимно однозначное соответствие между точками двух изображений назовем их разметкой. Соответствующие друг другу точки будем обозначать одной буквой (с разными индексами). Ясно, что описанным выше устанавливается взаимно однозначное соответствие между точками тела T и точками проекций S_i ($i = 1, 2, \dots$). Если a – 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 взаимно однозначное 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: $A \neq B$). 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. Оговорим обозначения для специального случая. Изображение A , состоящее из точек $a_i, b_i, c_i, \dots, u_i$ и прямой L_i , будем обозначать также через $A(a_i, b_i, c_i, \dots, u_i, L_i)$. Аналогично изображение B , состоящее из

точек $a_j, b_j, c_j, \dots, u_j$ и прямой L_j , будем обозначать также через $B(a_j, b_j, c_j, \dots, u_j, L_j)$. Назовем A и B a' -эквивалентными, если их можно перевести одно в другое аффинными преобразованиями, причем так, что точки $a_i, b_i, c_i, \dots, u_i$ совместятся с точками соответственно $a_j, b_j, c_j, \dots, u_j$ и прямая L_i совместится с прямой L_j .

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 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. Проекции из $\{S\}$ можно интерпретировать как изображения тела 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).

Лемма 5. *Если S_1 и S_2 – проекции тела T по двум разным направлениям α_1 и α_2 , $\tilde{S}_1 \approx S_1$, $\tilde{S}_2 \approx S_2$, то $\tilde{S}_1$ и $\tilde{S}_2$ a' -разные тогда и только тогда, когда T – не двумерное тело.*

Доказательство. Положим T – не двумерное тело. Покажем, что $\tilde{S}_1 \neq \tilde{S}_2$.

Рассмотрим в T четыре точки a, b, c и d , не лежащие в одной плоскости. Достаточно, очевидно, показать, что $S_1(a_1, b_1, c_1, d_1) \neq S_2(a_2, b_2, c_2, d_2)$. Пусть плоскость грани abc не параллельна направлениям проекции α_1 и α_2 (если это не

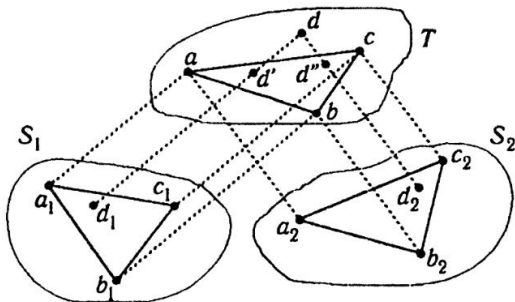


Рис. 13.

так, то возьмем вместо abc какую-либо другую грань в тетраэдре $abcd$). Возьмем в качестве плоскости проекции плоскость грани abc (рис. 13). Тогда по направлению α_1 проекция точек a, b, c, d будет состоять из точек a, b, c, d' , а по направлению α_2 — из точек a, b, c, d'' . Поскольку d' и d'' не совпадают, то $S_1(a_1, b_1, c_1, d_1) \neq S_2(a_2, b_2, c_2, d_2)$.

Пусть теперь $\tilde{S}_1 \neq \tilde{S}_2$. Если предположить, что T — двумерное изображение, то $T \approx \tilde{S}_1$ и $T \approx \tilde{S}_2$ и, значит, $\tilde{S}_1 \approx \tilde{S}_2$ — пришли к противоречию. Лемма доказана.

Как следствие из леммы 5 получаем, что тело двумерное тогда и только тогда, когда любые два изображения, a' -эквивалентные его проекциям по разным направлениям, a' -эквивалентны.

Лемма 5 и следствие из нее применимы, очевидно, и в отношении части тела и соответствующих частей на проекциях. Если, однако, часть тела двумерная, то должна выполняться сделанная выше оговорка — направления проекции не должны быть параллельными плоскости этой части.

Пусть S_1 и S_2 — a' -разные проекции тела T , $\tilde{S}_1 \approx S_1$, $\tilde{S}_2 \approx S_2$, $a_1 b_1 c_1$ на $\tilde{S}_1$ и $a_2 b_2 c_2$ на $\tilde{S}_2$ — треугольники. Будем говорить, что точка d_1 на $\tilde{S}_1$ лежит в плоскости треугольника $a_1 b_1 c_1$, если $\tilde{S}_1(a_1, b_1, c_1, d_1) \approx \tilde{S}_2(a_2, b_2, c_2, d_2)$. В этом случае и d_2 лежит в плоскости треугольника $a_2 b_2 c_2$. В противном случае говорим, что d_1 лежит вне плоскости треугольника $a_1 b_1 c_1$ (d_2 лежит вне плоскости треугольника $a_2 b_2 c_2$).

Рассмотрим на $\tilde{S}_1$ и $\tilde{S}_2$ четверку и троек точек $\langle (a_1 b_1 c_1)(a_2 b_2 c_2)(c_1 d_1 e_1)(c_2 d_2 e_2) \rangle$ такую, что $a_1 b_1 c_1$ и $a_2 b_2 c_2$ – треугольники, точки d_1 и e_1 лежат вне плоскости треугольника $a_1 b_1 c_1$ (и, соответственно, точки d_2 и e_2 лежат вне плоскости треугольника $a_2 b_2 c_2$), и из троек $c_1 d_1 e_1$ и $c_2 d_2 e_2$ 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. Очевидно, что $\tilde{L}_1$ и $\tilde{L}_2$ проходят через точки соответственно c_1 и c_2 . Для определения положения прямых $\tilde{L}_1$ и $\tilde{L}_2$, тем самым, достаточно определить еще по одной точке, лежащих на них.

Лемма 6. Если S_1 и S_2 – a' -разные проекции тела T , $\tilde{S}_1 \approx S_1$, $\tilde{S}_2 \approx S_2$, и u – правильная четверка на $\tilde{S}_1$ и $\tilde{S}_2$, то по $\tilde{S}_1$ и $\tilde{S}_2$ можно определить положение на них прямых $\tilde{L}_1$ и $\tilde{L}_2$.

Доказательство. 1. Положим, что только одна из троек $c_1 d_1 e_1$ и $c_2 d_2 e_2$ – треугольник, и пусть, для определенности, это тройка $c_2 d_2 e_2$. Следовательно, точки c_1, d_1 и e_1 лежат на одной прямой и, значит, направление проекции α_1 параллельно плоскости грани cde . В этом случае точки c_1, d_1 и e_1 лежат на прямой $\tilde{L}_1$, что и определяет ее положение.

Прямая L в теле T лежит в плоскости треугольника abc . Следовательно, $\tilde{S}_1(a_1, b_1, c_1, \tilde{L}_1) \approx \tilde{S}_2(a_2, b_2, c_2, \tilde{L}_2)$. Совместим аффинными преобразованиями изображения $\tilde{S}_1$ точки a_1, b_1 и c_1 с точками соответственно a_2, b_2 и c_2 на $\tilde{S}_2$. Прямая $\tilde{L}_1$ при этом совместится с прямой $\tilde{L}_2$, что и определяет положение $\tilde{L}_2$ на $\tilde{S}_2$.

2. Положим теперь, что тройки $c_1 d_1 e_1$ и $c_2 d_2 e_2$ – треугольники.

Если какая-либо из точек a_1 и b_1 лежит в плоскости треугольника $c_1 d_1 e_1$, то прямая $\tilde{L}_1$ проходит через эту точку и точку c_1 . Положение $\tilde{L}_2$ тоже определено тем, что она проходит через соответствующие точки на $\tilde{S}_2$.

Пусть теперь точки a_1 и b_1 лежат вне плоскости треугольника $c_1d_1e_1$.

Если в теле T отрезок de параллелен плоскости грани abc и отрезок ab параллелен плоскости грани cde , то оба отрезка de и ab должны быть параллельны прямой L и, следовательно, параллельны между собой. В этом случае отрезки a_1b_1 и d_1e_1 на $\tilde{S}_1$, и a_2b_2 и d_2e_2 на $\tilde{S}_2$ параллельны между собой. Их направлениями и определяются прямые $\tilde{L}_1$ и $\tilde{L}_2$.

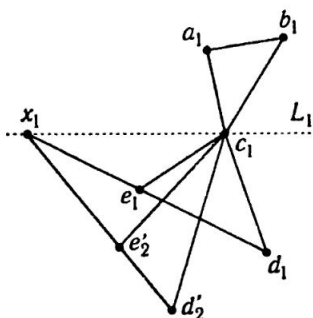


Рис. 14.

Положим теперь, что либо отрезок de не параллелен грани abc , либо отрезок ab не параллелен грани cde . Пусть, для определенности, отрезок de не параллелен грани abc . Обозначим через x точку пересечения отрезка de или его продолжения с прямой L , и соответственно через x_1 и x_2 — соответствующие точки на $\tilde{S}_1$ и $\tilde{S}_2$. Точки a, b, c и x лежат в теле T в одной плоскости, следовательно, $\tilde{S}_1(a_1, b_1, c_1, x_1) \approx \tilde{S}_2(a_2, b_2, c_2, x_2)$. Точки c, d, e и x в теле T тоже лежат в одной плоскости, и, значит, $\tilde{S}_1(c_1, d_1, e_1, x_1) \approx \tilde{S}_2(c_2, d_2, e_2, x_2)$. Совместим аффинными преобразованиями изображения $\tilde{S}_2$ его точки a_2, b_2 и c_2 с точками соответственно a_1, b_1 и c_1 на $\tilde{S}_1$. При этом точка x_2 совместится с точкой x_1 , а точки d_2 и e_2 займут положения, которые обозначим через d'_2 и e'_2 (рис. 14). Точка x_1 лежит на отрезке e_1d_1 или его продолжении. Точка x_2 — на отрезке e_2d_2 или его продолжении. После совмещения x_1 с x_2 точка x_1 должна, тем самым, быть пересечением отрезков e_1d_1 и $e'_2d'_2$ или их продол-

жений. Это и определяет положение прямой $\tilde{L}_1$ на $\tilde{S}_1$, а значит и $\tilde{L}_2$ на $\tilde{S}_2$. Лемма доказана.

Опишем процедуру 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 и 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 и c_1 с соответственно точками 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 и c_1 с точками соответственно a', b' и 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 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 и 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 и d_1 точки лежат в одной плоскости. В свою очередь, точки b_1, c_1, e_1 и d_1 лежат в одной плоскости, если при $a_1c_1e_1$ и $a_2c_2e_2$ – треугольниках $\tilde{S}_1(a_1, c_1, e_1, x_1) \approx \tilde{S}_2(a_2, c_2, e_2, x_2)$. 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 и d_1 должны лежать на одной прямой. Аналогично определяется принадлежность одной плоскости и точек b_1, c_1, e_1 и 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' .

26). Точки c , d и 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 4. Если S_1 и S_2 суть a' -разные проекции тела T , $\tilde{S}_1 \approx S_1$, $\tilde{S}_2 \approx S_2$, и тело T' построено по $\tilde{S}_1$ и $\tilde{S}_2$ использованием $\text{Alg } T'$ то тела T и T' – a' -эквивалентны.

Доказательство. Совместим аффинными преобразованиями точки a' , b' , c' и e' тела T' с точками соответственно a , b , c и e тела T . Покажем, что произвольная точка d' тела T' при этом совместится с соответствующей точкой d тела T . При этом будем опираться на ту последовательность действий и обозначений, которые использованы в описании $\text{Alg } T'$.

1. Если точка d_1 лежит в плоскости треугольника $a_1b_1c_1$, то $T(a, b, c, d) \approx \tilde{S}_1(a_1, b_1, c_1, d_1)$. По построению $T'(a', b', c', d') \approx \tilde{S}_1(a_1, b_1, c_1, d_1)$. Следовательно, $T(a, b, c, d) \approx T'(a', b', c', d')$ и, значит, при совмещении точек a' , b' , c' с точками соответственно a , b , c точка d' совместится с точкой d .

2. Пусть точка 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 и, для определенности, пусть $c_2d_2e_2$ – треугольник. По построению $T'(a', b', e', L') \approx \tilde{S}_1(a_1, b_1, c_1, \tilde{L}_1)$. Кроме того, по условию $T(a, b, e, L) \approx \tilde{S}_1(a_1, b_1, c_1, \tilde{L}_1)$. Отсюда $T'(a', b', e', L') \approx T(a, b, e, L)$ и, следовательно, при совмещении точек a' , b' , c' с точками соответственно a , b , c прямая L' совместится с прямой L .

По построению $T'(a', b', e', L') \approx \tilde{S}_2(a_2, b_2, e_2, \tilde{L}_2)$. По условию $T(a, b, e, L) \approx \tilde{S}_2(a_2, b_2, e_2, \tilde{L}_2)$. Отсюда $T(a, b, e, L) \approx \tilde{S}_2(a_2, b_2, e_2, \tilde{L}_2)$ и, следовательно, при совмещении прямой L' с прямой L и точки e' с точкой e точка d' совместится с точкой d .

Если $c_1d_1e_1$ и $c_2d_2e_2$ – не треугольники, то, в соответствии с описанием $\text{Alg } T'$, возможны два случая.

а). Точка d лежит на ребре ce (или его продолжении) тетраэдра $abcd$, то есть точки c , d , e лежат на одной прямой. В этом случае отношения длин отрезков ce , cd и ed при аффинных преобразованиях не меняются, то есть они одни и те же для троек $c_1d_1e_1$, $c_2d_2e_2$, ced и $c'e'd'$. Следовательно, при совмещении точек c и e с точками соответственно c' и e' точка d совместится с точкой d' .

Аналогично показываем совмещение точек d и d' и в тех случаях, когда точка d' лежит на ребрах ae и be тетраэдра $abcd$.

б). В этом случае точки c , d и e образуют в теле T грань, направления проекций α_1 и α_2 параллельны плоскости этой грани, и точка d не лежит на ребрах ae и be тетраэдра $abcd$. Тогда, например, из троек $a_1b_1d_1$ и $a_2b_2d_2$ хотя бы одна – треугольник, и, следовательно, рассуждения могут быть сведены к началу пункта 2 настоящего доказательства. Теорема доказана.

Содержательно теорема 4 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 $\tilde{S}_1$ и $\tilde{S}_2$ (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 тело восстановить нельзя. Из теоремы 4 следует, что для этого (с точностью до аффинных преобразований) достаточно двух проекций. Можно предположить, что это имеет некоторое отношение к тому обстоятельству, что зрительное восприятие в живых организмах повсеместно осуществляется посредством именно двух глаз.

Работа с восстановленным телом может включать в себя построение его проекций по произвольным направлениям, в том

числе и таким, для которых на каждом луче проекции может быть более чем одна точка. Такие проекции будем здесь называть неполными. Конечно, имея восстановленное тело, можно построить произвольную его проекцию. Покажем, 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.

Далее описывается процедура Alg S_0 построения по $\tilde{S}_1$ и $\tilde{S}_2$ некоторого изображения S_0 . Эта процедура похожа на Alg T' и это сходство мы постараемся при описании подчеркнуть, поскольку оно будет использовано при дальнейших доказательствах.

Пусть $a_1b_1c_1$ и $a_2b_2c_2$ – треугольники, и точка e_1 лежит вне плоскости треугольника $a_1b_1c_1$. Выберем на плоскости четыре точки a_0, b_0, c_0 и e_0 такие, что $a_0b_0c_0$ – треугольник. Точки a_0, b_0, c_0 и e_0 – точки строящегося изображения S_0 . Покажем, как строятся, причем однозначно, остальные точки изображения S_0 .

Возьмем на $\tilde{S}_1$ произвольную точку d_1 , отличную от точек a_1, b_1, c_1 и e_1 .

1. Пусть d_1 лежит в плоскости треугольника $a_1b_1c_1$. Совместим аффинными преобразованиями изображения $\tilde{S}_1$ точки a_1, b_1 и c_1 с точками соответственно a_0, b_0, c_0 . Точку, в которую перейдет при этом преобразовании d_0 , обозначим через d_1 .

2. Пусть d_1 лежит вне плоскости треугольника $a_1b_1c_1$.

Пусть хотя бы одна из троек точек $c_1d_1e_1$ 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 на $\tilde{S}_1$ и $\tilde{S}_2$ прямые $\tilde{L}_1$ и $\tilde{L}_2$. Аффинными преобразованиями изображения $\tilde{S}_1$ совместим точки a_1, b_1, c_1 с точками соответственно a_0, b_0, c_0 . Прямую, в которую преобразуется $\tilde{L}_1$, обозначим через $\tilde{L}_0$.

Пусть теперь, для определенности, $c_2d_2e_2$ – треугольник. Аффинными преобразованиями изображения $\tilde{S}_2$ совместим прямую $\tilde{L}_2$ с прямой $\tilde{L}_0$ и точки c_2 и e_2 – с точками соот-

ветственно c_0 и e_0 . Точку, в которую при этом перейдет d_2 , обозначим через d_0 .

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

а). Точки c , d и e в теле T лежат на одной прямой, то есть точка d лежит на ребре se тетраэдра $abce$ или на его продолжении. Распознается это обстоятельство так же, как и в соответствующем пункте описания $\text{Alg } T'$. Аффинными преобразованиями изображения $\tilde{S}_1$ совместим точки c_1 и e_1 с точками соответственно c_0 и e_0 . Точку, в которую перейдет при этом d_1 , обозначим через d_0 . Аналогичным образом поступаем и в случаях, когда точка d лежит на ребрах ae и be тетраэдра $abce$.

б). Точки c , d и e в теле T образуют грань и направления проекции α_1 и α_2 оба параллельны этой грани. Полагаем, что уже исключен случай расположения точки d на ребрах ae и be тетраэдра $abce$.

Таким образом, aed и bed – грани. Следовательно, грань, например, aed проецируется и на S_1 , и на S_2 в виде треугольника. Но тогда четверка $\langle (a_1b_1c_1)(a_2b_2c_2)(c_1d_1e_1)(c_2d_2e_2) \rangle$ – правильная, и построение точки d_0 сводится к началу пункта 2. Описание $\text{Alg } S_0$ закончено.

Из описания $\text{Alg } S_0$ нетрудно видеть, что если $S_0(a_0, b_0, c_0, e_0) \approx \tilde{S}_1(a_1, b_1, c_1, e_1)$ или $S_0(a_0, b_0, c_0, e_0) \approx \tilde{S}_2(a_1, b_1, c_1, e_1)$, то, соответственно $S_0 \approx \tilde{S}_1$ или $S_0 \approx \tilde{S}_2$. В противном случае изображение S_0 заведомо a' -разное и с $\tilde{S}_1$, и с $\tilde{S}_2$.

Теорема 5. Если S_1 и S_2 суть a' -разные проекции тела T , $\tilde{S}_1 \approx S_1$, $\tilde{S}_2 \approx S_2$ и изображение S_0 построено по $\tilde{S}_1$ и $\tilde{S}_2$ использованием $\text{Alg } S_0$, то существует тело T' , a' -эквивалентное телу T и такое, что изображение S_0 является его проекцией (возможно неполной) по некоторому направлению.

Доказательство. Выберем на $\tilde{S}_1$ и $\tilde{S}_2$ точки a_1, b_1, c_1, e_1 и a_2, b_2, c_2, e_2 такие, что $a_1b_1c_1$ и $a_2b_2c_2$ – треугольники, и точка e_1 лежит вне плоскости треугольника $a_1b_1c_1$. Используя алгоритмы $\text{Alg } T'$ и $\text{Alg } S_0$, построим тело T' и изображение S_0 .

Выберем некоторое направление проекции α' (не параллельное плоскости построенного изображения S_0). Проведем через каждую из точек a_0, b_0, c_0 и e_0 лучи проекции, параллельные α' , и возьмем на этих лучах по произвольной точке - обозначим их соответственно a'', b'', c'', e'' - с тем, однако, условием, чтобы они не лежали в одной плоскости. Совместим аффинными преобразованиями тела T' точки a', b', c', e' (см. Alg T') с точками соответственно a'', b'', c'' и e'' . Покажем, что тело T' после преобразования и есть искомое тело. Надо, очевидно, показать, что проекциями точек тела T по направлению α' будут после преобразования точки изображения S_0 .

Проекциями точек T' точки a', b', c', e' (совмещенными с a'', b'', c'' и e'') являются, по построению, точки a_0, b_0, c_0 и e_0 .

Рассмотрим произвольную точку d_1 (отличную от a_1, b_1, c_1 и e_1) на $\tilde{S}_1$. Ей соответствует на S_0 точка d_0 , и на T' - точка d' . Покажем, что проекцией точки d' является точка d_0 .

1. Пусть точка d_1 лежит в плоскости треугольника $a_1b_1c_1$. Тогда по построению $T'(a', b', c', d') \approx \tilde{S}_1(a_1, b_1, c_1, d_1)$ и $S_0(a_0, b_0, c_0, d_0) \approx \tilde{S}_1(a_1, b_1, c_1, d_1)$. Следовательно, $T'(a', b', c', d') \approx \tilde{S}_0(a_0, b_0, c_0, d_0)$. Известно, что параллельным проецированием точек одной плоскости на другую определяется аффинное преобразование точек одной плоскости в точки другой (и наоборот). Заданием направления проекции α' каждой точке плоскости изображения S_0 ставится в соответствие единственная точка плоскости изображения $T'(a', b', c', d')$, причем этим аффинным преобразованием точкам a_0, b_0 и c_0 ставятся в соответствие точки a', b' и c' . Следовательно, в силу того, что $T'(a', b', c', d') \approx \tilde{S}_0(a_0, b_0, c_0, d_0)$ точке d_0 будет поставлена в соответствие точка d' , то есть точка d_0 будет лежать на луче проекции от точки d' .

Итак, показано, что d_0 является проекцией точки d' для случая, когда d_1 лежит в плоскости треугольника $a_1b_1c_1$. Это соответствует случаю пунктов 1 в Alg T' и Alg S_0 . Повторяя аналогичным образом рассуждения для остальных пунктов (2, 2а и 2б), покажем, что произвольная точка d' имеет проекцией по направлению α' точку d_0 , что и завершит доказательство теоремы.

Посредством Alg S_0 получается изображение S_0 , про кото-

рое априори нельзя утверждать, что оно — проекция. Содержательный смысл теоремы 5 как раз и состоит в утверждении о том, что для любого таким образом построенного изображения S_0 можно так аффинно преобразовать тело T , что S_0 будет его проекцией по некоторому направлению.

В связи с этим встает такой вопрос: можно ли для тела T' , полученного из T аффинными преобразованиями (и, значит в частности, и для самого тела T), посредством $\text{Alg } S_0$ получить любую его проекцию?

Лемма 7. *Для любой (в том числе и неполной) проекции S_x объемного тела T существуют в T такие четыре точки, не лежащие в одной плоскости, что их проекции на S_x — разные точки.*

Доказательство. Возьмем на S_x точки a_x, b_x и c_x такие, что $a_x b_x c_x$ — треугольник. На S_x должны существовать точки, отличные по положению от a_x, b_x и c_x , ибо в противном случае T — не объемное тело. Пусть e_x — такая точка. Если соответствующая точка e лежит вне плоскости грани abc , то четверка точек a, b, c и e — искомая. Допустим противное, и точек на S_x , отличных по положению от точек a_x, b_x и c_x и таких, что соответствующие точки в теле T' лежат вне плоскости грани abc , нет. Тогда на каком-либо из лучей проекции, идущих от a, b и c , должна быть точка, отличная от a, b и c , ибо в противном случае T — двумерное тело. Пусть, для определенности, это будет точка v на луче проекции, идущем от точки a . Таким образом, проекция v_x точки v на S_x совпадает с a_x . Но тогда четверка v, e, b и c — искомая, если e не лежит на ребре bc или его продолжении.

Пусть e лежит на ребре bc или его продолжении. Тело T не может состоять только из точек a, b, c, e и v , ибо тогда T — не объемное тело. Значит, на каком-либо из лучей проекции, идущем от точек a, b и c , должна быть еще хотя бы одна точка, обозначим ее через u . Эта точка не может быть на луче от a , ибо в противном случае все точки тела T будут располагаться в пространстве на двух скрещивающихся прямых: точки a, v и u — на одной, и точки b, c и e — на другой. В этом случае тело T — не объемное. Значит точка u должна быть на луче либо от

точки b , либо от точки c . Положим, для определенности, что u расположена на луче от точки b . Тогда, например, четверка u, u, c и e — искомая. Лемма доказана.

Теорема 6. Если T — объемное тело, S_1 и S_2 — проекции его по разным направлениям α_1 и α_2 , $\tilde{S}_1 \approx S_1$, $\tilde{S}_2 \approx S_2$ и T' — некоторое тело, a' -эквивалентное телу T , то по $\tilde{S}_1$ и $\tilde{S}_2$ использованием $\text{Alg } S_0$ можно получить любую (включая неполные) проекцию тела T' .

Доказательство. Пусть S_x — произвольная (включая неполные) проекция тела T' по направлению α_x . Покажем, что она может быть получена по $\tilde{S}_1$ и $\tilde{S}_2$ использованием $\text{Alg } S_0$.

Рассмотрим четыре разные точки a_x, b_x, c_x и e_x на S_x такие, что соответствующие точки a', b', c' и e' в теле T' образуют тетраэдр. Согласно лемме 7 такие точки есть. Рассмотрим тетраэдр $a'b'c'e'$ в теле T' и соответствующий тетраэдр $abce$ в теле T . В каждом из этих тетраэдров по четыре грани. Каждое из направлений проекций α_x, α_1 и α_2 может быть параллельным только одной грани (в своем тетраэдре). Следовательно, есть такая грань в T' и соответствующая грань в T , что они не параллельны α_x в T' и α_1 и α_2 в T . Пусть, для определенности, это будет грань $a'b'c'$ в T' и соответственно грань abc в T . Следовательно, $a_x b_x c_x, a_1 b_1 c_1$ и $a_2 b_2 c_2$ — треугольники.

Выберем в качестве исходных четверок точек на $\tilde{S}_1$ и $\tilde{S}_2$ для $\text{Alg } S_0$ точки a_1, b_1, c_1, e_1 и a_2, b_2, c_2, e_2 . В качестве точек a_x, b_x, c_x, e_x возьмем точки соответственно a_0, b_0, c_0, e_0 . Остальные точки изображения S_0 строятся, при заданных a_x, b_x, c_x, e_x однозначно. Построениями, повторяющими использованные в теоремах 4 и 5, нетрудно показать, что каждая точка d_0 , строящаяся посредством $\text{Alg } S_0$, совместится с соответствующей точкой d_x на проекции S_x , что и завершает доказательство теоремы.

Теорему 5 можно рассматривать как в некотором смысле прямое утверждение: если есть две проекции тела T , то в результате использования $\text{Alg } S_0$ получается изображение, которое тоже является проекцией тела T (или тела, a' -эквивалентного телу T). Теорему 6 можно интерпретировать

как своего рода обратное утверждение: если S_x — проекция тела (или тела, a' -эквивалентного телу T), то S_x можно получить как результат использования Alg S_0 .

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 $S'_1 \approx S_1$, $S'_2 \approx S_2$ и точки x_{i_k} и y_{j_k} являются проекциями точки z_k . 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 $\tilde{S}_1(x_{i_a}, x_{i_b}, x_{i_c}, x_{i_e}) \neq \tilde{S}_2(y_{j_a}, y_{j_b}, y_{j_c}, y_{j_e})$. Проведем лучи, параллельные некоторому направлению a' через все точки изображения $\tilde{S}_1$. На лучах, проходящих через $x_{i_a}, x_{i_b}, x_{i_c}$ и x_{i_e} возьмем точки z'_a, z'_b, z'_c и z'_e так, чтобы они не лежали в одной плоскости. Если разметка приемлемая, то далее использованием 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 a'' , на лучах через точки $y_{j_a}, y_{j_b}, y_{j_c}$ и y_{j_e} отметим точки z''_a, z''_b, z''_c и 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 и z'_e с точками z''_a, z''_b, z''_c и z''_e . Если тела T' и T'' существуют, то, в силу теоремы 4, 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 $x_{i_a}, x_{i_b}, x_{i_c}$ и

x_{i_e} и y_{j_a} , y_{j_b} , y_{j_c} и 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 .

Глава II

Зрительный образ и геометрические преобразования изображений

Применительно к изображениям говорят о зрительных образах. Определения образа нет, и конкретный образ задается, как правило, перечислением примеров изображений, в него входящих. Если исходить из этого, то, тем самым, образ – совокупность изображений, объединенных в одну группу (потенциально бесконечную), причем по трудноформализуемым критериям. В словесном выражении эти критерии чаще всего звучат так: изображения группы «похожи», они «одной формы», «одного типа», одинаковы в своей «сущности», и т.п. За этим должны стоять какие-то алгоритмы, позволяющие делать такие суждения. Попытаемся по шагам проследить те правила, которые можно положить в основу формальных критериев объединения изображений в группы, с тем, чтобы они (эти группы) в как можно большей степени удовлетворяли бы интуитивным представлениям о зрительном образе.

Наиболее наглядный и вследствие этого часто рассматриваемый случай – образы букв и цифр, поэтому будем вести рассуждения на их примере. Возьмем, для конкретности, образ цифры два. Во множество изображений, объединяемых под общим названием «двойка», входят очень разные изображения. Это может быть и кривая линия на листе бумаги, и гигантская фигура на рекламном щите, и значок из стандартного шрифта на экране компьютера, и пр. Вопрос в том, как конкретизировать, на чем основывать интуитивно очевидное сходство в форме этих изображений? Можно положить, в качестве первого шага, что нагляднее и очевиднее всего «одинаковость формы» для двух изображений имеет место, если их можно просто наложить друг на друга и они при этом в точности совпадут.

Наложить, очевидно, движениями, то есть, в этом случае, параллельными переносами, поворотами и, может быть, преобразованиями симметрии. Таким образом, если взять пример на рис. 15, то в одну группу с ним попадут все изображения, которые можно получить из этого рисунка переносами и поворотами.



Рис. 15.

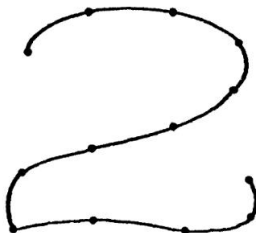


Рис. 16.



Рис. 17.

Конечно, такое определение будет недостаточным хотя бы потому, что в эту группу не попадут изображения, отличающиеся от двойки на рис. 15 размерами. Поэтому на следующем шаге добавим изменение масштаба, и, тем самым, будем рассматривать все, что можно получить из рис. 15 преобразованиями подобия.

Наконец, добавление возможности растягивать и сжимать изображение (рисунки 16 и 17) по любому направлению приводит к аффинным преобразованиям в целом. Этими преобразованиями из рисунка 15 порождается уже довольно широкий класс изображений, они, очевидно, отличаются друг от друга довольно значительно. В главе I рассмотрен код, позволяющий описывать любое изображение, получаемое из данного произвольным аффинным преобразованием, и только такие изображения. В этом смысле код – эквивалент группы (континуальной по мощности) изображений, получаемых из данного изображения произвольными аффинными преобразованиями.

Разумеется, класс изображений, отличающихся друг от дру-

га произвольными аффинными преобразованиями – еще далеко не то, что содержательно понимается под образом. Если на рис. 15 какую-либо из точек несколько сдвинуть, оставляя остальные точки неподвижными, то такую фигуру уже никакими аффинными преобразованиями не сделать совпадающей с исходной. А ведь именно своеобразная «инвариантность» по отношению к локальным изменениям формы должна иметь место, если ориентироваться на интуитивные, содержательные представления о том, что есть образ.

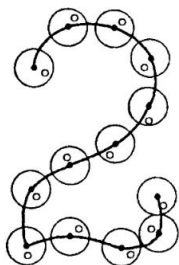


Рис. 18.

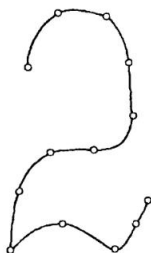


Рис. 19.

В этой главе делается попытка сделать очередной шаг в приближении к понятию зрительного образа на основе следующих соображений. Пусть дано изображение цифры два. Обозначим его через A . Положим, что каждая из точек изображения A – центр круга радиуса r (рис. 18). Отметим в каждом круге по точке (кружочки на рис. 18). В совокупности они тоже составляют точечное изображение (центр кружочка – точка), обозначим его через B . Изображение B смещениями и поворотами можно вынести за пределы изображения A и рассматривать отдельно (рис. 19). Ясно, что в общем случае оно отлично от изображения A в том смысле, что никакими переносами и поворотами нельзя совместить все точки рисунка 19 с точками рисунка 18. Вместе с тем из «происхождения» рисунка 19 следует, что параллельными переносами и поворотами можно его так расположить на рисунке 18, что каждая точка изображения B будет отстоять от соответствующей точки изображения A на расстояние, не большее r . Назовем такое взаиморасположение для A и B искомым. Тогда именно это обстоятельство

– возможность искомым образом расположить изображения A и B – и может служить проявлением критерия «одинаковости формы», который мы здесь и полагаем рассматривать.

Если рисунок 19 дан «сам по себе», без указания на его «происхождение», то априори неизвестно, можно ли A и B расположить искомым образом. Поскольку множество возможных взаиморасположений рисунков 18 и 19, очевидно, континуальное по мощности, то эту задачу нельзя решить перебором. Поэтому в первую очередь нужно ответить на вопрос о том, можно ли и как решить эту задачу конечными средствами, то есть использованием некоторой конечной процедуры. Под конечной процедурой понимается здесь решение задачи геометрическим построением, или сведением к решению некоторой системы уравнений. Можно выразиться так, что решение должно быть технологичным по характеру.

Вопрос, связанный с определением искомого положения, можно усилить: каков минимальный возможный радиус r , при котором еще можно так расположить изображения A и B , что их соответствующие точки взаимоудалены на расстояние, не большее r , и какими средствами (конечными) осуществить такое их расположение? Именно этот вопрос и то, что с ним связано, рассматриваются в этой главе. В §1 рассматривается случай, когда преобразованиями, которым подвергаются изображения, являются преобразования параллельного переноса, в §2 – в целом изометрические преобразования.

Ясно, что решение этой задачи может быть положено в основу процедур распознавания изображений. Двойку на рис. 19 расположить на изображении другой цифры так, чтобы соответствующие точки находились на расстоянии, не большем того же r , очевидно, уже не удастся, и на этом можно основывать их различение.

Рассмотрения §1 и §2 недостаточны, если иметь в виду то, что на практике изображения могут разниться, например, по размерам. Поскольку в целом эта работа ориентирована на применение в конечном счете к реальному распознаванию образов, то такого рода различия должны быть учтены в общем решении. Поэтому в §3 рассматривается та же задача, но вместо изометрических преобразований рассматривается более широ-

кий класс – преобразования подобия, в §4 – аффинные преобразования.

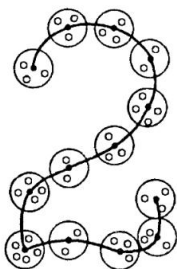


Рис. 20.

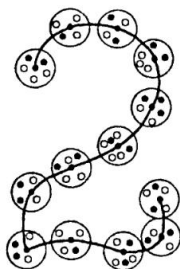


Рис. 21.

Везде выше соответствие между точками изображений A и B предполагалось взаимно однозначным, то есть изображения A и B предполагались состоящими из одинакового числа точек. Это действительно основной, базовый случай, и именно он рассматривается далее. Однако, нетрудно видеть, решение для этого случая позволит сделать потом естественные его расширения: когда группа точек изображения B сопоставляется одной точке изображения A (рис. 20), и когда сопоставляются друг другу группы точек изображений A и B (рис. 21).

§1. Преобразования параллельного переноса

Содержательно задачу этого параграфа можно описать следующим образом. Пусть даны два изображения A и B , и каким-либо образом точки из A и B взаимно однозначно сопоставлены друг другу. Этим определена взаимоудаленность для каждой пары соответствующих друг другу точек на A и B – она есть длина отрезка, концами которого являются эти точки. Можно считать характеристикой данного взаиморасположения изображений A и B величину наибольшего из этих отрезков и называть ее взаимоудаленностью точек в A и B . Будем теперь двигать A и B (параллельными переносами) всеми возможными способами. Можно ли этим так установить изображения относительно друг друга, чтобы взаимоудаленность точек в A и B была бы наименьшей? Ясно, что это нельзя сделать просто

перебором, поскольку разных положений для A и B при параллельных переносах бесконечное множество. Можно ли найти нужное взаиморасположение конечной процедурой?

Итак, изображением называем конечное (непустое) множество точек на плоскости. Пусть изображение A состоит из n ($n \geq 1$) точек. Перенумеруем некоторым образом точки из A , точку с номером i будем обозначать через a_i ($i = 1, \dots, n$). Через $\bar{A}$ обозначим множество всех изображений, которые получаются из A параллельными переносами. При этом если $A' \in \bar{A}$, то полагаем, что на нем сохраняется нумерация точек, порожденная A , то есть через a'_i на A' обозначена точка, в которую при соответствующем параллельном переносе переходит точка a_i из A . Изображения A' и A'' из $\bar{A}$ называем эквивалентными.

Любая часть точек $a_{i_1}, \dots, a_{i_k}$ ($k \leq n$) из A может также рассматриваться как изображение. Его называем частью изображения A (при $k < n$ – собственной частью).

Пусть изображения A и B состоят из точек соответственно $a_1, \dots, a_n$ и $b_1, \dots, b_n$. Будем рассматривать все возможные биекции на множествах точек изображений A и B . Зафиксируем какую-либо из них и обозначим ее через ψ . Без ограничения общности можем полагать, что ею точке a_i ставится в соответствие точка b_i ($i = 1, \dots, n$). Точки a_i и b_i называем соответствующими, отрезки $(a_i a_j)$ и $(b_i b_j)$ тоже называем соответствующими.

Введем в рассмотрение изображение C , которое будем называть характеристическим, и которое получается следующим образом. Фиксируем произвольную точку O на плоскости. Параллельным переносом отрезка $(a_i b_i)$ совмещаем точку a_i с точкой O . Точку, в которую переходит при этом b_i , обозначаем через c_i . Точку c_i называем порожденной точками a_i и b_i , а точки a_i и b_i – соответствующими точке c_i . Изображение C , полагаем, состоит из точки O и точек c_i ($i = 1, \dots, n$). Точку O называем центром изображения C , а часть, составленную из точек $c_1, \dots, c_n$ – его ядром. При таком определении несколько точек $b_{i_1}, \dots, b_{i_k}$ могут при параллельном переносе перейти в одну. Такую точку называем кратной и сохраняем приписанные ей номера, то есть обозначаем ее через $c_{i_1}, \dots, c_{i_k}$. Изображение с

кратными точками называем особым. Тем самым, следовательно, C может быть особым изображением.

Взяв вместо центра O иную точку O' на плоскости, описанной процедурой получим характеристическое изображение C' из точки O' и точек c'_i ($i = 1, \dots, n$). Очевидно, что C' и C эквивалентны. Точки c_i из C и c'_i из C' , порожденные одной и той же парой точек a_i и b_i , называем соответствующими. Класс изображений, получаемых таким образом при всех возможных центрах на плоскости, обозначим через $(B - A)$. Все изображения из $(B - A)$, следовательно, переводимы друг в друга параллельными переносами.

Поменяв ролями A и B , можно определить $(A - B)$. Изображения из $(B - A)$ и $(A - B)$ переводимы друг в друга параллельными переносами и преобразованиями симметрии.

По A и любому $C \in (B - A)$ построим изображение, которое обозначим через $A + (B - A)$. Берем отрезок (Oc_i) из изображения C и параллельным переносом совмещаем точку O с точкой a_i . Точка, в которую при этом переходит c_i , обозначаем через b_i . Нетрудно видеть, что изображение из точек b_i ($i = 1, \dots, n$), возникающее таким образом, есть изображение B .

Пусть $r_\psi(A, B)$ — длина наибольшего из отрезков (Oc_i) ($i = 1, \dots, n$) в изображении C , то есть то, что выше было названо взаимоудаленностью точек в A и B . Очевидно, что эта величина не зависит от конкретного $C \in (B - A)$. Ясно, что и $r_\psi(A, B) = r_\psi(B, A)$, где $r_\psi(B, A)$ определяется по произвольному $C' \in (A - B)$. Отметим, что если на C рассматривать круг с центром в точке O и радиуса $r_\psi(A, B)$, то все точки изображения C будут находиться внутри круга или на окружности.

Рассмотрим бинарное отношение P на декартовом произведении $\bar{A} \times \bar{B}$: полагаем, что пары (A_1, B_1) и (A_2, B_2) из $\bar{A} \times \bar{B}$ находятся в отношении P , если (A_1, B_1) как целое переводится параллельным переносом в пару (A_2, B_2) . По отношению P множество $\bar{A} \times \bar{B}$ разбивается на множество P^* классов эквивалентности, и, очевидно, P^* имеет мощность континуума.

Содержательно каждый класс P' из P^* задает один вариант взаиморасположения изображений A и B с точностью до параллельных переносов пары (A, B) как целого на плоскости.

Из этого следует, что всеми парами (A, B) из P' задается одно и то же, с точностью до параллельных переносов, характеристическое изображение, то есть класс попарно эквивалентных изображений. Этот класс обозначим через $\bar{C}'$. Очевидно, что $\bar{C}'$ совпадает с классом $(B - A)$, порожденным любой парой (A, B) из P' .

Разным классам P_1 и P_2 из P^* соответствуют разные классы $\bar{C}_1$ и $\bar{C}_2$ характеристических изображений. Действительно, рассмотрим такие две пары изображений из соответственно P_1 и P_2 , у которых первые элементы совпадают, то есть (A, B') и (A, B'') . Это интерпретируется как фиксированное изображение A и сдвинутые по отношению друг к другу параллельным переносом изображения B' и B'' . Последнее означает, что все отрезки (a_i, b'_i) разнятся с отрезками (a_i, b''_i) ($i = 1, \dots, n$) по длине или углу наклона, и, следовательно, параллельными переносами не совмещаются. Но тогда, соответственно, и $\bar{C}_1$ с $\bar{C}_2$ не совпадают.

Множество классов характеристических изображений обозначим через C^* . Каждому классу из P^* соответствует, как показано, один и только один класс из C^* . Все изображения из одного класса $\bar{C}'$ характеристических изображений имеют одну и ту же величину $r_\psi(A, B)$, которая, тем самым, есть и характеристика соответствующего класса P' из P^* . Пусть на классе P_0 из P^* достигается минимум величины $r_\psi(A, B)$. Если такой класс P_0 существует, то назовем его главным.

Итак, изображения внутри каждого класса в C^* попарно эквивалентны, изображения из разных классов – не эквивалентны. Каждое такое изображение состоит из центра O и точек $c_1, \dots, c_n$ ядра.

Лемма 1. *Ядра всех характеристических изображений из всех классов в C^* эквивалентны.*

Доказательство. Рассмотрим какую-либо пару (A, B) из какого-либо класса в P^* . Зафиксируем некоторую точку изображения A , пусть это будет точка a_i (рис. 22). Рассмотрим характеристическое изображение C пары (A, B) . Без ограничения общности можем полагать, что центр O характеристического изображения совпадает с точкой a_i . Тогда точка c_i

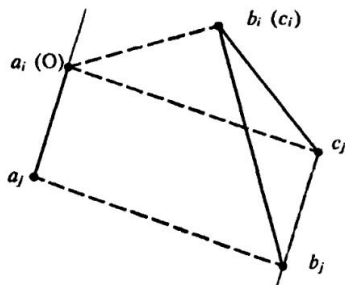


Рис. 22.

ядра совпадает с точкой b_i . Определяем остальные точки c_j ($j = 1, \dots, n, j \neq i$) ядра. Отрезок (Oc_i) получается параллельным переносом отрезка $(a_j b_j)$ с совмещением точки a_j с точкой O , то есть с точкой a_i . В параллелограмме $a_i a_j b_j c_j$ сторона $(c_j b_j)$ может рассматриваться как полученная параллельным переносом отрезка $(a_i a_j)$. Рассмотрим треугольник $c_i c_j b_j$. Он определен однозначно: сторона $(c_i b_j)$ есть отрезок $(b_i b_j)$ из изображения B , сторона $(c_j b_j)$ получается параллельным переносом отрезка $(a_i a_j)$ из изображения A с совмещением точки a_j с точкой b_j . Тем самым отрезок $(c_i c_j)$ ($j = 1, \dots, n, j \neq i$) определен однозначно, а значит определены и все точки ядра. Нетрудно видеть, что в построении треугольника $c_i c_j b_j$ никак не сказывается конкретное взаиморасположение изображений A и B , что и доказывает лемму.

Замечание. В треугольнике $c_i c_j b_j$ угол, противолежащий стороне $(c_i c_j)$, обозначим через α_{ij} и будем называть углом в характеристическом изображении, определяемым соответствующими отрезками $(a_i a_j)$ и $(b_i b_j)$. Длина отрезка $(c_i c_j)$ может быть вычислена, поскольку соответствующие отрезки и угол α_{ij} , определяемый ими, при заданных A и B известны. Тем самым все попарные расстояния между точками ядра можно считать известными.

Поскольку характеристические изображения из разных классов в C^* не эквивалентны, то, следовательно, при совмещении параллельными переносами точек их ядер у них не совпа-

дуг именно центры. Таким образом, содержательно различие между не эквивалентными характеристическими изображениями состоит в различии положения центра относительно точек ядра.

Можно ли утверждать, что и произвольно взятая в качестве центра точка вместе с точками ядра есть в совокупности некоторое характеристическое изображение?

Пусть характеристическое изображение C состоит из центра O и точек $c_1, \dots, c_n$ ядра. Изображение из точек $c_1, \dots, c_n$ и произвольной точки O_x обозначим через C_x .

Лемма 2. *Изображение C_x является характеристическим.*

Доказательство. Достаточно показать, что существует такая пара из $\bar{A} \times \bar{B}$, для которой изображение C_x является характеристическим.

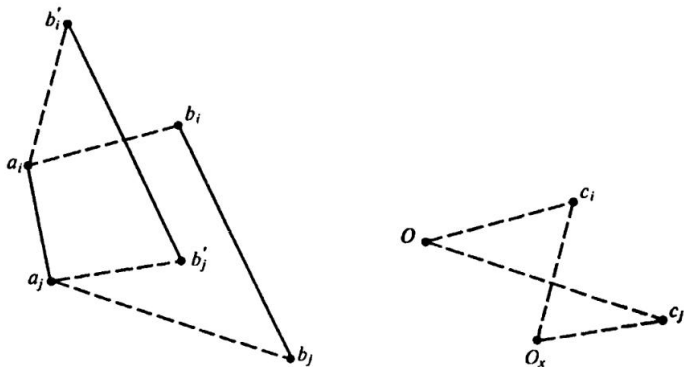


Рис. 23.

На рис. 23 точки a_i и a_j принадлежат изображению A , точки b_i и b_j — изображению B . Строим точки b'_i и b'_j , принадлежащие изображению $(A + C_x)$: точку b'_i получаем параллельным переносом отрезка $(O_x c_i)$ с совмещением точки O_x с точкой a_i , и точку b'_j — параллельным переносом отрезка $(O_x c_j)$ с совмещением точки O_x с точкой a_j . Рассмотрим треугольники $b'_i a_i b_i$ и $O c_i O_x$. Стороны $(a_i b_i)$ и $(c_i O)$, $(a_i b'_i)$ и $(c_i O_x)$ у них равны и параллельны, углы $b'_i a_i b_i$ и $O a_i O_x$ равны. Отсюда сторона

$(b'_i b_i)$ равна и параллельна стороне (OO_x) . Из аналогичного рассмотрения для треугольников $b'_j a_j b_j$ и $O c_j O_x$ получаем, что и $(b'_j b_j)$ равна и параллельна (OO_x) . Таким образом, $(b'_i b_i)$ равна и параллельна $(b'_j b_j)$ ($i \neq j$, $i, j = 1, \dots, n$), и, следовательно, изображение из точек b'_i ($i = 1, \dots, n$) может быть получено параллельным переносом изображения B . Лемма доказана.

Итак, произвольная точка O_x вместе с точками ядра дает характеристическое изображение C_x , которое принадлежит ровно одному классу в C^* . Поскольку классам в C^* взаимно однозначно соответствуют классы в P^* , то, тем самым, содержательно C_x определяет ровно один вариант взаиморасположения изображений A и B . Взаимоудаленность точек изображений A и B определяется длиной наибольшего из отрезков $(O_x c_i)$ ($i = 1, \dots, n$), то есть радиусом окружности с центром в O_x , вмещающей в себя все точки ядра. Ясно, что существует единственная наименьшая по радиусу такая окружность, ее центр и радиус определяются однозначно. Назовем эту окружность ключевой, соответствующий класс из P^* есть, очевидно, то, что выше было названо главным классом.

Предшествующими рассуждениями, включающими леммы 1 и 2, доказано следующее утверждение.

Теорема 1. *В P^* существует и единственен главный класс.*

Содержательно это означает, что существует единственный вариант взаиморасположения изображений A и B (назовем его тоже главным), при котором взаимоудаленность их точек минимальна.

При заданных изображениях A и B главное их взаиморасположение геометрически строится следующим образом. Выбираем произвольную точку O на плоскости в качестве центра, и параллельными переносами отрезков $a_i b_i$ ($i = 1, \dots, n$), с совмещением точки a_i с точкой O , строим характеристическое изображение. Этим определяется совокупность точек c_i ($i = 1, \dots, n$) ядра. Строим ключевую окружность, ее центр O_x вместе с ядром образуют характеристическое изображение C_x . Далее строим $A + C_x$, что дает изображение B' , эквивалентное B . Взаиморасположение A и B' - главное.

Рассмотрим, в качестве иллюстрации, примеры, в которых изображение A состоит из точек a_1, a_2, a_3 , изображение B – из точек b_1, b_2, b_3 , треугольники $a_1a_2a_3$ и $b_1b_2b_3$ подобны и стороны их параллельны.

Пример 1. Этот случай можно назвать вырожденным, треугольники $a_1a_2a_3$ и $b_1b_2b_3$ в нем равны (рис. 24), точки ядра c_1, c_2, c_3 слиты в одну.

Пример 2. В этом примере треугольник $c_1c_2c_3$ тупоугольный (рис. 25), поэтому центр O_x располагается на середине наибольшей стороны треугольника. На ключевой окружности находятся две точки ядра, одна точка ядра – внутри круга.

Пример 3. В этом случае треугольник $c_1c_2c_3$ остроугольный (рис. 26(а)), центр O_x ключевой окружности – центр описанной окружности треугольника $c_1c_2c_3$. На ключевой окружности – три точки ядра. На рис. 26(б) построено главное взаиморасположение изображений A и B .

Радиус ключевой окружности может быть определен не только геометрическим построением, но и вычислением. В соответствии с замечанием к лемме 1 все попарные расстояния между точками ядра можно считать известными. Далее вычисления разбиваются на два случая, иллюстрируемых примерами 2 и 3 выше.

Первый случай. На ключевой окружности находятся две точки, c_{i_1} и c_{i_2} , причем отрезок $(c_{i_1}c_{i_2})$ – диаметр окружности, а все остальные точки ядра находятся внутри круга. Необходимым и достаточным условием этого является то, что отрезок $(c_{i_1}c_{i_2})$ – наибольший по длине, и треугольники $c_{i_1}c_{i_2}c_i$ ($i = 1, \dots, n, i \neq i_1, i_2$) – тупоугольные (включая сюда случай, когда c_i лежит на диаметре). При известных попарных расстояниях между точками ядра это условие проверяется несложно.

Сюда же отнесем и, в некотором смысле, пограничный случай, когда на ключевой окружности, помимо точек c_{i_1} и c_{i_2} , есть еще и другие точки ядра. В этом случае они должны с точками c_{i_1} и c_{i_2} образовывать прямоугольный треугольник.

Второй случай. Положим, что первый случай уже исключен. Это значит что на ключевой окружности лежат три и больше точек, причем никакая пара из них не образует диаметр окружности. В этом случае ключевая окружность является описанной

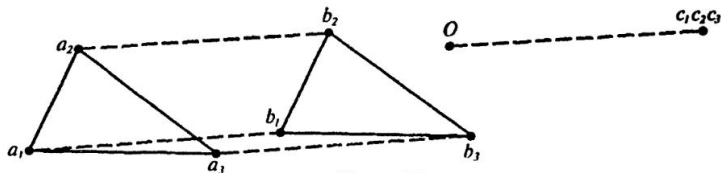


Рис. 24.

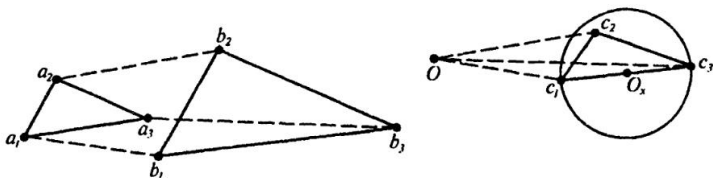


Рис. 25.

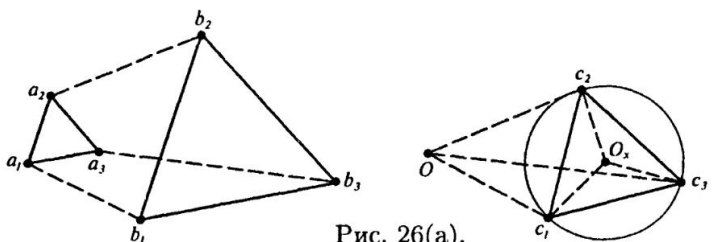


Рис. 26(а).

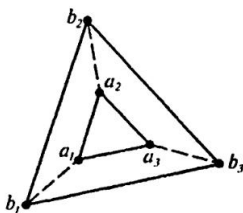


Рис. 26(б).

около треугольника из точек, лежащих на окружности, и этот треугольник – остроугольный. Для нахождения радиуса ключевой окружности надо перебрать все остроугольные треугольники с вершинами в точках ядра и определить треугольник с наибольшей по радиусу описанной окружностью. Она и будет ключевой, все остальные точки ядра будут содержаться внутри круга (или, в крайнем случае, на самой окружности,

когда остроугольных треугольников с наибольшим радиусом описанной окружности несколько).

Предшествующее описание относилось к случаю, когда биекцией ψ точки $a_1, \dots, a_n$ изображения A сопоставлялись точкам $b_1, \dots, b_n$ изображения B . Обозначим через $r(A, B) = \min_{\psi} \{r_{\psi}(A, B)\}$, где минимум берется на множестве всех величин $r_{\psi}(A, B)$, полученных при всех возможных биекциях на множествах точек изображений A и B . Содержательно $r(A, B)$ определяет наименьшую взаимоудаленность точек изображений A и B при всех возможных параллельных переносах изображений и при всех возможных вариантах сопоставлений их точек друг другу.

§2. Изометрические преобразования изображений

Содержательно задача для этого параграфа повторяет задачу для параграфа предшествующего, но с добавлением возможности поворачивать изображения относительно друг друга и использовать преобразования симметрии. Таким образом, для двух изображений нужно параллельными переносами, преобразованиями симметрии и поворотами, то есть изометрическими преобразованиями, так расположить их, чтобы взаимоудаленность их точек была бы наименьшей.

Через $\tilde{A}$ обозначим множество всех изображений, получаемых из изображения A параллельными переносами и поворотами. При этом для $A' \in \tilde{A}$ полагаем, что сохраняется нумерация точек, порожденная изображением A , то есть через a'_i обозначена та точка на A' , в которую перешла точка a_i ($i = 1, \dots, n$) при соответствующем вращении и переносе изображения A . Изображения A' и A'' из $\tilde{A}$ называем эквивалентными по переносу и повороту (далее, для краткости, просто эквивалентными), точки a'_i и a''_i ($i = 1, \dots, n$) — тоже эквивалентными. Части изображений A' и A'' , состоящие из эквивалентных точек, называем эквивалентными.

Рассматриваем биекцию ψ , которая точкам $a_1, \dots, a_n$ изображения A ставит в соответствие точки соответственно

$b_1, \dots, b_n$ изображения B . Для изображений $A' \in \tilde{A}$ и $B' \in \tilde{B}$ точки a'_i и b'_i ($i = 1, \dots, n$) называем соответствующими. Части изображений A' и B' , состоящие из соответствующих точек, называем соответствующими.

Рассматриваем декартово произведение $\tilde{A} \times \tilde{B}$. Введем понятие угла φ между изображениями в паре (A, B) . Для этой цели зафиксируем в изображениях A и B пары отрезков $(a_p a_q)$ ($p \neq q$), и соответственно $(b_p b_q)$. В качестве угла φ выберем угол, образованный этими отрезками (или их продолжениями). Для произвольной пары изображений (A', B') из $\tilde{A} \times \tilde{B}$ угол φ между ними есть угол между соответствующими отрезками $(a'_p a'_q)$ и $(b'_p b'_q)$ или их продолжениями. Без ограничения общности можно полагать, что у исходных изображений A и B отрезки $(a_p a_q)$ и $(b_p b_q)$ параллельны и угол φ равен нулю, а для всех пар из $\tilde{A} \times \tilde{B}$ этот угол находится в промежутке от 0 до 2π .

На множестве $\tilde{A} \times \tilde{B}$ рассматриваем бинарное отношение P : пары $(A_1 B_1)$ и $(A_2 B_2)$ находятся в отношении P , если углы между изображениями в парах одинаковы. Это отношение порождает разбиение множества $\tilde{A} \times \tilde{B}$ на классы эквивалентности. Каждому классу соответствует свое значение угла φ между изображениями A и B , поэтому класс обозначим через P'_φ , множество классов – через $\{P'_\varphi\}_{0 \leq \varphi < 2\pi}$.

Пусть Q' и Q'' – ядра характеристических изображений произвольных пар (A', B') и (A'', B'') из P'_φ , c'_i и c'_j – точки из Q' , c''_i и c''_j – соответствующие точки из Q'' ($i, j = 1, \dots, n$, $i \neq j$). Длины отрезков $(c'_i c'_j)$ и $(c''_i c''_j)$ определяются только длинами соответствующих отрезков $(a'_i a'_j)$, $(b'_i b'_j)$ и $(a''_i a''_j)$, $(b''_i b''_j)$ и углом между ними, и потому одинаковы. Множество ядер всех характеристических изображений всех пар (A', B') из P'_φ обозначим через Q_φ . Из сказанного следует, что все изображения из Q_φ переводятся друг в друга преобразованиями параллельного переноса и поворота. Поэтому можно говорить об общей для всех ядер из Q_φ величине радиуса ключевой окружности, обозначим ее через R_φ . Через P_φ обозначим подкласс класса P'_φ , состоящий из всех таких пар, у которых центр характеристического изображения совпадает с центром ключевой окружности.

Величина R_φ есть функция от φ . Минимальное значение

R_φ ($0 \leq \varphi < 2\pi$) обозначим через $R_\varphi^{\min}$. Угол φ_0 такой, что $R_{\varphi_0} = R_\varphi^{\min}$ назовем искомым. Взаиморасположение A' и B' в парах (A', B') из P_{φ_0} тоже назовем искомым. Содержательно искомое взаиморасположение A' и B' означает, что их точки взаимоудалены на минимальное расстояние, какого можно добиться параллельными переносами и поворотами изображений A и B .

Мы используем далее в целом эквивалентное прежнему, но более наглядное и удобное представление о множестве классов $\{P'_\varphi\}_{0 \leq \varphi < 2\pi}$ и об искомом взаиморасположении. Зафиксируем какое-либо изображение из $\tilde{A}$, пусть, для определенности, это будет исходное изображение A . Из каждого класса P_φ возьмем ту пару (A, B_φ) , у которой первый элемент совпадает с A . Если продельвать это последовательно для всех φ от 0 до 2π , то наглядно это можно представить как неподвижное изображение A и последовательно поворачивающееся относительно A изображение B , причем в каждый момент взаиморасположение A и B – главное.

Точки, из которых состоит B_φ , обозначим через $b_1^\varphi, \dots, b_n^\varphi$. Для пары (A, B_φ) строится характеристическое изображение, состоящее из точек $c_1^\varphi, \dots, c_n^\varphi$ ядра и центра O , совпадающего с центром ключевой окружности. Точка O фиксирована и центры всех ключевых окружностей (при разных φ) совпадают с точкой O . При таком представлении величина R_φ и изображение из точек $c_1^\varphi, \dots, c_n^\varphi$ – функции от угла φ . Вопрос в том, при каких углах поворота величина R_φ наименьшая?

Рассмотрим сначала для примера частный случай, в котором поставленная задача решается просто. Пусть изображения A и B таковы, что они переводимы друг в друга подобным преобразованием. Положим для конкретности, что B получено из A увеличением или уменьшением в размерах с коэффициентом k и поворотом на угол α . При повороте затем B на произвольный угол φ попарные расстояния между точками c_i^φ и c_j^φ ядра ($i, j = 1, \dots, n, i \neq j$) находятся из формулы $(c_i^\varphi c_j^\varphi)^2 = (a_i a_j)^2 + (b_i b_j)^2 - 2(a_i a_j)(b_i b_j) \cos(\alpha + \varphi)$. Поскольку $(b_i b_j) = k(a_i a_j)$, то $(c_i^\varphi c_j^\varphi)^2 = (a_i a_j)^2 (1 + k^2 - 2k \cos(\alpha + \varphi))$. Это значит, что при вращении изображения B ядро характе-

ристического изображения пары (A, B) меняется, оставаясь подобным самому себе (и, заметим, изображениям A и B). Наименьшими все попарные расстояния между точками ядра, а, следовательно, и наименьший радиус ключевой окружности, будут при $(\alpha + \varphi) = 2\pi$. Таким образом, искомое взаиморасположение для A и B характеризуется тем, что все соответствующие отрезки на этих изображениях (то есть $(a_i a_j)$ и $(b_i b_j)$, $i, j = 1, \dots, n$, $i \neq j$) параллельны друг другу.

В рассмотренном примере ядро, меняясь в размерах, остается подобным себе, и потому на ключевой окружности при всех углах поворота находятся одни и те же точки (то есть те точки, которые выше были названы соответствующими). В общем случае точки ядра меняют расположение относительно друг друга, и потому при разных φ на ключевой окружности могут оказываться разные точки.

Итак, имеем изображения A из точек $a_1, \dots, a_n$, B_φ из точек $b_1^\varphi, \dots, b_n^\varphi$ и C^φ из точек $c_1^\varphi, \dots, c_n^\varphi$; R_φ — радиус ключевой окружности для C^φ . Пусть изображение A_i есть часть A из точек $a_{i_1}, \dots, a_{i_k}$, B_i^φ есть часть B_φ из точек $b_{i_1}^\varphi, \dots, b_{i_k}^\varphi$, C_i^φ есть часть C^φ из точек $c_{i_1}^\varphi, \dots, c_{i_k}^\varphi$ ($1 \leq k \leq n$). Изображение C_i^φ называем порожденным изображениями A_i и B_i^φ , а сами изображения A_i и B_i^φ называем соответствующими C_i^φ . Через r_i^φ обозначим радиус ключевой окружности изображения C_i^φ .

Изображение C_i^φ назовем монотонным в точке φ , если существует такое $\Delta\varphi > 0$, что в промежутке от $\varphi - \Delta\varphi$ до $\varphi + \Delta\varphi$ радиус r_i^φ либо монотонно возрастает, либо монотонно убывает. В противном случае C_i^φ называем немонотонным в точке φ . Значения углов φ , в которых C_i^φ немонотонно, называем критическими для C_i^φ , взаиморасположение соответствующих C_i^φ изображений B_i^φ и A_i при критических углах тоже называем критическим.

Пусть угол φ_0 — искомый. Тогда ядро C^φ — немонотонное изображение в точке φ_0 , следовательно, угол φ_0 для C^φ — критический.

Идея дальнейших построений состоит в выделении на C^φ особых частей, минимальных по количеству точек, для которых φ_0 — тоже критический угол. Эти особые части назовем ключевыми изображениями и определим следующим образом.

Пусть для C_i^φ – части изображения C^φ – выполняется условие: угол φ_0 для C_i^φ – критический, и $r_i^{\varphi_0} = R_{\varphi_0}$. Пусть для C_j^φ – любой части изображения C_i^φ – это условие не выполняется. Тогда C_i^φ называем ключевым изображением, а соответствующие C_i^φ изображения A_i и B_i^φ тоже называем ключевыми.

Из последнего определения следует, что искомый угол может находиться только среди критических углов ключевых изображений. Все возможные ключевые изображения разобьем на три класса: 1) из двух точек, 2) из трех точек, 3) из четырех и более точек.

Рассмотрим последовательно эти три случая.

Первый случай. Ключевые изображения состоят из двух точек: $c_{i_1}^{\varphi_0}, c_{i_2}^{\varphi_0}$ из C^{φ_0} , a_{i_1}, a_{i_2} из $\tilde{A}$ и $b_{i_1}^{\varphi_0}, b_{i_2}^{\varphi_0}$ из B_{φ_0} . В этом случае отрезок $(c_{i_1}^{\varphi_0} c_{i_2}^{\varphi_0})$ должен быть диаметром ключевой окружности, и $(c_{i_1}^{\varphi_0} c_{i_2}^{\varphi_0})^2 = (a_{i_1} a_{i_2})^2 + (b_{i_1}^{\varphi_0} b_{i_2}^{\varphi_0})^2 - 2(a_{i_1} a_{i_2})(b_{i_1}^{\varphi_0} b_{i_2}^{\varphi_0}) \cos \alpha_{i_1 i_2}$, где $\alpha_{i_1 i_2}$ – угол в характеристическом изображении, определяемый соответствующими отрезками $(a_{i_1} a_{i_2})$ и $(b_{i_1}^{\varphi_0} b_{i_2}^{\varphi_0})$. Отсюда следует, что критические углы для рассматриваемых ключевых изображений определяются необходимым условием: $|\cos \alpha_{i_1 i_2}| = 1$. Это значит, что соответствующие отрезки $(a_{i_1} a_{i_2})$ и $(b_{i_1}^{\varphi_0} b_{i_2}^{\varphi_0})$ должны быть параллельными, то есть угол $\alpha_{i_1 i_2}$ должен быть равен нулю или π . Ясно, что φ_0 определяется равенством $\alpha_{i_1 i_2} = 0$. При этом R_{φ_0} равен половине модуля разности длин отрезков $(a_{i_1} a_{i_2})$ и $(b_{i_1}^{\varphi_0} b_{i_2}^{\varphi_0})$.

Обозначим через U_1 множество значений угла φ , при которых углы в характеристическом изображении, определяемые соответствующими отрезками $(a_i a_j)$ и $(b_i b_j)$ ($i, j = 1, \dots, n, i \neq j$) равны нулю.

Второй случай. Ключевые изображения состоят из трех точек: $c_{i_1}^{\varphi_0}, c_{i_2}^{\varphi_0}, c_{i_3}^{\varphi_0}$ из C^{φ_0} , $a_{i_1}, a_{i_2}, a_{i_3}$ из A и $b_{i_1}^{\varphi_0}, b_{i_2}^{\varphi_0}, b_{i_3}^{\varphi_0}$ из B_{φ_0} . Положим сначала, что никакие две из точек $c_{i_j}^{\varphi_0}$ ($j = 1, 2, 3$) не сливаются в одну, то есть $c_{i_1}^{\varphi_0} c_{i_2}^{\varphi_0} c_{i_3}^{\varphi_0}$ – остроугольный треугольник (рис. 27).

Длины отрезков $(a_{i_1} b_{i_1}^{\varphi_0}), (a_{i_2} b_{i_2}^{\varphi_0}), (a_{i_3} b_{i_3}^{\varphi_0})$ одинаковы и равны R_{φ_0} . Отрезки будем полагать направленными с направлением от точек a_{i_j} к $b_{i_j}^{\varphi_0}$ или наоборот. Пусть, для определенности, отрезки направлены от a_{i_j} к $b_{i_j}^{\varphi_0}$ ($j = 1, 2, 3$). Рассмотрим направленные прямые L_1, L_2, L_3 , на которых лежат эти отрезки,

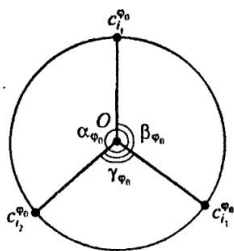


Рис. 27.

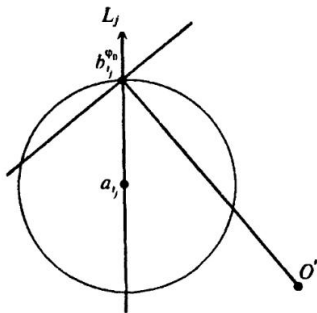


Рис. 28.

с направлениями для всех трех прямых, совпадающими с направлениями отрезков или противоположными им. Пусть, для определенности, направления прямых совпадают с направлениями отрезков $(a_{i_j} b_{i_j}^{\varphi_0})$ ($j = 1, 2, 3$), лежащих на них.

Лемма 3. Для того, чтобы взаиморасположение изображений из точек $a_{i_1}, a_{i_2}, a_{i_3}$ и $b_{i_1}^{\varphi_0}, b_{i_2}^{\varphi_0}, b_{i_3}^{\varphi_0}$ было критическим, необходимо, чтобы прямые L_1, L_2, L_3 пересекались в одной точке.

Доказательство. Заметим сначала, что в треугольнике $c_{i_1}^{\varphi} c_{i_2}^{\varphi} c_{i_3}^{\varphi}$ длины его сторон являются непрерывными функциями от угла φ . В точке φ_0 треугольник $c_{i_1}^{\varphi_0} c_{i_2}^{\varphi_0} c_{i_3}^{\varphi_0}$ по условию остроугольный, поэтому существует окрестность точки φ_0 такая, что при углах φ из этой окрестности он остается остроугольным.

Прямая L_j ($j = 1, 2, 3$), если смотреть по ее направлению, делит плоскость на правую полуплоскость и левую. Пусть O' — произвольная точка в правой полуплоскости (рис. 28). Правая часть перпендикуляра к отрезку $(O' b_{i_j}^{\varphi_0})$ в точке $b_{i_j}^{\varphi_0}$ находится вне круга радиуса R_{φ_0} с центром в точке a_{i_j} , для левой части перпендикуляра участок, примыкающий к $b_{i_j}^{\varphi_0}$, находится внутри круга. Поэтому существует угол $\Delta\varphi > 0$ такой, что поворот вправо точки $b_{i_j}^{\varphi_0}$ с центром вращения в O' на угол, не больший $\Delta\varphi$, удаляет точку $b_{i_j}^{\varphi_0}$ от a_{i_j} , а поворот влево на такой же угол — приближает. Аналогично для любого центра

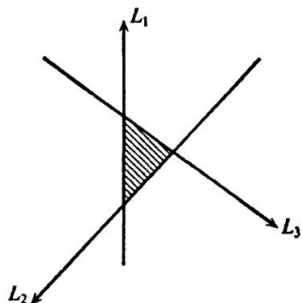


Рис. 29(а).

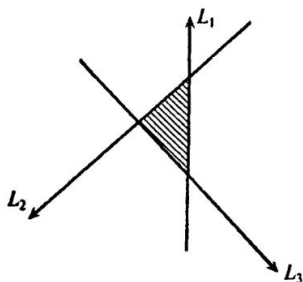


Рис. 29(б).

вращения O'' в левой полуплоскости поворот вправо приближает точку $b_{ij}^{\varphi_0}$, поворот влево – удаляет ее от a_{ij} .

Предположим, что L_1, L_2, L_3 не пересекаются в одной точке. Тогда они образуют треугольник, все внутренние точки которого являются точками их правых полуплоскостей (рис. 29(а)), или левых полуплоскостей (рис. 29(б)). В любом из этих случаев вращение B_{φ_0} с центром вращения в любой из внутренних точек этого треугольника в одну сторону удаляет одновременно все точки $b_{ij}^{\varphi_0}$ от соответствующих точек a_{ij} ($j = 1, 2, 3$), в другую – одновременно приближает их к ним. Длины отрезков $(a_{ij}, b_{ij}^{\varphi_0})$ ($j = 1, 2, 3$) и углы между ними при таком вращении являются непрерывными функциями от угла φ (в некоторой окрестности точки φ_0). Изображение из точек $c_{i_1}^{\varphi_0}, c_{i_2}^{\varphi_0}, c_{i_3}^{\varphi_0}$, следовательно, монотонно в точке φ_0 – пришли к противоречию. Лемма доказана.

Пересекающиеся в одной точке прямые L_1, L_2, L_3 назовем трехосником, сами прямые – осями, точку O_L пересечения – центром трехосника. На каждой оси часть ее от центра в направлении оси называем положительной полуосью, оставшуюся часть – отрицательной полуосью.

Все возможные варианты расположения точек $a_{i_1}, a_{i_2}, a_{i_3}$ и $b_{i_1}^{\varphi_0}, b_{i_2}^{\varphi_0}, b_{i_3}^{\varphi_0}$ на трехоснике можно представить четырьмя классами.

1. Класс S_1 : отрезки $(a_{ij}, b_{ij}^{\varphi_0})$ ($j = 1, 2, 3$) все одновременно лежат в положительных полуосях (рис. 30(а)), или все одновременно в отрицательных полуосях (рис. 30(б)).

2. Класс S_2 : один из отрезков $(a_{ij}, b_{ij}^{\varphi_0})$ ($j = 1, 2, 3$) лежит в положительной полуоси, два других – в отрицательных полуосях (рис. 31(а)), или, наоборот, один отрезок лежит в отрицательной полуоси, два других – в положительной (рис. 31(б)).

Классами S_1 и S_2 исчерпываются такие варианты размещения точек на осях, когда каждый из отрезков $(a_{ij}, b_{ij}^{\varphi_0})$ ($j = 1, 2, 3$) целиком лежит либо в положительной полуоси, либо в отрицательной.

3. Класс S_3 : для всех трех отрезков $(a_{ij}, b_{ij}^{\varphi_0})$ ($j = 1, 2, 3$) центр трехосника расположен между точками a_{ij} и $b_{ij}^{\varphi_0}$ (рис. 32).

4. Класс S_4 : у одного или двух отрезков из $(a_{ij}, b_{ij}^{\varphi_0})$ ($j = 1, 2, 3$) центр трехосника лежит на отрезке, остальные целиком лежат в отрицательной или положительной полуосях (два примера на рис. 33).

Классы S_1 и S_2 представляют варианты искомого взаиморасположения: нетрудно видеть, что вращение B_{φ_0} относительно центра трехосника в любую сторону (на угол, не превышающий некоторый $\Delta\varphi$) удаляет точки $b_{ij}^{\varphi_0}$ от соответствующих точек a_{ij} ($j = 1, 2, 3$). В S_3 взаиморасположение точек критическое, но не искомое: вращение B_{φ_0} относительно центра трехосника в любую сторону приближает соответствующие точки друг к другу. В классе S_4 только часть взаиморасположений искомые.

Итак, искомое взаиморасположение находится среди взаиморасположений изображений из точек a_{i1}, a_{i2}, a_{i3} и $b_{i1}^{\varphi_0}, b_{i2}^{\varphi_0}, b_{i3}^{\varphi_0}$, представленных классами $S_1 - S_4$. Следующий вопрос: как найти эти взаиморасположения при конкретных изображениях A и B ?

Условимся о следующих обозначениях. Отрезки $(a_{ij}, b_{ij}^{\varphi_0})$ ($j = 1, 2, 3$) будем представлять величиной R , где $|R| = R_{\varphi_0}$, причем будем считать R положительным, если направление отрезка совпадает с направлением оси, на которой он лежит, и отрицательным в противном случае.

Отрезки $(O_L a_{i1}), (O_L a_{i2}), (O_L a_{i3})$ считаем направленными от O_L к a_{ij} ($j = 1, 2, 3$), и представляем их величинами, соответственно, x, y, z . Длину отрезка $(O_L a_{i1})$ полагаем равной $|x|$, x полагаем положительным, если направление $(O_L a_{i1})$ совпадает

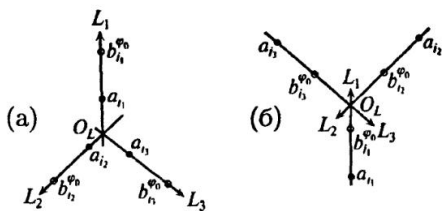


Рис. 30.

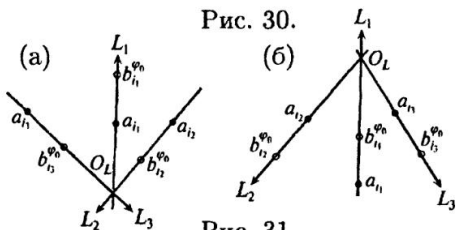


Рис. 31.

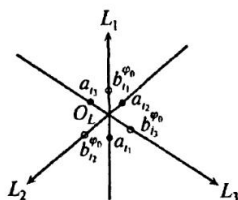


Рис. 32.

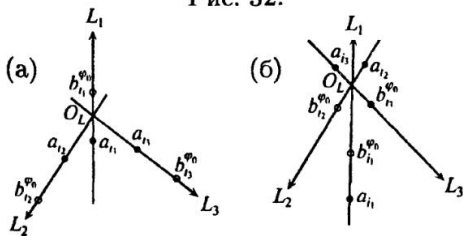


Рис. 33.

с направлением оси L_1 , и отрицательным в противном случае. Аналогичное полагаем и для 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 обозначим через, соответственно, α , β , γ . На рис. 34 для двух примеров показаны соответствующие обозначения.

Ясно, что задание конкретных значений для x , y , z , α , β , γ ,

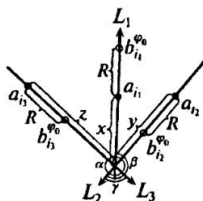
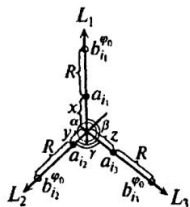


Рис. 34.

R определяет конкретное взаиморасположение точек a_{i1} , a_{i2} , a_{i3} и b_{i1} , b_{i2} , b_{i3} , а значит и конкретное значение угла φ между изображениями A и B .

Лемма 4. Все варианты взаиморасположения точек a_{i1} , a_{i2} , a_{i3} и b_{i1} , b_{i2} , b_{i3} , представленные в классах S_1 , S_2 , S_3 , S_4 , находятся среди решений следующей системы уравнений:

$$\begin{cases} (a_{i1}a_{i2})^2 = x^2 + y^2 - 2xy \cos \alpha \\ (b_{i1}b_{i2})^2 = (x+R)^2 + (y+R)^2 - 2(x+R)(y+R) \cos \alpha \\ (a_{i1}b_{i3})^2 = x^2 + z^2 - 2xz \cos \beta \\ (b_{i1}b_{i3})^2 = (x+R)^2 + (z+R)^2 - 2(x+R)(z+R) \cos \beta \\ (a_{i2}a_{i3})^2 = y^2 + z^2 - 2yz \cos \gamma \\ (b_{i2}b_{i3})^2 = (y+R)^2 + (z+R)^2 - 2(y+R)(z+R) \cos \gamma \\ \alpha + \beta + \gamma = 2\pi \end{cases}$$

Доказательство. Проводится непосредственной проверкой: написанием системы уравнений для каждого из вариантов, представленных классами $S_1 - S_4$.

Рассмотрим произвольную тройку точек $\tilde{a}$ из A и тройку соответствующих точек $\tilde{b}$ из B . Для изображений $\tilde{a}$ и $\tilde{b}$ при определении для них искомого взаиморасположения возможны только первый и второй случаи. Если имеет место первый случай, то определяемое им значение угла φ между изображениями A и B находится среди углов из множества U_1 . Во втором случае искомым угол находится среди набора $\tilde{u}$ углов, определяемых решениями системы уравнений из леммы 4. Обозначим

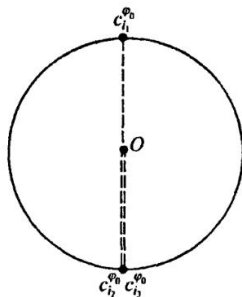
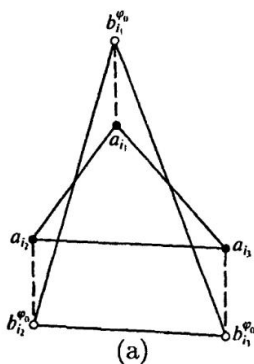
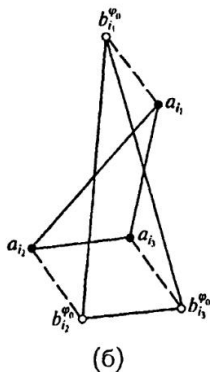


Рис. 35.



(a)



(б)

Рис. 36.

через U_2 объединение множеств $\bar{u}$, полученных для всех возможных троек точек из A и соответствующих троек точек из B .

Везде выше предполагалось, что точки $c_{i_1}^{\varphi_0}$, $c_{i_2}^{\varphi_0}$, $c_{i_3}^{\varphi_0}$ – разные. Рассмотрим теперь вырожденный случай, когда две из этих точек, для определенности, $c_{i_2}^{\varphi_0}$ и $c_{i_3}^{\varphi_0}$, сливаются в одну (рис. 35). Поскольку длина отрезка $(c_{i_2}^{\varphi_0} c_{i_3}^{\varphi_0})$ в этом случае равна нулю, то необходимым условием для этого является равенство по длине отрезков $(a_{i_2} a_{i_3})$ и $(b_{i_2} b_{i_3})$ и равенство нулю угла $\alpha_{i_2 i_3}$. Отсюда следует, что определяемые необходимым условием для всех возможных вырожденных случаев углы между изображениями A и B находятся среди углов множества U_1 .

На рисунке 36 представлены примеры, относящиеся к вырожденному случаю и выполнению для него необходимого условия: на (а) – искомое взаиморасположение, на (б) – не искомое. **Третий случай.** Ключевые изображения состоят из четырех и более точек. Рассмотрим четыре из них: $c_{i_1}^{\varphi_0}, c_{i_2}^{\varphi_0}, c_{i_3}^{\varphi_0}, c_{i_4}^{\varphi_0}$. Полагаем, что никакие две из них не слиты в одну. Из теоремы Птолемея следует, что для того, чтобы точки $c_{i_1}^{\varphi_0}, c_{i_2}^{\varphi_0}, c_{i_3}^{\varphi_0}, c_{i_4}^{\varphi_0}$ находились на одной окружности, необходимо и достаточно, чтобы выполнялось одно из трех равенств:

$$\pm(c_{i_1}^{\varphi_0} c_{i_2}^{\varphi_0})(c_{i_3}^{\varphi_0} c_{i_4}^{\varphi_0}) \pm (c_{i_1}^{\varphi_0} c_{i_3}^{\varphi_0})(c_{i_4}^{\varphi_0} c_{i_2}^{\varphi_0}) \pm (c_{i_1}^{\varphi_0} c_{i_4}^{\varphi_0})(c_{i_2}^{\varphi_0} c_{i_3}^{\varphi_0}) = 0.$$

Величина каждого из отрезков в этих уравнениях есть функция от угла φ между изображениями A и B . Следовательно, задача нахождения возможных критических углов для этого случая сводится к решению этих уравнений, неизвестным в которых является угол φ . Такие уравнения составляем для всех четверок точек $a_{i_1}, a_{i_2}, a_{i_3}, a_{i_4}$ из A и соответствующих четверок точек $b_{i_1}, b_{i_2}, b_{i_3}, b_{i_4}$ из B . Множество углов φ , являющихся решениями получаемых уравнений, обозначим через U_3 .

Обозначим через U объединение множеств U_1, U_2, U_3 . Предшествующими рассуждениями, включающими леммы 3 и 4, доказано следующее утверждение.

Теорема 2. *Искомый угол φ_0 находится среди углов множества U .*

Итак, пусть определен угол φ_0 , следовательно, найдено значение $R_{\varphi_0} = R_{\psi}^{\min}$. Обозначим через $R'(A, B) = \min_{\psi} R_{\psi}^{\min}$, где минимум берется на множестве величин $R_{\psi}^{\min}$, полученных при всех возможных биекциях на множествах точек изображений A и B .

Обозначим через $\hat{B}$ изображение, полученное из B преобразованием симметрии относительно какой-либо прямой, и пусть $R''(A, B) = R'(A, \hat{B})$.

Наконец, через $R(A, B)$ обозначим $\min(R'(A, B), R''(A, B))$. Величина $R(A, B)$ определяет наименьшую взаимоудаленность точек изображений A и B при всех возможных сопоставлениях

этих точек друг другу и при всех возможных изометрических преобразованиях этих изображений.

§3. Преобразования подобия изображений

Пусть изображение A состоит из точек $a_1, \dots, a_n$ ($n \geq 1$), изображение B — из точек $b_1, \dots, b_n$. Сразу исключаем из рассмотрения тривиальный случай, когда A и B переводятся друг в друга преобразованиями подобия. Следовательно, $n \geq 3$.

Рассмотрим некоторое взаимно однозначное соответствие ψ между точками изображений A и B . Без ограничения общности можем полагать, что этим соответствием точке a_i сопоставляется точка b_i ($i = 1, \dots, n$).

Назовем сокращенными преобразованиями подобия преобразования параллельного переноса, поворота, изменения масштаба (с коэффициентом подобия $k > 0$) и их комбинации, то есть преобразования подобия без симметрии относительно прямой. Пусть A^* и B^* — множества всех изображений, получаемых из A и B сокращенными преобразованиями подобия. Если $A' \in A^*$, то через a'_i обозначаем точку из A' , в которую переходит при соответствующем преобразовании точка a_i из A ($i = 1, \dots, n$). Аналогично трактуется обозначение точки b'_i из B' , где $B' \in B^*$. Точки a'_i и b'_i называем соответствующими, соответствующими называем и отрезки $(a'_i a'_j)$ и $(b'_i b'_j)$ ($i, j = 1, \dots, n, i \neq j$).

Рассмотрим бинарное отношение Q на декартовом произведении $A^* \times B^*$: пары (A_1, B_1) и (A_2, B_2) из $A^* \times B^*$ находятся в отношении Q , если пара (A_1, B_1) как целое переводится сокращенным преобразованием подобия в пару (A_2, B_2) . По этому отношению $A^* \times B^*$ разбивается на множество классов эквивалентности. Изображения во всех парах из одного такого класса можно, очевидно, трактовать как расположенные одинаково по отношению друг к другу. В качестве числовой характеристики взаимоудаленности точек изображений в некоторой паре (A', B') возьмем длину наибольшего из отрезков $(a'_i b'_i)$ ($i = 1, \dots, n$), отнесенную к длине заранее выделенного отрезка между парой точек на изображении либо A' , либо B' . Будем для определенности полагать, что этот выделенный от-

резок обозначается двумя точками a_u и a_v на A (и, следовательно, точками a'_u и a'_v на A'). Содержательно это означает, что мы берем наибольший по длине отрезок между соответствующими точками пары изображений (A', B') , причем в качестве масштабной единицы при измерении его длины используем отрезок $(a'_u a'_v)$. Очевидно, что эта характеристика одинакова для всех пар изображений из одного класса эквивалентности разбиения Q^* .

Введем понятие угла φ между изображениями в паре (A, B) . Для этой цели зафиксируем в изображениях A и B пары отрезков $(a_p a_q)$ ($p \neq q$), и соответственно $(b_p b_q)$. В качестве угла φ выберем угол, образованный этими отрезками (или их продолжениями). Для произвольной пары изображений (A', B') из $A^* \times B^*$ угол φ между ними есть угол между соответствующими отрезками $(a'_p a'_q)$ и $(b'_p b'_q)$ или их продолжениями. Без ограничения общности можно полагать, что у исходных изображений A и B отрезки $(a_p a_q)$ и $(b_p b_q)$ параллельны и угол φ равен нулю, а для всех пар из $A^* \times B^*$ этот угол находится в промежутке от 0 до 2π .

Отметим, что для всех пар изображений из одного класса эквивалентности в Q^* угол φ между изображениями один и тот же.

Разбиению Q^* можно дать наглядную интерпретацию, на которую в дальнейших рассуждениях удобно ориентироваться. Возьмем в каждом классе эквивалентности пару (очевидно, единственную), у которой первым элементом является исходное изображение A . Тогда получающееся множество пар представится как «неподвижное» изображение A и всеми возможными способами (конечно, из числа сокращенных преобразований подобия) трансформируемое изображение B . Каждое такое трансформированное B' охарактеризуем в отношении исходного изображения B парой параметров (φ, k) : здесь φ – угол между B' и A , и k – коэффициент подобия (то есть отношение любой пары соответствующих отрезков из B' и B). При этом B' – не единственное изображение с фиксированной парой (φ, k) параметров. Однако если обозначить через $\{B\}_{\varphi, k}$ множество всех таких изображений, то состоит оно, нетрудно видеть, из изображений, переводимых друг в друга преобразо-

ваниями параллельного переноса. Как следует из результатов §1, существует и единственно изображение $B'' \in \{B\}_{\varphi, k}$, на котором достигается минимум взаиморасстояния изображения A и изображений из $\{B\}_{\varphi, k}$. Этот минимум – обозначим его через $r(\varphi, k)$ – представляется тем самым как функция от пары (φ, k) параметров. Далее вопрос сводится к тому, чтобы определить минимум величин $r(\varphi, k)$ при всех возможных углах φ ($0 \leq \varphi < 2\pi$) и всех коэффициентах подобия k ($k > 0$), и найти расположение соответствующим образом трансформированного изображения B относительно изображения A . Пару (φ_0, k_0) , на которой достигается минимум для $r(\varphi, k)$, будем называть искомой, состоит она из искомого угла φ_0 и искомого коэффициента подобия k_0 . Величину $r(\varphi_0, k_0)$ будем обозначать через r_0^ψ и называть минимальной взаиморасстоянием точек изображений A и B (при заданной биекции ψ), изображения во всех парах (A', B') из $A^* \times B^*$ со взаиморасстоянием, равной r_0^ψ (искомых парах), будем называть взаиморасположенными искомым образом. При «фиксированном» изображении A это значит, что нужно найти $B_0 \in B^*$ такое, что в паре (A, B_0) изображения расположены искомым образом.

Возьмем пару (A, B') , где $B' \in B^*$, и рассмотрим длину отрезка $(c_i c_j)$ из характеристического изображения C этой пары: $(c_i c_j)^2 = (a_i a_j)^2 + (b'_i b'_j)^2 - 2(a_i a_j)(b'_i b'_j) \cos \alpha_{ij}$. Поскольку угол α_{ij} зависит от угла φ между изображениями A и B' и величина $(b'_i b'_j)$ зависит от коэффициента подобия k , то длина отрезка $(c_i c_j)$ есть функция от величин пары (φ, k) . Длина отрезка $(c_i c_j)$ равна нулю, то есть точки c_i и c_j слиты в одну, только при равенстве отрезков $(a_i a_j)$ и $(b'_i b'_j)$ и их параллельности, поскольку угол α_{ij} тоже должен быть равен нулю. Следовательно, существует лишь конечное множество пар (φ, k) , при которых это может иметь место. Обозначим это множество через $\hat{U}_1$. В дальнейших рассуждениях исходим из предположения, что уже никакие две точки ядра не слиты в одну. Положим, что мы имеем искомую пару (A, B_0) изображений. Далее априори возможны три случая: на ключевой окружности находятся соответственно две, три и больше трех точек

Первый случай. Этот случай при подобных преобразованиях невозможен. Действительно, положим, что это не так. Тог-

да точки c_{i_1} и c_{i_2} ($i_1, i_2 \in (1, \dots, n)$) на ключевой окружности должны быть концами ее диаметра. Но тогда отрезки $(a_{i_1} a_{i_2})$ и $(b_{i_1}^0 b_{i_2}^0)$ должны быть параллельными, и угол $\alpha_{i_1 i_2}$ в характеристическом изображении, определяемый соответствующими отрезками $(a_{i_1} a_{i_2})$ и $(b_{i_1}^0 b_{i_2}^0)$, должен быть равен нулю, ибо в противном случае всегда возможен поворот изображения B_0 , уменьшающий длину отрезка $(c_{i_1} c_{i_2})$. При $\alpha_{i_1 i_2} = 0$ длина отрезка $(c_{i_1} c_{i_2})$ равна модулю разности длин отрезков $(a_{i_1} a_{i_2})$ и $(b_{i_1}^0 b_{i_2}^0)$. Это значит, что $(c_{i_1} c_{i_2})$ всегда можно уменьшить увеличением или уменьшением изображения B_0 в размерах, что противоречит предположению о том, что пара (A, B_0) – искомая.

Второй случай. Пусть 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 либо остроугольным, либо прямоугольным.

Пусть для начала этот треугольник остроугольный. Будем рассматривать направленные прямые L_1, L_2, L_3 , на которых лежат отрезки соответственно $(a_{i_1} b_{i_1}^0)$, $(a_{i_2} b_{i_2}^0)$, $(a_{i_3} b_{i_3}^0)$. Согласно лемме 3 прямые L_1, L_2, L_3 пересекаются в одной точке (точке O_L). Далее рассматриваем трехосник с обозначениями $x, y, z, \alpha, \beta, \gamma, R$ теми же, что и в §2.

Все варианты взаиморасположения точек $a_{i_1}, a_{i_2}, a_{i_3}$ и $b_{i_1}^0, b_{i_2}^0, b_{i_3}^0$ на трехоснике находятся среди решений системы уравнений из леммы 4. Если стороны треугольников $a_{i_1} a_{i_2} a_{i_3}$ и $b_{i_1}^0 b_{i_2}^0 b_{i_3}^0$ известны, то в этой системе из семи уравнений семь неизвестных: $x, y, z, \alpha, \beta, \gamma, R$. Однако теперь известны только стороны треугольника $b_{i_1} b_{i_2} b_{i_3}$, а стороны треугольника $b_{i_1}^0 b_{i_2}^0 b_{i_3}^0$ получаются из сторон треугольника $b_{i_1} b_{i_2} b_{i_3}$ умножением на коэффициент подобия k . Следовательно, возникает дополнительное восьмое неизвестное k .

Введем в рассмотрение направленные прямые P_1, P_2, P_3 , проходящие через точки $b_{i_1}^0, b_{i_2}^0, b_{i_3}^0$ соответственно и перпендикулярные прямым L_1, L_2, L_3 . Направление прямой P_j ($j = 1, 2, 3$) определяется тем условием, что поворот ее на 90° против часовой стрелки с центром вращения в точке $b_{i_j}^0$ должен совмещать ее с прямой L_j .

Лемма 5. Прямые P_1, P_2, P_3 должны пересекаться в одной точке.

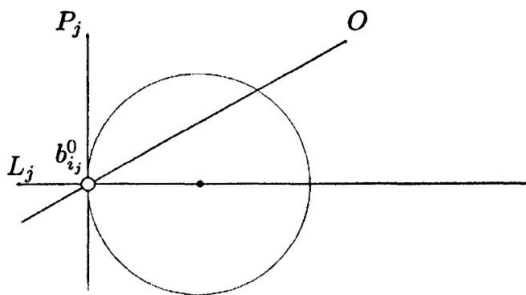


Рис. 37.

Доказательство Прямая P_j ($j = 1, 2, 3$) делит плоскость на две полуплоскости: правую, если смотреть по ее направлению (рис. 37), и левую. Пусть O – произвольная точка в правой полуплоскости. Тогда часть отрезка $(Ob_{i,j}^0)$, примыкающая к точке $b_{i,j}^0$, находится внутри круга радиуса R с центром в точке $a_{i,j}$, а продолжение отрезка $(Ob_{i,j}^0)$, примыкающее к точке $b_{i,j}^0$, находится вне этого круга. Поэтому гомотетия с центром в точке O и с коэффициентом $k > 1$ удалит точку $b_{i,j}^0$ от точки $a_{i,j}$, а с коэффициентом $0 < k < 1$ – приблизит.

Аналогично гомотетия с любым центром в левой полуплоскости и с коэффициентом $k > 1$ будет приближать $b_{i,j}^0$ к $a_{i,j}$, а с коэффициентом $0 < k < 1$ – удалять.

Предположим теперь, что прямые P_1, P_2, P_3 не пересекаются в одной точке. Тогда они образуют треугольник, все внутренние точки которого являются точками их правых полуплоскостей (рис. 38(а)), или левых полуплоскостей (рис. 38(б)). В любом из этих случаев гомотетия с центром в любой из внутренних точек этого треугольника с одним коэффициентом удаляет все точки $b_{i,j}^0$ от соответствующих точек $a_{i,j}$, а с другим коэффициентом – приближает их к ним. Тем самым приходим к противоречию с предположением о том, что взаиморасположение изображений A и B_0 – искомого. Лемма доказана.

Точку пересечения прямых P_1, P_2, P_3 обозначим через O_P .

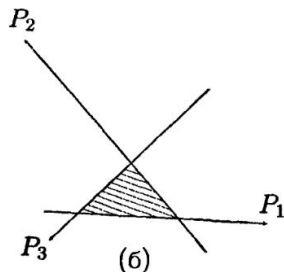
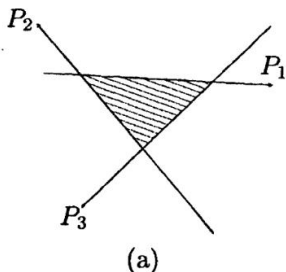


Рис. 38.

Лемма 6. Точки $b_{i_1}^0, b_{i_2}^0, b_{i_3}^0, O_L, O_P$ коцикличны.

Доказательство Углы $O_L b_{i_1}^0 O_P, O_L b_{i_2}^0 O_P, O_L b_{i_3}^0 O_P$ прямые и опираются на один и тот же отрезок $(O_L O_P)$. Следовательно, $(O_L O_P)$ является диаметром общей для треугольников $O_L b_{i_1}^0 O_P, O_L b_{i_2}^0 O_P, O_L b_{i_3}^0 O_P$ описанной окружности. Лемма доказана.

Лемма 7. Взаиморасположение точек $a_{i_1}, a_{i_2}, a_{i_3}$ и $b_{i_1}^0, b_{i_2}^0, b_{i_3}^0$, определяющее искомое взаиморасположение изображений A и B_0 , находится среди решений четырех систем уравнений, получаемых поочередным присоединением к системе (1) уравнений (2), (3), (4), (5):

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

$$\pm (b_{i_1} b_{i_2})|z + R| \pm (b_{i_1} b_{i_3})|y + R| \pm (b_{i_2} b_{i_3})|x + R| = 0 \quad (2)$$

$$x + R = 0 \quad (3)$$

$$y + R = 0 \quad (4)$$

$$z + R = 0 \quad (5)$$

Доказательство Система (1) совпадает с ранее выписанной системой уравнений из леммы 4. Отмечалось, что в ней восемь неизвестных: $x, y, z, \alpha, \beta, \gamma, R, k$. Поскольку в (1) семь уравнений, то нужно добавить к ней еще одно уравнение.

Уравнение (2) возникает из тех соображений, что поскольку, согласно лемме 6, точки $b_{i_1}^0, b_{i_2}^0, b_{i_3}^0, O_L$ коцикличны, необходимо и достаточно (теорема Птолемея), чтобы выполнялось одно из трех равенств:

$$\pm(b_{i_1}^0 b_{i_2}^0)(b_{i_3}^0 O_L) \pm (b_{i_1}^0 b_{i_3}^0)(b_{i_2}^0 O_L) \pm (b_{i_2}^0 b_{i_3}^0)(b_{i_1}^0 O_L) = 0$$

Однако условие, проистекающее из теоремы Птолемея, не работает, если точка O_L совпадает с какой-либо из вершин треугольника $b_{i_1}^0 b_{i_2}^0 b_{i_3}^0$. Но такое совпадение, например, O_L с вершиной $b_{i_1}^0$, означает дополнительное уравнение $x + R = 0$, то есть уравнение (3). Аналогично из предположений о совпадении O_L с вершинами $b_{i_2}^0$ и $b_{i_3}^0$ получаются уравнения (4) и (5). Лемма доказана.

Рассмотрим все тройки точек $a_{i_1}, a_{i_2}, a_{i_3}$ из изображения A и тройки соответствующих точек $b_{j_1}, b_{j_2}, b_{j_3}$ из изображения B . Для каждой такой тройки решаются системы уравнений из леммы 7. Каждое решение определяет пару (φ, k) . Множество всех таким образом получаемых пар (φ, k) обозначим через $\hat{U}_2$.

Везде выше предполагалось, что треугольник $c_{i_1} c_{i_2} c_{i_3}$ остроугольный. Положим теперь, что он прямоугольный и покажем, что взаиморасположение A и B в этом случае не может быть искомым. Пусть $(c_{i_1} c_{i_2})$ – гипотенуза треугольника и, следовательно, диаметр ключевой окружности. Далее возможны два варианта:

1. Отрезки $(a_{i_1} a_{i_2})$ и $(b_{i_1}^0 b_{i_2}^0)$ не параллельны. Пример такой ситуации представлен на рис. 39. Прямые L_1, L_2, L_3 не пересекаются в одной точке. Теми же построениями, которые использованы в §2 для доказательства пересечения прямых L_1, L_2, L_3 в одной точке, нетрудно показать, что в этом случае всегда существует точка (например, точка O на рис. 39), поворот относительно которой изображения B_0 приближает точки $b_{i_j}^0$ к точкам a_{i_j} ($j = 1, 2, 3$).

2. Отрезки $(a_{i_1} a_{i_2})$ и $(b_{i_1}^0 b_{i_2}^0)$ параллельны (пример на рис. 40). В этом случае прямые P_1, P_2, P_3 не пересекаются в одной точке

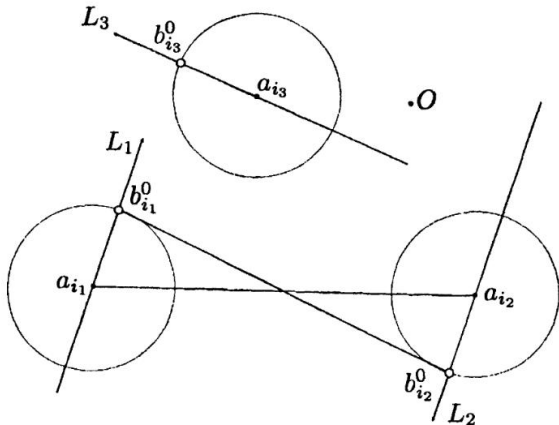


Рис. 39.

и потому существуют точки (например, точка O на рис. 40), которые могут служить центром гомотетии, посредством которой можно приблизить точки b_{ij}^0 к точкам a_{ij} ($j = 1, 2, 3$).

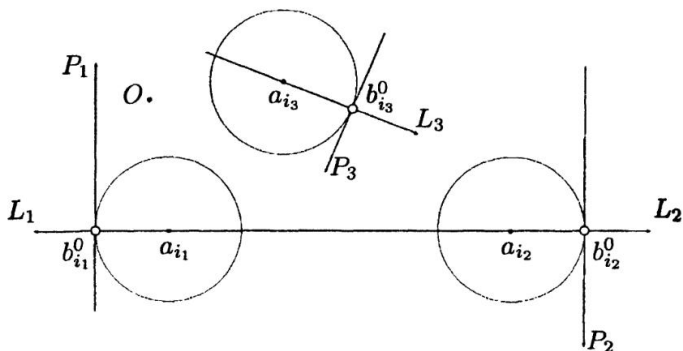


Рис. 40.

Отметим, что если точки b_{i1}, b_{i2}, b_{i3} лежат на одной прямой, то решение для этого случая должно принадлежать множеству $\hat{U}_1$.

Третий случай. Осталось рассмотреть случай, когда на

ключевой окружности находятся четыре и более точек ядра. Попарные расстояния между точками на ключевой окружности известны и являются при заданных A и B функциями от φ и k . Известна формула, представляющая радиус описанной окружности треугольника через длины его сторон. Приравняем попарно радиусы описанных окружностей треугольников $c_{i_1} c_{i_2} c_{i_3}$ и $c_{i_1} c_{i_2} c_{i_4}$, и треугольников $c_{i_1} c_{i_3} c_{i_4}$ и $c_{i_2} c_{i_3} c_{i_4}$. Это даст систему из двух уравнений, неизвестными в которых будут φ и k . Множество всех пар (φ, k) , полученных в качестве решений систем уравнений, составленных для всех возможных четверок точек $a_{j_1}, a_{j_2}, a_{j_3}, a_{j_4}$ и соответствующих четверок точек $b_{j_1}, b_{j_2}, b_{j_3}, b_{j_4}$, обозначим через $\hat{U}_3$.

Обозначим через $\hat{U}$ объединение множеств $\hat{U}_1, \hat{U}_2, \hat{U}_3$.

Теорема 3. *Искомая пара (φ_0, k_0) находится среди элементов множества $\hat{U}$.*

Доказательство Следует из того, что при искомом взаиморасположении изображений A и B на ключевой окружности находятся либо три, либо четыре и более точек. Все пары (φ, k) , этому соответствующие, принадлежат множеству $\hat{U}$ (подмножества $\hat{U}_2$ и $\hat{U}_3$). Там же находятся и пары (φ, k) , соответствующие вырожденным случаям, когда две и более точек ядра совпадают (подмножество $\hat{U}_1$). Теорема доказана.

Величина r_0^ψ получена при заданной биекции ψ и сокращенных преобразованиях подобия. Наименьшую из величин r_0^ψ , полученную при всех возможных биекциях ψ , обозначим через $r_0(A, B)$. Обозначим через $\bar{B}$ изображение, полученное из B преобразованием симметрии относительно какой-либо прямой, и через $R_0(A, B)$ — меньшую из величин $r_0(A, B)$ и $r_0(A, \bar{B})$. Значением $R_0(A, B)$ и определяется наименьшая взаимоудаленность изображений A и B при подобных преобразованиях в целом.

Для случая, когда на ключевой окружности находятся три точки, искомое взаиморасположение можно найти геометрическими построениями. Ниже описываются основные моменты таких построений.

Треугольники $a_{i_1} a_{i_2} a_{i_3}$ и $b_{i_1}^0 b_{i_2}^0 b_{i_3}^0$, полагаем, расположены искомым образом, причем точки $b_{i_1}^0, b_{i_2}^0, b_{i_3}^0$ не лежат на одной прямой и треугольник $c_{i_1} c_{i_2} c_{i_3}$ из точек ядра остроугольный.

Леммы 5 и 6 определяют условия, необходимые для искомого взаиморасположения. На рис. 41–43 представлены примеры расположения треугольников, удовлетворяющие этим условиям.

Леммой 6 углы α, β, γ между осями трехосника определены. Действительно, это углы, дополнительные к углам при вершинах треугольника $b_{i_1} b_{i_2} b_{i_3}$. Тем самым и трехосник можно считать известным с точностью до параллельных переносов и поворотов. Покажем теперь, что на осях L_1, L_2, L_3 трехосника можно разместить точки соответственно $a_{i_1}, a_{i_2}, a_{i_3}$. Разместим точки a_{i_2} и a_{i_3} на осях L_2 и L_3 . Существует четыре варианта такого размещения: обе точки в положительной или отрицательной полуосях (рис. 44), или одна точка в положительной полуоси, другая – в отрицательной (рис. 45). Дополнительными (вырожденными) вариантами можно назвать совпадение какой-либо из точек a_{i_2} и a_{i_3} с центром O_L трехосника. Перемещая отрезок $(a_{i_2} a_{i_3})$ так, чтобы точки a_{i_2} и a_{i_3} перемещались по осям L_2 и L_3 , можно последовательно получить все эти варианты. Если отрезку $(a_{i_2} a_{i_3})$ придать направленность, например, от a_{i_2} к a_{i_3} , то из рисунков 44 и 45 видно, что направленный отрезок $(a_{i_2} a_{i_3})$ совершает при таком перемещении полный оборот. Точка a_{i_1} при фиксированном положении отрезка $(a_{i_2} a_{i_3})$ может занимать два положения a'_{i_1} и a''_{i_1} , симметрично относительно прямой, на которой лежит $(a_{i_2} a_{i_3})$ (рис. 44). Ясно, что эти два положения для a_{i_1} соответствуют изображениям A и $\bar{A}$, получаемым друг из друга преобразованием симметрии. При вращении отрезка $(a_{i_2} a_{i_3})$ каждая из точек a'_{i_1} и a''_{i_1} описывает замкнутую кривую, причем каждая из этих кривых симметрична относительно центра O_L . Значит, каждая из этих кривых пересекает ось L_1 , что и дает искомые положения для точки a_{i_1} .

Описанное относится к вопросу о существовании нужного размещения точек $a_{i_1}, a_{i_2}, a_{i_3}$ на трехоснике. Собственно же построение нужного размещения может осуществляться следующим образом. На сторонах треугольника $a_{i_1} a_{i_2} a_{i_3}$ как на хордах строим окружности такие, что стороны треугольни-

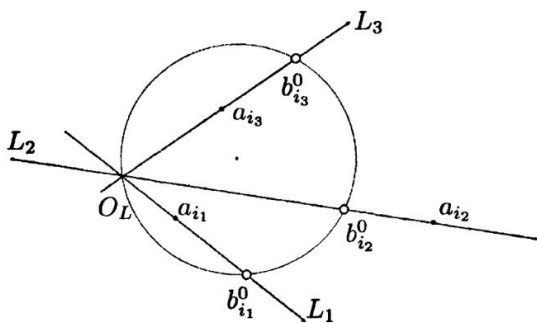


Рис. 41.

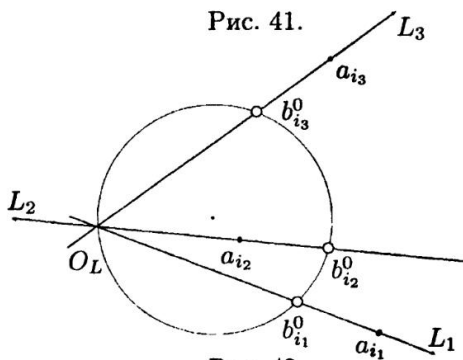


Рис. 42.

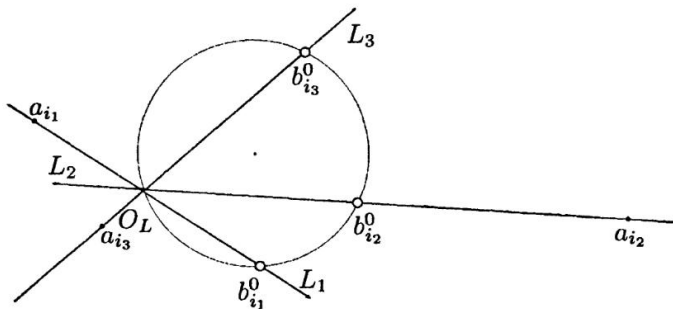


Рис. 43.

ка стягивают дуги, равные соответствующим углам треугольника $b_{i1}b_{i2}b_{i3}$ или дополнительные к ним. Так, например, стороне $(a_{i1}a_{i2})$ соответствует угол при вершине b_{i3} треугольни-

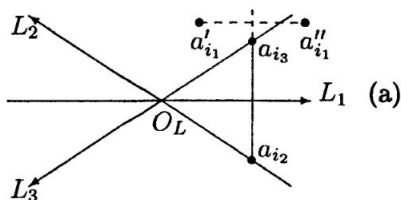


Рис. 44.

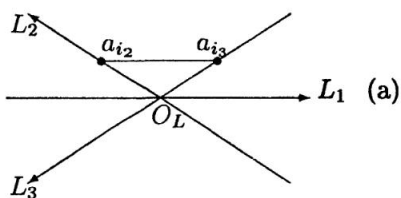
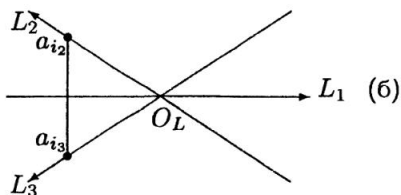
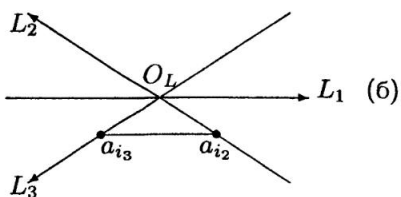


Рис. 45.



ка $b_{i_1} b_{i_2} b_{i_3}$. Для каждой стороны треугольника $a_{i_1} a_{i_2} a_{i_3}$ таких окружностей будет две (пример на рис. 46). Искомый центр O_L трехосника определится пересечением в одной точке трех окружностей, соответствующих трем сторонам треугольника $a_{i_1} a_{i_2} a_{i_3}$.

Итак, полагаем теперь известным трехосник и размещение

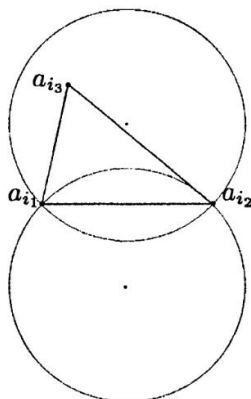


Рис. 46.

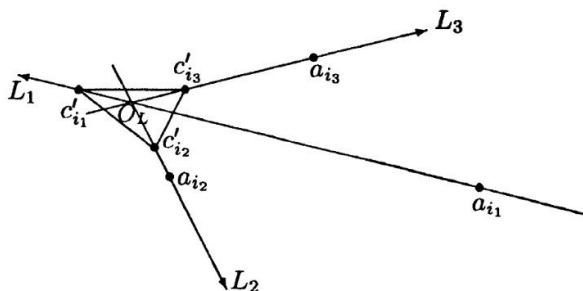


Рис. 47.

точек a_{i1}, a_{i2}, a_{i3} на нем. Если на трехоснике отложить по осям положительные равные отрезки, то у возникающего треугольника (треугольник $c'_{i1}c'_{i2}c'_{i3}$ на рис. 47) стороны параллельны сторонам треугольника с вершинами на ключевой окружности. Следовательно, можно считать известными (с точностью до параллельных переносов) прямые, на которых лежат стороны ключевого треугольника.

Будем далее ориентироваться на пример, изображенный на рис. 48. Что в сравнении с рисунком 47 нужно знать, чтобы получить рис. 48? Достаточно, очевидно, иметь угол поворота треугольника $b_{i1}b_{i2}b_{i3}$ относительно $a_{i1}a_{i2}a_{i3}$ и коэффициент

подобия, позволяющий из $b_{i_1} b_{i_2} b_{i_3}$ получить $b_{i_1}^0 b_{i_2}^0 b_{i_3}^0$. Для знания угла поворота достаточно иметь любой из углов между соответствующими сторонами треугольников $b_{i_1}^0 b_{i_2}^0 b_{i_3}^0$ и $a_{i_1} a_{i_2} a_{i_3}$, а для знания коэффициента подобия – любое из отношений соответствующих сторон этих треугольников.

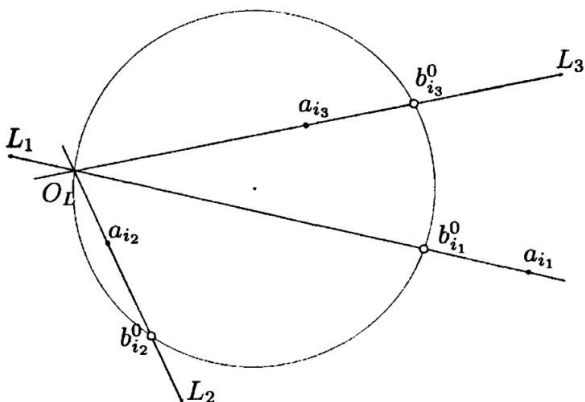


Рис. 48.

Проведем некоторые преобразования рисунка 48. Параллельным переносом треугольника $b_{i_1}^0 b_{i_2}^0 b_{i_3}^0$ какую-либо из точек $b_{i_j}^0$ совместим с точкой a_{i_j} ($j = 1, 2, 3$). Для определенности положим, что точка $b_{i_3}^0$ совмещается с точкой a_{i_3} (рис. 49). Точки $b_{i_1}^0$ и $b_{i_2}^0$ должны очутиться на прямых K_1 и K_2 . Эти прямые проводятся через точки a_{i_1} и a_{i_2} и параллельны сторонам $(c'_{i_2} c'_{i_3})$ и $(c'_{i_1} c'_{i_3})$ треугольника $c'_{i_1} c'_{i_2} c'_{i_3}$ и, следовательно, могут быть построены заранее. Достаточно построить положение какой-либо из точек $b_{i_1}^0$ или $b_{i_2}^0$ на прямых K_1 и K_2 – это, очевидно, определит и искомый угол, и коэффициент подобия.

Предположим, что построен треугольник $a_{i_3} a'_{i_2} b'_{i_2}$, равный треугольнику $a_{i_3} a_{i_2} b_{i_2}^0$ так, чтобы прямые, на которых лежат отрезки $(a_{i_3} b'_{i_2})$ и $(a_{i_3} b_{i_1}^0)$, совместились (рис. 49). Прямую, на которой лежит отрезок $(a'_{i_2} b'_{i_2})$, обозначим через K'_2 . Такое построение наглядно можно представить как поворот вокруг точки a_{i_3} на рис. 49 комплекса из треугольника $a_{i_3} a_{i_2} b_{i_2}^0$ и прямой K_2 до пересечения точкой $b_{i_2}^0$ отрезка $a_{i_3} b_{i_1}^0$ или его продолже-

ния. Пересечение прямых K_1 и K'_2 обозначим через x . Через a_{i_3} проводим прямую, параллельную K_1 до пересечения с K'_2 . Точку пересечения обозначим через y .

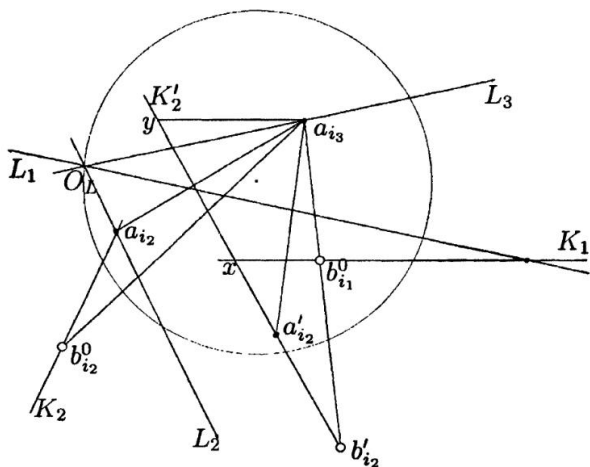


Рис. 49.

Однако в этих рассуждениях пока можно считать известными лишь точку a'_{i_2} , прямую K'_2 , и кроме того, точку x и отрезок $(a_{i_3}y)$, которые находятся построением. Искомое положение треугольника $b^0_{i_1} b^0_{i_2} b^0_{i_3}$ относительно $a_{i_1} a_{i_2} a_{i_3}$ на рис. 49 определится, если будет определен отрезок $(b^0_{i_1}x)$, поскольку этим, очевидно, определится и положение точки $b^0_{i_1}$ на прямой K_1 .

Треугольники $a_{i_3} y b'_{i_2}$ и $b^0_{i_1} x b'_{i_2}$ подобны, и коэффициент подобия определяется известным отношением сторон $(a_{i_3} b'_{i_2})$ и $(b^0_{i_1} b'_{i_2})$ (на рис. 49 отрезок $(b^0_{i_1} b'_{i_2})$ равен модулю разности длин отрезков $(a_{i_3} b^0_{i_1})$ и $(a_{i_3} b'_{i_2})$). Так как отрезок $(a_{i_3} y)$ известен из построения, то отсюда и находим длину искомого отрезка $(b^0_{i_1} x)$.

§4. Аффинные преобразования изображений

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 — взаимно однозначное 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 взаимно однозначных 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 другой – 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. Из полученных ранее результатов следуют некоторые необходимые условия, которым должны удовлетворять изображения в паре (A, B_0) . Они коротко (для аффинного случая) состоят в следующем. Пусть биекцией ψ точке a_{i_k} из A сопоставляется точка b'_{j_k} из B' ($k = 1, \dots, n$). Возьмем на плоскости произвольную точку O и параллельными переносами отрезков $(a_{i_k} b'_{j_k})$ совместим точку a_{i_k} с точкой O . Точку, в которую перейдет при этом b'_{j_k} , обозначим через $c_{i_k j_k}$ и назовем порожденной парой соответствующих точек a_{i_k} и b'_{j_k} . Изображение, состоящее из точки O – центра, и точек $c_{i_k j_k}$ ($k = 1, \dots, n$) – точек ядра, называем характеристическим изображением C для пары (A, B') . При таком построении некоторые из точек ядра могут совпасть (слиться). Такие точки называем кратными (l -кратными, если слились l точек), и сохраняем в их обозначениях символы слившихся точек. Окружность наименьшего по радиусу круга, включающего все точки ядра, называем ключевой. Показано, что для пары (A, B_0) центр характеристического изображения с необходимостью должен совпадать с центром ключевой окружности, и на ключевой окружности (ненулевого радиуса) должны находиться (при преобразованиях подобия) не менее трех точек ядра.

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 существен-

ное 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} .

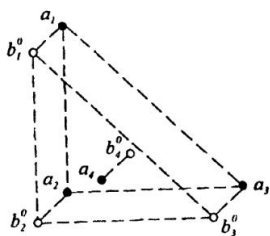


Рис. 50.

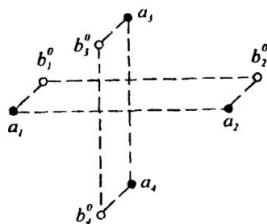


Рис. 51.

На рис. 50 и 51 представлены примеры изображений B_0 (из четырех точек), согласованных с A (и искомых для A). Точке a_i на этих рисунках соответствует точка b_i^0 .

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 .

Доказательство. На ключевой окружности не может быть меньше четырех точек ядра. Действительно, пусть это не так и на окружности только три точки $c_{i_1 j_1}, c_{i_2 j_2}, c_{i_3 j_3}$ ($i_1, j_1, i_2, j_2, i_3, j_3 \in (1, \dots, n)$). Это значит, что точки $b_{j_1}^0, b_{j_2}^0, b_{j_3}^0$ находятся от соответствующих точек $a_{i_1}, a_{i_2}, a_{i_3}$ на одинаковых расстояниях, равных $R_A(B)$. Однако ε -аффинным преобразованием изображения B_0 вершины треугольника $b_{j_1}^0 b_{j_2}^0 b_{j_3}^0$ всегда можно придвинуть к соответствующим вершинам треугольника $a_{i_1} a_{i_2} a_{i_3}$, и, тем самым, уменьшить $R_A(B)$.

Положим теперь, что на окружности четыре точки $c_{i_1 j_1}, c_{i_2 j_2}, c_{i_3 j_3}, c_{i_4 j_4}$ ($i_1, j_1, i_2, j_2, i_3, j_3, i_4, j_4 \in (1, \dots, n)$), и все они попарно не совпадают. Из определения характеристического изображения следует, что при аффинных преобразованиях пары (A, B_0) как целого этим же преобразованиям подвергается и характеристическое изображение C этой пары, то есть если (A', B'_0) получено из (A, B_0) аффинным преобразованием, то изображение C' , полученное из C этим же преобразованием, будет характеристическим для (A', B'_0) . Следовательно, точки $c'_{i_1 j_1}, c'_{i_2 j_2}, c'_{i_3 j_3}, c'_{i_4 j_4}$ из C' тоже разные и тоже должны лежать на окружности. Это, однако, равносильно тому, что и исходная четверка точек $c_{i_1 j_1}, c_{i_2 j_2}, c_{i_3 j_3}, c_{i_4 j_4}$, и каждая из четверок $c'_{i_1 j_1}, c'_{i_2 j_2}, c'_{i_3 j_3}, c'_{i_4 j_4}$ лежат на бесконечном множестве эллипсов, что невозможно.

Положим, что из точек $c_{i_1 j_1}, c_{i_2 j_2}, c_{i_3 j_3}, c_{i_4 j_4}$ совпадают только две, например, $c_{i_1 j_1}$ и $c_{i_2 j_2}$. Это значит, что отрезки $(a_{i_1} a_{i_2})$ и $(b_{j_1}^0 b_{j_2}^0)$ равны по длине, параллельны и однонаправлены. Треугольник из двукратной точки $(c_{i_1 j_1}, c_{i_2 j_2})$ и точек $c_{i_3 j_3}$ и $c_{i_4 j_4}$ в характеристическом изображении должен быть остроугольным. Покажем, что изображение B_0 не может быть искомым, поскольку некоторым ε -аффинным преобразованием изображения B_0 можно уменьшить $R_A(B)$, приблизив все или часть точек $b_{j_1}^0, b_{j_2}^0, b_{j_3}^0, b_{j_4}^0$ к соответствующим точкам $a_{i_1}, a_{i_2}, a_{i_3}, a_{i_4}$.

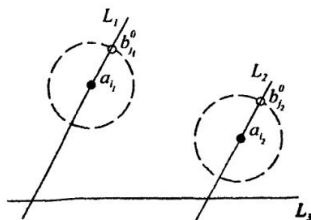


Рис. 52.

Рассмотрим параллельные прямые L_1 и L_2 , на которых лежат отрезки $(a_{i_1} b_{j_1}^0)$ и $(a_{i_2} b_{j_2}^0)$, и прямую L_3 , проходящую через точки $b_{j_3}^0$ и $b_{j_4}^0$. Положим, L_1 и L_2 имеют пересечение с L_3 и точки $b_{j_1}^0$ и $b_{j_2}^0$ лежат по одну сторону от L_3 (пример на рис. 52). Осуществим теперь ε -аффинное преобразование изображения B_0 , оставляющее неподвижной прямую L_3 , и придвигающее $b_{j_1}^0$ по прямой L_1 к точке a_{i_1} . При таком преобразовании $b_{j_2}^0$ тоже останется на прямой L_2 и придвинется к точке a_{i_2} , а это в целом уменьшит $R_A(B)$.

Аналогичные рассуждения можно провести и в случае, когда L_1 и L_2 параллельны и лежат по одну сторону от L_3 .

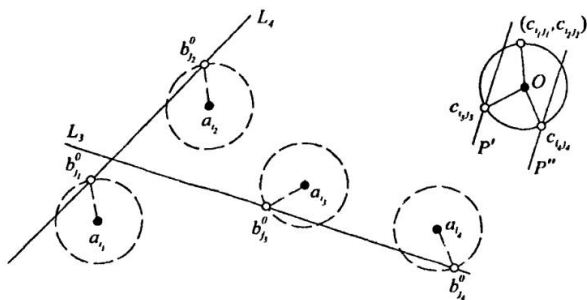


Рис. 53.

Пусть теперь точки $b_{j_3}^0$ и $b_{j_4}^0$ лежат по одну сторону от прямой L_4 , проходящей через $b_{j_1}^0$ и $b_{j_2}^0$ (пример на рис. 53, справа на рисунке представлено характеристическое изображение).

Треугольник из точек $(c_{i_1 j_1}, c_{i_2 j_2})$, $c_{i_3 j_3}$ и $c_{i_4 j_4}$ – остроугольный, угол между $(Oc_{i_3 j_3})$ и $(Oc_{i_4 j_4})$ – меньше π , поэтому всегда существуют параллельные прямые P' и P'' , проходящие через точки $c_{i_3 j_3}$ и $c_{i_4 j_4}$ и такие, что их части, примыкающие к этим точкам и находящиеся внутри ключевой окружности, расположены по одну сторону от прямой, на которой лежат точки $c_{i_3 j_3}$ и $c_{i_4 j_4}$. Сжатие или растяжение по направлению прямых P' и P'' , оставляющее неподвижной прямую L_4 , придвинет точки $b_{j_3}^0$ и $b_{j_4}^0$ к точкам a_{i_3} и a_{i_4} , что невозможно для искомого изображения.

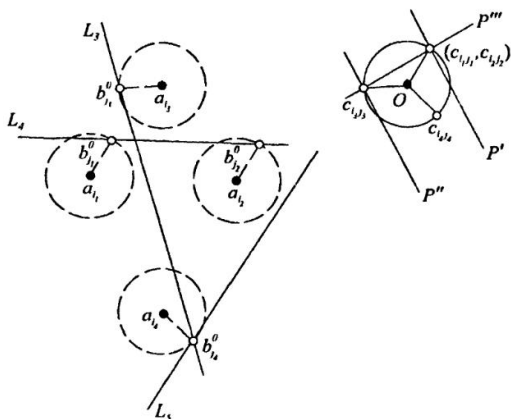


Рис. 54.

Остается случай, когда и точки $b_{j_1}^0$ и $b_{j_2}^0$ лежат по разные стороны от прямой L_3 , и точки $b_{j_3}^0$ и $b_{j_4}^0$ лежат по разные стороны от прямой L_4 (пример на рис. 54). Выберем на характеристическом изображении параллельные прямые P' и P'' , проходящие через кратную точку $(c_{i_1 j_1}, c_{i_2 j_2})$ и одну из точек $c_{i_3 j_3}$ и $c_{i_4 j_4}$ (на рис. 54 прямая P'' проходит через точку $c_{i_3 j_3}$) таким образом, чтобы части этих прямых, примыкающие к точкам и находящиеся внутри ключевой окружности, были расположены по одну сторону от прямой, проходящей через выбранные точки (прямая P''' на рис. 54). Через точку в B_0 , соответствующую оставшейся точке на ключевой окружности (точки $b_{j_4}^0$ и $c_{i_4 j_4}$ на рис. 54) проводим прямую L_5 так, чтобы точки $b_{j_1}^0$,

$b_{j_2}^0, b_{j_3}^0$ были по одну сторону от этой прямой. Это можно сделать, поскольку четырехугольник $b_{j_1}^0 b_{j_2}^0 b_{j_3}^0 b_{j_4}^0$ в этом случае – выпуклый. Теперь сжатие или растяжение изображения по направлению прямых P' и P'' , оставляющее неподвижной прямую L_5 , придвинет точки $b_{j_1}^0, b_{j_2}^0, b_{j_3}^0$ к соответствующим точкам $a_{i_1}, a_{i_2}, a_{i_3}$, что невозможно для искомого изображения.

Итак, все четыре точки на ключевой окружности разными быть не могут, и только две из них совпадать тоже не могут. Остается вариант, когда совпадают либо три точки, например, $c_{i_1 j_1}, c_{i_2 j_2}, c_{i_3 j_3}$ (пример на рис. 50), либо по две точки в двух парах, например, $c_{i_1 j_1}, c_{i_2 j_2}$ и $c_{i_3 j_3}, c_{i_4 j_4}$ (пример на рис. 51), что дает доказываемое утверждение.

Пусть теперь на ключевой окружности – пять точек ядра. Все они разными, или только две из них слившимися (то есть на окружности четыре точки, из них одна – кратная) быть не могут по тем же причинам, что и в рассмотренном выше случае с четырьмя разными точками на окружности. Значит должны быть по меньшей мере либо две пары слившихся точек, либо одна тройка слившихся точек.

Если на ключевой окружности больше пяти точек, то такие рассуждения применимы к любым пяти из них. Теорема доказана.

Отметим, что ранее в главе I были введены коды изображений, задающие их с точностью до аффинных преобразований. То, что в паре (A, B_0) изображения должны быть согласованы, означает, что коды этих изображений, в определенном смысле, сравнимы.

Итак, имеем теперь два необходимых условия для пары (A, B_0) : искомое изображение B_0 должно быть согласовано с A , и центр характеристического изображения пары (A, B_0) должен совпадать с центром ключевой окружности. Сочетание этих двух условий позволит нам далее вычленивать из B^* конечное подмножество U изображений, среди которых только и может находиться искомое изображение.

Действительно, согласовывать B' с A можно по разным парам отрезков. Пусть, например, задана пара отрезков $(a_{i_1} a_{i_2})$ и $(a_{i_3} a_{i_4})$, и пусть эти отрезки равны, параллельны и однонаправлены с соответствующими отрезками из B' , то есть с

отрезками $(b'_{\psi(i_1)} b'_{\psi(i_2)})$ и $(b'_{\psi(i_3)} b'_{\psi(i_4)})$. Однако этим условием определяется не единственное изображение из B^* , а некоторое их множество, и изображения в этом множестве переводимы друг в друга параллельными переносами. Но, как показано в §1, среди них есть только единственное – обозначим его через $B_{i_1 i_2, i_3 i_4}$ – для которого в паре с A центр ключевой окружности совпадает с центром характеристического изображения.

Отрезки $(a_{i_1} a_{i_2})$ и $(a_{i_3} a_{i_4})$ можно выбрать в A не более чем $(C_n^2)^2$ способами. Соответственно не больше будет и изображений $B_{i_1 i_2, i_3 i_4}$. Обозначим множество их через U^ψ . Объединение множеств U^ψ для всех возможных биекций ψ обозначим через U . Таким образом, приходим к следующему утверждению.

Теорема 5. *Искомое изображение B^0 может находиться только среди изображений множества U .*

Глава III

О сжатии информации при обработке и распознавании изображений

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 целиком всю информацию на сетчатке глаза. Должно существовать некоторое «сито» и для изображений из среды, и для визуальной информации из памяти при оперировании с ними. Парадокс, однако, состоит в том, что попытки до собственно распознавания выделить на изображениях «опорные точки», «важные детали», и пр. есть тоже распознавание, то есть возникает до некоторой степени замкнутый круг. Ясно, что нужно избегать этого.

Трудно рассчитывать эту задачу – построение «сита» для изображений – решить неким единовременным усилием. В настоящей работе предлагается некоторый шаг в ее решении в рамках строящейся здесь модели. Наложение изображений друг на друга осуществляем движениями, то есть геометрическими преобразованиями. У представленных процедур, однако, есть недостатки, причем существенные, если основывать на них реальное распознавание: сравниваемые изображения предполагаются состоящими из одинакового числа n точек, и при сравнении изображений нужно исследовать $n!$ вариантов соответствия точек друг другу. Ясно, что с ростом n такая процедура становится неприемлемой из-за вычислительных трудностей. В §1 устраняются эти недостатки, и зависимость числа вариантов сравнения от числа точек в каждом из изображений становится полиномиальной.

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. Поэтому в §2 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).

В §3 показывается, что для функции, характеризующей схожесть изображений, выполняется неравенство треугольника. Это позволяет сократить перебор при сравнении распознаваемого изображения с изображениями, хранящимися в памяти.

В §4 и §5 построения распространяются на более широкий, чем движения, класс геометрических преобразований – преобразования подобия и аффинные преобразования.

§1. Об определении искомого взаиморасположения двух изображений

Самое трудоемкое в построении множества U для изометрических преобразований – необходимость в переборе $n!$ биекций ψ . Вместе с тем, как это следует из описанного в главе II, 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 искомого взаиморасположение (см. описание первого случая в §2 главы II). 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 $\tilde{U}_1$.

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 искомого 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} (рис. 55), либо еще две точки c_{i_3} и c_{i_4} (рис. 56), 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

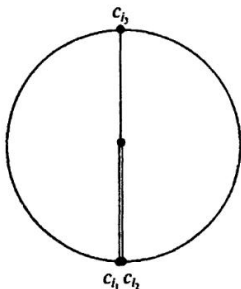


Рис. 55.

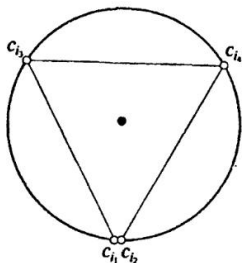


Рис. 56.

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 $c_{i_1} c_{i_2} c_{i_3} c_{i_4}$ (рис. 55 и 56). Отметим, что формируемое таким образом множество $\tilde{U}_0$ включает, очевидно, каждое из множеств U_0^ψ , $\bar{U}_0^\psi$ 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, как нетрудно видеть, следующее утверждение (аналог теоремы 2 главы II).

Теорема 1. *Искомое изображение может находиться только среди изображений множества $\tilde{U}$.*

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 A , 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 взаимное однозначное 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 соответствуют 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 (A, B') изображений 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.

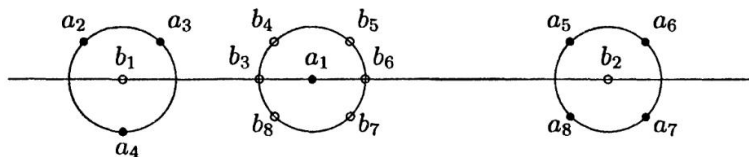


Рис. 57.

Измененное понятие искомого изображения более оправдано с содержательной точки зрения, чем прежнее. Если взять рис. 18, который использовался для иллюстрации понятия искомого взаиморасположения изображений, то применительно к нему прежнее определение и новое совпадают. Действительно, на расстоянии не большем R от каждой точки из A находится ровно одна точка из B и наоборот. Однако если взять

пример на рис. 57, то для него эти два определения дадут разные расстояния между изображениями A (из точек) и B (из кружочков). Каждое изображение состоит из 8 точек, и рис. 57 иллюстрирует их искомое взаиморасположение в рамках второго определения. Величина $R(A, B)$ не больше радиуса больших кругов на этом рисунке. При этом некоторым точкам одного изображения сопоставляются «кучки» точек другого изображения: $Q_{a_1} = \{b_3, b_4, b_5, b_6, b_7, b_8\}$, $Q_{b_1} = \{a_2, a_3, a_4\}$, $Q_{b_2} = \{a_5, a_6, a_7, a_8\}$. Ясно, что если потребовать биективного соответствия между точками изображений A и B_0 (в рамках прежнего определения), то естественное соответствие между точками изображений, которое представлено на рис. 57, разрушится.

§2. Приближение изображений их покрытиями

Зададимся некоторым числом r ($r \geq 0$). Выберем на A подмножество A^r точек изображения A с тем условием, чтобы каждая точка из A находилась на расстоянии, не большем r , от какой-либо точки из A^r . Ясно, что A^r – фрагмент изображения A , в частном случае A^r может совпадать с A . Изображение A^r можно трактовать как покрытие изображения A кругами радиуса r с центрами в точках из A^r , причем так, что каждая точка исходного изображения попадает хотя бы в один из этих кругов.

Содержательно A^r интерпретируется как «огрубление» изображения A , устранение на нем излишних деталей и подробностей, поэтому A^r называем эскизом изображения A . При заданном параметре r существует некоторое множество $\{A^r\} = \{A_1^r, \dots, A_k^r\}$ эскизов, причем исходное изображение A входит в это множество. Однако интерпретировать как «огрубление» A можно, очевидно, лишь те эскизы из $\{A^r\}$, которые состоят из меньшего числа точек, чем исходное изображение A . В частности, в $\{A^r\}$ есть изображение (может быть, не одно) с наименьшим среди $A_1^r, \dots, A_k^r$ числом точек. Такое изображение будем называть остовом изображения A .

Можно, очевидно, говорить об иерархии остовов, определя-

емой различными значениями r . Остовов (как и эскизов) – конечное множество, и наименьший по числу точек остов всегда (и для всех изображений) состоит из одной точки, наибольший – совпадает с исходным изображением A .

При оценке похожести двух изображений достаточно, может быть, оценить их сходство в целом, в главных чертах, без деталей. Такого рода интуитивные соображения неявно предполагают, что вместо собственно изображений при сравнении будут использованы другие изображения, полученные из исходных удалением излишних деталей, то есть некоторые наброски, эскизы. При этом нужно знать, как сходство эскизов связано со сходством оригиналов. Это и составляет содержательную суть следующего ниже утверждения.

Пусть изображения A и B взаиморасположены искомым образом. Обозначим через R_0 расстояние $R(A, B)$. Пусть A^{r_1} и B^{r_2} – эскизы изображений A и B . Обозначим через $R_{r_1 r_2}$ расстояние $R(A^{r_1}, B^{r_2})$ между эскизами, через r – большее из чисел r_1 и r_2 .

Теорема 2. *Для расстояния R_0 между изображениями A и B справедлива следующая оценка:*

$$(R_{r_1 r_2} - r) \leq R_0 \leq (R_{r_1 r_2} + r).$$

Доказательство. Докажем, что $(R_{r_1 r_2} - r) \leq R_0$, или $R_{r_1 r_2} \leq (R_0 + r)$.

Пусть a_i – произвольная точка из A^{r_1} . Покажем, что на расстоянии не большем $(R_0 + r)$ от a_i есть некоторая точка из B^{r_2} . Действительно, a_i – точка из A , значит на расстоянии не большем R_0 от нее существует точка из B , обозначим ее через b_j . Если точка b_j входит в B^{r_2} , то она искомая. Если b_j не входит в B^{r_2} то существует круг радиуса r_2 с центром в некоторой точке b_x из B^{r_2} , в котором точка b_j содержится. Расстояние от a_i до b_x не больше чем $(R_0 + r_2)$, и значит точка b_x – искомая.

Эти рассуждения можно повторить, поменяв ролями A^{r_1} и B^{r_2} , и получить, что для произвольной точки из B^{r_2} существует в A^{r_1} точка, расстояние до которой не больше $(R_0 + r_1)$. Отсюда $R_{r_1 r_2} \leq (R_0 + r)$.

Докажем, что $R_0 \leq (R_{r_1 r_2} + r)$.

Изображения A и B расположены, по условию, искомым образом. Из этого, однако, не следует, что и их фрагменты A^{r_1} и B^{r_2} тоже расположены искомым образом. Изометрическими преобразованиями расположим B^{r_2} относительно A^{r_1} искомым образом и обозначим через $B_0^{r_2}$. Это определит и некоторую трансформацию для B , которую обозначим через B_0 .

Пусть a_i – произвольная точка из A . Покажем, что на расстоянии не большем $(R_{r_1 r_2} + r_1)$ от a_i есть некоторая точка из B_0 . Если a_i входит в A^{r_1} , то это очевидно. Если не входит, то существует в A^{r_1} точка, находящаяся от a_i на расстоянии не большем r_1 , и значит есть в $B_0^{r_2}$ точка, удаленная от a_i на расстояние, не большее $(R_{r_1 r_2} + r_1)$.

Эти рассуждения можно повторить, поменяв ролями A и B_0 , и получить, что для произвольной точки из B_0 существует в A точка, удаленная не более чем на $(R_{r_1 r_2} + r_2)$. Отсюда $R_0 \leq (R_{r_1 r_2} + r)$. Теорема доказана.

Теорема 2 позволяет заменить вычисление расстояний между изображениями A и B оценками этого расстояния, причем чем меньше величина r , тем уже интервал, в который заключено $R(A, B)$, и тем ближе оценка к значению $R(A, B)$.

Поскольку при заданном r для A и B возможны эскизы с разным числом точек, то, имея ввиду упрощение вычислений, целесообразно брать для оценок эскизы с как можно меньшим числом точек. Лучше всего брать остовы. Но при этом нет необходимости брать «в точности» остовы, ибо нахождение остова может потребовать некоторого перебора. Можно использовать любое «приближение» к остову с достаточно небольшим числом точек.

Отметим, что биективное соответствие между точками изображений A и B иллюстрировалось рис. 18. Изменения в определениях из §1 иллюстрировались рис. 57 (точке одного изображения может соответствовать группа точек другого). Этот случай можно дополнить примерами на рис. 58 и 59 (на рис. 59 контуры фигур такие же, как на рис. 18, но более плотно заполнены точками). Наконец, соответствие между точками в A^{r_1} и B^{r_2} можно трактовать как соответствие между группами, «кучками» точек в исходных изображениях A и B (рис. 60), причем группы могут пересекаться. Другая возможная содер-

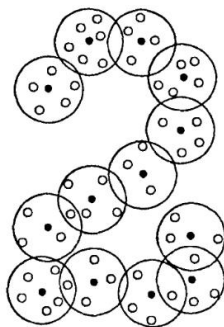


Рис. 58.

жательная трактовка этому – размытость, нечеткость сопоставляемых друг другу частей изображений.



Рис. 59.

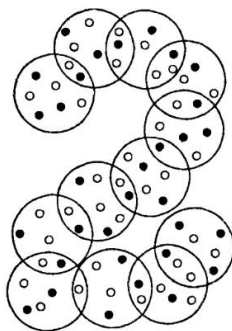


Рис. 60.

Пусть простая процедура распознавания, которую назовем первичной, состоит в следующем. Распознаваемое изображение A сравнивается с каждым из изображений-эталонов $B_1, \dots, B_k$, то есть вычисляются все расстояния $R(A, B_i)$, $i = 1, \dots, k$. Результатом распознавания будет изображение B_x с номером $x \in (1, \dots, k)$, расстояние до которого наименьшее, и отображения, сопоставляющие точки изображений A и B_x друг другу.

Теорема 2 позволяет в первичной процедуре распознавания заменить вычисление точных значений $R(A, B_i)$ ($i = 1, \dots, k$) их оценками. Действительно, пусть $\Delta'_x \leq R(A, B_x) \leq \Delta''_x$,

$\Delta'_i \leq R(A, B_i) \leq \Delta''_i, i = 1, \dots, k, i \neq x$, и пусть $\Delta''_x < \Delta'_i$ (то есть интервал для B_x не пересекается с интервалами для остальных $B_1, \dots, B_k$). Тогда заведомо из всех $R(A, B_i)$ ($i = 1, \dots, k$) расстояние $R(A, B_x)$ – наименьшее.

Поскольку для оценок нужно вычислять расстояния только между эскизами, то выигрыш за счет «упрощения» изображений может оказаться значительным. Нагляднее всего это можно продемонстрировать, показав возможность распространения предшествующих определений и результатов на случай изображений, бесконечных по числу точек.

На плоскости рассмотрим $\tilde{a}$ – замкнутую ограниченную область с границей, представляющей собой замкнутую непрерывную линию. Эта область $\tilde{a}$ – аналог точки в прежнем изображении, и, будем полагать, в частном, «вырожденном» случае может представлять собой точку. Ориентируясь на содержательную интерпретацию, $\tilde{a}$ можно назвать «мазком». Изображением A назовем теперь множество точек, представимое как объединение конечного числа мазков.

Расположение изображения B относительно «неподвижного» изображения A 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^* , где B^* – множество всех возможных изображений, получаемых из изометрическими преобразованиями.

Далее можно повторить определения и провести построения, сделанные выше для изображений с конечным числом точек, хотя, разумеется, процедуру вычисления $R(A, B)$ и определения искомого взаиморасположения для этого случая применить нельзя (из-за бесконечного числа фрагментов изображений). Однако теорема 2 остается справедливой и для этого случая. Поскольку в рамках процедуры распознавания не обязательно знать точные значения расстояний между изображениями – достаточно иметь их оценки – то, следовательно, можно использовать вместо A и B их остовы, а уж остовы при любом $r \neq 0$ состоят из конечного числа точек.

Такого рода рассмотрения представляют интерес, по-видимому, только как демонстрация возможности оперировать с бесконечными по числу точек изображениями. Реальные изображения естественнее полагать все же конечными.

§3. Сокращение перебора в памяти

Изображения-эталоны $B_1, \dots, B_k$ в первичной процедуре распознавания будем интерпретировать как хранящиеся в памяти распознающего устройства, изображение A – как изображение из окружающей среды, поступающее для распознавания. Для изображений в памяти можно считать известными все парные расстояния $R(B_i, B_j)$ ($i, j = 1, \dots, k$), поскольку их вычисления могут быть проведены заранее, как предобработка памяти, и до появления распознаваемого изображения A .

Возникает вопрос: всегда ли нужно при распознавании вычислять (или оценивать) расстояния от A до всех изображений памяти? Из содержательных соображений ясно, что если A очень похоже на одно из изображений B_x памяти, а B_x , в свою очередь, мало похоже на остальные изображения памяти, то вряд ли и A окажется похоже на эти остальные изображения.

Такого рода содержательные соображения имеют в основе своей, можно полагать, следующее утверждение.

Теорема 3. *Для изображений A, B_i, B_j ($i, j = 1, \dots, k$) выполняется неравенство: $R(A, B_j) \geq R(B_i, B_j) - R(A, B_i)$.*

Доказательство. Пусть каждое из изображений B_i и B_j расположено по отношению к A искомым образом. При этом по отношению друг к другу расположение B_i и B_j , конечно, не обязательно искомое.

Пусть b' – произвольная точка B_i . Тогда существует точка a из изображения A такая, что b' содержится в круге радиуса $R(A, B_i)$ с центром в a . В круге радиуса $R(A, B_j)$ с центром в этой же точке a должна содержаться хотя бы одна точка из B_j , обозначим ее через b'' . Отсюда расстояние между b' и b'' не больше $R(A, B_i) + R(A, B_j)$.

Этими же рассуждениями, поменяв ролями B_i и B_j , можно получить, что и для произвольной точки из B_j существует на

§4. О распознавании изображений, заданных с точностью до подобия

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. И хотя в этом случае есть свои особенности, в целом определения и построения не слишком меняются при расширении на преобразования подобия.

Действительно, несмотря на отсутствие первого случая и некоторые отличия для второго, третьего и вырожденного случаев, в основном решение для подобных преобразований повторяет решение для преобразований изометрических. Поэтому и переход к решению, исключающему $n!$ биекций, проводится так же, как и в §1. Смысл, как и ранее, состоит в том, чтобы вместо рассмотрения $n!$ биекций и фрагментов $\tilde{a}$ и $\tilde{b}$ в рамках каждой биекции, рассматривать фрагменты $\tilde{a}$ и $\tilde{b}$ отдельно, вне рамок заранее заданных биекций. Формируемое таким образом множество $\tilde{U}$ изображений из B^* будет включать искомое изображение.

Искомое соответствие между точками изображений A и B_0 как и в §1 определим тем, чтобы каждой точке из A (соответственно из B_0) ставились в соответствие все точки из B_0 (соответственно из A), находящиеся от нее на расстоянии, не большем $R(A, B)$.

Обозначим через R_0 расстояние $R(A, B)$, B_0 – искомое изображение. Рассматриваем эскизы A^{r_1} и B^{r_2} , и пусть преобразование, переводящее B в B_0 , переводит B^{r_2} в $B_0^{r_2}$.

Обозначим через $R_{r_1 r_2}$ расстояние между эскизами A^{r_1} и B^{r_2} , и пусть $B_{00}^{r_2}$ – трансформированный эскиз B^{r_2} , искомым образом расположенный относительно A^{r_1} . Пусть преобразование, переводящее B^{r_2} в $B_{00}^{r_2}$, переводит B в B_{00} .

Теорема 4. Для расстояния R_0 между изображениями A и B справедлива следующая оценка:

$$(R_{r_1 r_2} - \max(r_1, r_2^0, r_2^{00})) \leq R_0 \leq (R_{r_1 r_2} + \max(r_1, r_2^0, r_2^{00})).$$

Доказывается теорема практически повторением доказательства теоремы 2. Однако в использовании теоремы есть трудности. Для изометрических преобразований величины r_2 , r_2^0 , r_2^{00} совпадают, что и давало возможность считать известным $r = \max(r_1, r_2^0, r_2^{00})$. Для преобразований подобия мы можем считать известными r_1 и r_2 , поскольку заданы эскизы A^{r_1} и B^{r_2} , и величину r_2^{00} , поскольку считаем вычисленным $R_{r_1 r_2}$ и известным $B_{00}^{r_2^{00}}$. Величина же r_2^0 неизвестна, поскольку изображение B_0 и эскиз $B_0^{r_2^0}$ не заданы.

Можно попытаться выйти из положения, взяв величину r_2 настолько маленькой, чтобы наибольшей из r_1 , r_2^0 , r_2^{00} была бы заведомо величина r_1 . Однако остается вопросом, насколько малой должна быть величина r_2 , чтобы это выполнялось.

Есть, однако, «вырожденный» случай, когда это условие заведомо выполняется – случай $r_2 = 0$. При этом, конечно, и $r_2^0 = 0$, $r_2^{00} = 0$. Однако это означает, что B^{r_2} совпадает с B и выигрыш, который дает использование эскизов, обеспечивается только эскизами изображения A , а вот на изображениях памяти $B_1, \dots, B_k$ придется учитывать все точки. Показательно плохо и то, что перейти, например, к использованию изображений, бесконечных по числу точек (в памяти), в этом случае не удастся.

Однако каким-то образом неизвестная величина r_2^0 с известными A , B , B_{00} и r_2^{00} должна быть связана. Будем рассуждать следующим образом. Точки изображения B_{00} в целом находятся к точкам изображения A не ближе, чем точки изображения B_0 , то есть $l(B_{00}) \geq l(B_0) = R(A, B)$. Пусть b_x – произвольная точка в B_{00} . Ближайшая к b_x точка изображения A находится на расстоянии, не большем $(R_{r_1 r_2} + r_2^{00})$. Следовательно, две точки $b_{x'}$ и $b_{x''}$ из изображения B_{00} , расстояние между которыми наибольшее среди пар точек изображения B_{00} , находятся друг от друга на расстоянии, не большем $(a_{x'} a_{x''}) + 2(R_{r_1 r_2} + r_2^{00})$, где $(a_{x'} a_{x''})$ – наибольший по длине отрезок между парами точек в изображении A . Пусть $(b_y, b_{y''})$ – наибольший по длине

отрезок между парами точек изображения B . Обозначим через k отношение $(a_x, a_{x''}) + 2(R_{r_1 r_2} + r_2^{00})$ к $(b_y, b_{y''})$. Сдержательно это коэффициент подобия такой, что после преобразования B с этим коэффициентом (преобразованное B обозначим через $\tilde{B}$) наибольший отрезок в $\tilde{B}$ будет равен по длине $(a_x, a_{x''}) + 2(R_{r_1 r_2} + r_2^{00})$. Этот отрезок обозначим через $(\tilde{b}_y, \tilde{b}_{y''})$. Если он соответствует отрезку $(a_x, a_{x''})$ в A , то порождаемый отрезками $(\tilde{b}_y, \tilde{b}_{y''})$ и $(a_x, a_{x''})$ отрезок (c', c'') в характеристическом изображении имеет длину $2(R_{r_1 r_2} + r_2^{00})$. Следовательно, радиус ключевой окружности равен по меньшей мере $(R_{r_1 r_2} + r_2^{00})$. Если $(\tilde{b}_y, \tilde{b}_{y''})$ сопоставляется любому другому отрезку в A , то радиус ключевой окружности может от этого только увеличиться. Отсюда следует, что r_2^0 не может быть больше $r_2 \cdot k$. Поэтому в задаче о выборе малого r_2 достаточно выбрать величину r_2 настолько малой, чтобы $\max(r_1, r_2 \cdot k, r_2^{00}) = r_1$.

§5. О распознавании изображений, заданных с точностью до аффинных преобразований

Самое трудоемкое в построении множества U из §4 главы II – необходимость в переборе $n!$ биекций ψ . Рассмотрим, как можно обойти их использование.

Согласованность B_0 с A означает, что среди точек ядра на ключевой окружности характеристического изображения пары (A, B_0) есть трехкратная (или большей кратности) точка (первая ситуация), или есть две двухкратные точки (вторая ситуация). Оценим число возможных вариантов первой ситуации (множество $\tilde{U}_1$ изображений из B^*), и второй ситуации (множество $\tilde{U}_2$ изображений из B^*).

Наличие трехкратной точки означает, что некоторые два треугольника с вершинами в точках соответственно из A и B' , переводятся друг в друга параллельным переносом. Этим условием, однако, определяется не единственное B' , а некоторое подмножество $\{B'\}$ изображений из B^* , переводимых друг в друга параллельными переносами. Пар треугольников с вер-

пинами из A и B можно выбрать и сопоставить друг другу не более чем $(3!)(C_n^3)^2$ способами (с учетом возможности по разному сопоставлять друг другу вершины в треугольниках).

Трехкратная точка не может быть единственной точкой на ключевой окружности. Должны быть либо еще одна точка, образующая с кратной точкой концы диаметра ключевой окружности, либо еще две точки, образующие с кратной точкой остроугольный треугольник. В первом случае в $\{B'\}$ существует единственное изображение, для которого в паре с A центр характеристического изображения совпадает с серединой диаметра. Эту дополнительную точку на ключевой окружности можно породить, очевидно, не более $(n-3)^2$ способами сопоставления по одной точке из A и B' друг другу. Во втором случае в $\{B'\}$ существует единственное изображение, для которого в паре с A центр характеристического изображения совпадает с центром описанной окружности остроугольного треугольника. Дополнительные две точки на ключевой окружности можно породить не более чем $2(C_{n-3}^2)^2$ способами сопоставления пар точек из A и B' друг другу.

В целом получаем, что в $\tilde{U}_1$ может быть не более чем $3!(C_n^3)^2((n-3)^2 + 2(C_{n-3}^2)^2)$ изображений.

Рассмотрим теперь вторую ситуацию, когда на ключевой окружности находятся две двухкратные точки. Четверок точек из A и B' , порождающих эти двухкратные точки, можно выбрать и сопоставить друг другу не более чем $4(C_n^2)^4$ способами. Далее возможен случай, когда двухкратные точки являются концами диаметра ключевой окружности, и случай, когда есть еще дополнительная точка, образующая с кратными точками остроугольный треугольник. В первом случае однозначно определяется изображение из B^* , для которого в паре с A центр характеристического изображения совпадает с серединой диаметра. Во втором случае существует в B^* единственное изображение, у которого в паре с A центр характеристического изображения совпадает с центром описанной окружности остроугольного треугольника. Дополнительную точку на окружности можно породить выбором и сопоставлением друг другу по одной точке из A и B' . Вариантов такого выбора будет, очевидно, не более $(n-4)^2$.

В целом получаем, что в $\tilde{U}_2$ может быть не более чем $4(C_n^2)^4(1 + (n - 4)^2)$ изображений.

Обозначим через $\tilde{U}$ объединение множеств $\tilde{U}_1$ и $\tilde{U}_2$. Из рассмотренного вытекает следующее утверждение (аналогичное теореме 5 главы II).

Теорема 5. *Искомое изображение B_0 может находиться только среди изображений множества $\tilde{U}$.*

Формирование множества $\tilde{U}$, в отличие от U , не требует перебора $n!$ взаимно однозначных соответствий между точками изображений A и B . Если фрагментом изображения называть некоторое подмножество точек в нем, то при построении $\tilde{U}$ в сущности сопоставляются друг другу только фрагменты изображений A и B , а не изображения в целом. 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 одинаковом расстоянии, которое обозначим через $R_{\tilde{a}}(\tilde{b})$. Если B' – искомое изображение, то и все остальные точки из B' должны находиться от соответствующих точек из A 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 всех тех точек из A , 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 взаимно однозначное 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 попрежнему выбирать изображение B_0 с наименьшей величиной $R_{\tilde{a}}(\tilde{b})$, но среди приемлемых изображений, а не узкоприемлемых. Точке a_i , будем полагать, соответствует теперь не одна точка, а все точки из множества Q_{a_i} , точке b_j^0 соответствуют 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. Так измененное определение более оправдано с содержательной точки зрения.

Литература

1. Глезер В.Д. Зрение и мышление. СПб.: Наука, 1993. 284 с.
2. Клейн Ф. Геометрия. М.: Наука, 1987.
3. Козлов В.Н. Математическое моделирование зрительно-го восприятия // Математические вопросы кибернетики. Вып. 6. М.: Наука, 1996. С. 321–338.
4. Козлов В.Н. О распознавании аффинно разных дискретных изображений // Интеллектуальные системы. 1998. Том 3. Вып. 3–4. С. 95–122.
5. Козлов В.Н. О зрительном образе, математических подходах к определению этого понятия и о распознавании изображений // Журнал вычислительной математики и математической физики. 1999. Том 39. №11. С. 1929–1946.
6. Крушинский Л.В., Козлов В.Н., Кудрявцев В.Б. О некоторых результатах применения математики к моделированию в биологии // Математические вопросы кибернетики. Вып. 1. М.: Наука, 1988. С. 52–88.
7. Фу К., Гонсалес Р., Ли К. Робототехника. М.: Мир, 1989. (Fu K.S., Gonzalez R.C., Lee C.S.G. Robotics: Control, Sensing, Vision, and Intelligence. McGraw-Hill Book Company, New York, 1987.)
8. Хорн Б.К.П. Зрение роботов. М.: Мир, 1989. (Horn B.K.P., Robot Vision. The MIT Press, Cambridge, Massachusetts, 1986.)
9. Хьюбел Д. Глаз, мозг, зрение. М.: Мир, 1990. (Hubel D.H. Eye, Brain, and Vision. Scientific American Library, New York, 1988.)

10. Kozlov V.N. Image Coding and Recognition and Some Problems of Stereovision // Pattern Recognition and Image Analysis. 1997. Vol. 7. N 4. P. 448–466.
11. Kozlov V.N. Visual Pattern and Geometric Transformations of Images // Pattern Recognition and Image Analysis. 2000. Vol. 10. N 3. P. 321–342.

Козлов Вадим Никитович

Элементы математической теории зрительного восприятия
М., Издательство Центра прикладных исследований при механико-
математическом факультете МГУ, 128 стр.

*Оригинал макет изготовлен издательской группой
механико-математического факультета МГУ*

Подписано в печать 10.10.2001 г.

Формат 60×90 1/16. Объем 8 п.л.

Заказ Тираж 250 экз.

Издательство ЦПИ при механико-математическом факультете МГУ
г. Москва, Ленинские горы.

Лицензия на издательскую деятельность ЛР №040746,
от 12.03.1996 г.

Отпечатано на типографском оборудовании механико-математичес-
кого факультета и Франко-русского центра им. А.М.Ляпунова.