Распознавание изображений, представляемых конечным множеством точек

В. Н. КОЗЛОВ
Московский государственный университет им. М. В. Ломоносова
e-mail: vnkozlov@mail.ru

УДК 519.95

Ключевые слова: образ, код изображения, распознавание, стереозрение, аффинное преобразование.

Аннотация

В работе предлагается новый подход к решению проблемы распознавания зрительных образов, основанный на использовании специального кодирования изображений. Это кодирование является инвариантным по отношению к аффинным преобразованиям множества дискретных точек (дискретных фигур).

Abstract

V. N. Kozlov, Recognition of the images represented by final set of points, Fundamentalnaya i prikladnaya matematika, vol. 15 (2009), no. 5, pp. 95—110.

We describe a new approach to the problem of visual pattern recognition. This approach is based on figure encoding. The encoding used is invariant under affine transformations of discrete sets of points (discrete figures).

Говоря об изображениях и их распознавании с позиций математики, важно сразу как можно яснее определить, что понимается под изображением. В этой работе мы будем понимать под изображением конечное (непустое) множество точек на плоскости. Содержательным обоснованием этому может служить то, что любое реальное (нецветное) изображение можно аппроксимировать изображением из точек, причём в нужной мере можно передать все градации «серого цвета» разной плотностью точек в разных частях изображения. Такое представление не закрывает дорогу и к рассмотрению цветных изображений, поскольку, как известно, цветное изображение можно представить тремя нецветными.

Перенумеруем некоторым образом точки изображения A так, чтобы номера были попарно различны. Обозначим через M_A множество этих номеров. Пусть S_{mnu} и S_{ksp} — площади треугольников с вершинами в тройках точек с номерами m, n, u и k, s, p, и пусть $\rho_{mnu,ksp} = S_{mnu}/S_{ksp}$. Полагаем, что порядок номеров в тройках не важен, сами тройки различны и при $S_{ksp} = 0$ значение $\rho_{mnu,ksp}$ не определено. Множество индексированных чисел $\rho_{mnu,ksp}$ для всех таких пар троек обозначим через T_A. Код изображения A — пара $\langle M_A, T_A\rangle$.

Фундаментальная и прикладная математика, 2009, том 15, № 5, с. 95—110.

Изображения, все точки которых расположены на одной прямой, мы не рассматриваем, поскольку код для них не определён. Изображения A и B с кодами $\langle M_A, T_A\rangle$ и $\langle M_B, T_B\rangle$ назовём эквивалентными, если существует такая биекция $\psi\colon M_A \to M_B$, что для любых m, n, u, k, s, p из M_A выполнено $\rho_{mnu,ksp} = \rho_{\psi(m)\psi(n)\psi(u),\psi(k),\psi(s),\psi(p)}$. Ясно, что эквивалентность изображений содержательно означает одинаковость их кодов с точностью до перенумерации точек. Два изображения будем называть аффинно эквивалентными, если они переводимы друг в друга аффинными преобразованиями. Изображение будем называть плоским, если не все его точки лежат на одной прямой или на двух параллельных прямых.

Теорема 1. *Два плоских изображения эквивалентны тогда и только тогда, когда они аффинно эквивалентны.*

Пусть $i_1, \ldots, i_k$ — номера k ($k \geqslant 3$) точек изображения A. Рассмотрим выпуклый многоугольник, включающий все эти точки, часть из которых образует его множество вершин. Ясно, что такой многоугольник определяется однозначно. Пусть $S_{i_1,\ldots,i_k}$ — площадь этого многоугольника, которую мы будем называть площадью k-точечника. Если точки $i_1, \ldots, i_k$ расположены на одной прямой, то полагаем $S_{i_1,\ldots,i_k} = 0$. Пусть $j_1, \ldots, j_k$ — другой k-точечник из A. Полагаем $\rho_{i_1,\ldots,i_k,j_1,\ldots,j_k} = S_{i_1,\ldots,i_k}/S_{j_1,\ldots,j_k}$. При этом если $S_{j_1,\ldots,j_k} = 0$, то считаем, что значение $\rho_{i_1,\ldots,i_k,j_1,\ldots,j_k}$ не определено. Множество всех таких индексированных чисел $\rho_{i_1,\ldots,i_k,j_1,\ldots,j_k}$ обозначим через T_A. Кодом изображения A назовём пару $\langle M_A, T_A\rangle$. Изображения A и B с кодами $\langle M_A, T_A\rangle$ и $\langle M_B, T_B\rangle$ назовём k-эквивалентными, если существует такая биекция $\psi\colon M_A \to M_B$, что для любых $i_1, \ldots, i_k, j_1, \ldots, j_k$ из M_A выполнено $\rho_{i_1,\ldots,i_k,j_1,\ldots,j_k} = \rho_{\psi(i_1),\ldots,\psi(i_k),\psi(j_1),\ldots,\psi(j_k)}$.

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

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

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

Назовём трёхмерным изображением A или телом конечное множество точек в трёхмерном евклидовом пространстве. Занумеруем попарно различными номерами точки изображения A. Пусть $\bar{M}_A$ — множество этих номеров, V_{mnuv} и V_{kspq} — объёмы тетраэдров с вершинами в четвёрках точек с номерами m, n, u, v и k, s, p, q, и пусть $\rho_{mnuv,kspq} = V_{mnuv}/V_{kspq}$. Полагаем, что порядок номеров в четвёрках не важен, сами четвёрки различны, для случая, когда $V_{kspq} = 0$, $\rho_{mnuv,kspq}$ не определено. Множество индексированных чисел $\rho_{mnuv,kspq}$ для всех таких пар четвёрок обозначим через $\bar{T}_A$. Кодом тела A назовём пару $\langle \bar{M}_A, \bar{T}_A\rangle$. Тела A и B с кодами $\langle \bar{M}_A, \bar{T}_A\rangle$ и $\langle \bar{M}_B, \bar{T}_B\rangle$ назовём эквивалентными, если существует такая биекция $\psi\colon \bar{M}_A \to \bar{M}_B$, что для любых m, n, u, v, k, s, p, q из $\bar{M}_A$ выполнено $\rho_{mnuv,kspq} = \rho_{\psi(m)\psi(n)\psi(u)\psi(v),\psi(k)\psi(s)\psi(p)\psi(q)}$.

Тела, все точки которых расположены в одной плоскости, будем называть двумерными, и для них рассматриваемый код не определён. Тела будем называть аффинно эквивалентными, если они переводятся друг в друга аффинными преобразованиями. Трёхмерное изображение назовём объёмным, если не все его точки лежат в одной плоскости или в двух параллельных плоскостях.

Теорема 3. *Два объёмных изображения эквивалентны тогда и только тогда, когда они аффинно эквивалентны.*

Восстановление трёхмерного изображения по плоским проекциям, с одной стороны, служит предположительной основой механизмов стереоскопического зрения в живых организмах, с другой — является важной задачей в рамках машинного зрения для робототехники.

Рассмотрим тело T и прямую, называемую направлением проекции. Направления проекции назовём разными, если они не параллельны. Проведём через каждую точку тела T прямые, параллельные направлению проекции α, называемые лучами. Полагаем α таким, что на каждом луче находится только одна точка тела. Таких направлений проекции бесконечное множество, не таких — только конечное. Назовём плоскость, пересекающую лучи, плоскостью проекции, изображение, образованное точками пересечения лучей с плоскостью проекции — проекцией тела (на данную плоскость и по данному направлению). Будем рассматривать проекции тела T по разным направлениям и на разные плоскости. Оговорим, что если T — двумерное изображение, то мы предполагаем, что α не параллельно плоскости этого изображения. Взаимно-однозначное соответствие между точками двух изображений назовём 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. Таким образом устанавливается 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 назовём a'-эквивалентными, если можно перевести одно из них в другое аффинными преобразованиями так, что совместятся соответствующие друг другу точки (обозначение: $A \approx B$). В противном случае A и B назовём a'-разными (обозначение: $A \not\approx B$). Часть изображения A, состоящую из его точек $a, b, \dots, v$, будем обозначать $A(a, b, \dots, v)$. Три точки тела, не лежащие на одной прямой, назовём гранью, а определяемую этими точками плоскость — плоскостью грани. Три точки проекции, не лежащие на одной прямой, назовём треугольником.

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

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

Рассмотрим на S_1 и S_2 четвёрку троек точек

$$\langle (a_1b_1c_1)(a_2b_2c_2)(c_1d_1e_1)(c_2d_2e_2)\rangle,$$

где $a_1b_1c_1$ и $a_2b_2c_2$ — треугольники, точки d_1 и e_1 лежат вне плоскости треугольника $a_1b_1c_1$ (и точки d_2 и e_2 лежат вне плоскости треугольника $a_2b_2c_2$) и из троек $c_1d_1e_1$ и $c_2d_2e_2$ хотя бы одна — треугольник. Такую четвёрку назовём правильной. Её точки на S_1 и S_2 являются проекциями точек a, b, c, d, e тела T, причём abc и cde — грани и плоскости этих граней разные. Обозначим через L линию пересечения этих плоскостей, через L_1 и L_2 — проекции 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 трансформированные L_1 и L_2 на них. Доказано, что по $\tilde{S}_1$ и $\tilde{S}_2$ можно определить положение $\tilde{L}_1$ и $\tilde{L}_2$ на них.

Опишем процедуру $\mathrm{Alg}\,T'$ 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. Проведём через точки a_1, b_1, c_1, e_1 лучи, параллельные α', и на каждом луче возьмём по точке (соответственно a', b', c', e') так, чтобы они не лежали в одной плоскости. Пусть теперь d_1 — произвольная точка на $\tilde{S}_1$. Построим соответствующую ей точку d' тела T'.

1. Пусть d_1 лежит в плоскости треугольника $a_1b_1c_1$. Совместим аффинными преобразованиями изображения $\tilde{S}_1$ точки a_1, b_1, c_1 с точками a', b', c'. Точка, в которую перейдёт 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 четвёрка. Построим по $\tilde{S}_1$ и $\tilde{S}_2$ прямые $\tilde{L}_1$ и $\tilde{L}_2$ на них. Аффинными преобразованиями изображения $\tilde{S}_1$ совместим точки a_1, b_1, c_1 с точками a', b', c'. Прямую, в которую преобразуется $\tilde{L}_1$, обозначим через L. Пусть для определённости $c_2d_2e_2$ — треугольник. Аффинными преобразованиями изображения $\tilde{S}_2$ (с прямой $\tilde{L}_2$) совместим $\tilde{L}_2$ с L и точку e_2 с точкой e'. Точка, в которую перейдёт при этом d_2, есть d'.

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

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

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

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

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

Теорема 4. *Если S_1 и S_2 — a'-разные проекции тела T, $S_1 \approx \tilde{S}_1$, $S_2 \approx \tilde{S}_2$ и тело T' построено по $\tilde{S}_1$ и $\tilde{S}_2$ процедурой $\operatorname{Alg} T'$, то тела T и T' a'-эквивалентны.*

Содержательно теорема состоит в следующем. В процедуре $\operatorname{Alg} T'$ есть много возможностей выбора. Можно, например, выбирать разные $\tilde{S}_1$ и $\tilde{S}_2$, по-разному брать исходную четвёрку точек на $\tilde{S}_1$ (или на $\tilde{S}_2$) и т. д. Варьируя выбор, можно построить некоторое множество $\{T\}$ тел. Теорема, однако, утверждает, что все они a'-эквивалентны телу T и, значит, a'-эквивалентны между собой. Тем самым посредством процедуры $\operatorname{Alg} T'$ тело восстанавливается с точностью до аффинных преобразований. При этом нет необходимости знать расстояние между проекциями S_1 и S_2 (аналог расстояния между сетчатками глаз), тело T' строится по произвольным образом сдвинутым, повернутым, сжатым или растянутым, уменьшенным или увеличенным проекциям тела T.

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

До сих пор предполагалось, что точки на $\tilde{S}_1$ и $\tilde{S}_2$ размечены. Положим теперь, что разметка неизвестна. Если каждое из изображений $\tilde{S}_1$ 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 существуют тело T' из точек $z_1, \dots, z_n$ и такие его проекции S'_1 и S'_2, что $S'_1 \approx \tilde{S}_1$, $S'_2 \approx \tilde{S}_2$ и точки x_{i_k} и y_{j_k} являются проекциями точки z_k, $k = 1, \dots, n$. Тогда данный вариант разметки назовём приемлемым вариантом или решением. По условию одно решение (с исходным телом T) существует, однако в общем случае решение не единственное. Пример нетрудно построить с использованием симметричного объёмного тела.

Опишем процедуру проверки данной разметки на приемлемость. Выберем на $\tilde{S}_1$ и $\tilde{S}_2$ такие 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}, что $x_{i_a}x_{i_b}x_{i_c}$ и $y_{j_a}y_{j_b}y_{j_c}$ — треугольники и $\tilde{S}_1(x_{i_a}, x_{i_b}, x_{i_c}, x_{i_e}) \not\approx \tilde{S}_2(y_{j_a}, y_{j_b}, y_{j_c}, y_{j_e})$. Проведём лучи, параллельные некоторому направлению α', через все точки изображения $\tilde{S}_1$. На лучах, проходящих через x_{i_a}, x_{i_b}, x_{i_c}, x_{i_e}, возьмём точки z'_a, z'_b, z'_c, z'_e так, чтобы они не лежали в одной плоскости. Если разметка приемлемая, то далее использованием $\operatorname{Alg} T'$ можно было бы построить остальные точки тела T' и они лежали бы на лучах от соответствующих точек на $\tilde{S}_1$. Проделаем аналогичные построения для $\tilde{S}_2$: проведём лучи по направлению α'', на лучах через точки y_{j_a}, y_{j_b}, y_{j_c}, y_{j_e} отметим точки z''_a, z''_b, z''_c, z''_e, не лежащие в одной плоскости, и далее, если разметка приемлемая, можно построить тело T''. Возьмём теперь всю конструкцию из точек изображения $\tilde{S}_1$ и лучей проекции через эти точки и аффинными преобразованиями её как целого совместим точки z'_a, z'_b, z'_c, z'_e с точками z''_a, z''_b, z''_c, z''_e. Если тела T' и T'' существуют, то по теореме 4 они a'-эквивалентны, значит, они совместятся 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 четвёрках x_{i_a}, x_{i_b}, x_{i_c}, x_{i_e} и y_{j_a}, y_{j_b}, y_{j_c}, y_{j_e} лучи проекции от некоторой точки x на $\tilde{S}_1$ и соответствующей ей (при данной разметке) точки y на $\tilde{S}_2$ не сходятся. Тогда они не будут сходиться и при всех вариантах разметок, в которых друг другу в соответствие ставятся точки данных четвёрок и точка x соответствует точке y. Все такие разметки могут быть заранее исключены как неприемлемые. С учётом этого оценка числа вариантов разметки для проверки на приемлемость может быть понижена до n^4.

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

Сравнение «на похожесть» двух изображений предлагается осуществлять таким наложением одного из них на другое, при котором соответствующие точки на двух изображениях оказываются как можно ближе друг к другу. Рассмотрим следующий пример. Каждое из изображений A и B, например цифры «2», состоит из n точек. Фигуру B движениями, т. е. параллельными переносами, поворотами и «переворотами» (преобразованиями симметрии относительно прямой) расположили специальным образом относительно A. Это расположение характеризуется тем, что каждая точка из B оказывается отстоящей от ровно одной соответствующей ей точки из A на расстояние, не большее некоторого R. Положим, что это R уже нельзя уменьшить никакими движениями фигуры B. В этом случае можно считать, что фигура B своими контурами повторяет фигуру A, а величина R — характеристика рассогласования в контурах. Если на изображение A пытаться наложить движениями не «двойку», а другую фигуру, например «четвёрку», можно ожидать, что характеристика рассогласования будет существенно больше. На этом и предлагается основывать процедуры различения фигур по форме и распознавания изображений в целом.

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

Множество $B^p \cup \bar{B}^p$ обозначим через B^*. Ясно, что B^* — множество всех изображений, которые можно получить из B изометрическими преобразованиями. Далее остаётся выбрать из величин $R_\psi(A, B)$, полученных для всех возможных ψ, наименьшую. Этот минимум обозначим через $R(A, B)$, биекцию, на которой он достигается, обозначим через ψ_0 и назовём искомым соответствием между точками изображений A и B. Изображение $B_0^{\psi_0}$ обозначим через B_0 и назовём искомым, взаиморасположение A и B_0 тоже искомое. Величину $R(A, B)$ будем называть также расстоянием между A и B.

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

Введём понятие угла между изображениями. Зафиксируем некоторые индексы p и q и в качестве угла между изображениями в паре (A, B') возьмём угол φ, образованный отрезками $(a_p\, a_q)$ и $(b'_p\, b'_q)$ или их продолжениями. Без ограничения общности можно полагать, что для исходных изображений A и B отрезки $(a_p\,a_q)$ и $(b'_p\,b'_q)$ параллельны и однонаправленны (т. е. если в отрезке $(a_p\,a_q)$ слева направо идёт сначала точка a_p, а затем a_q, то и в отрезке $(b'_p\,b'_q)$ слева направо идут последовательно точки b'_p и b'_q) и угол φ равен нулю. Для произвольного изображения B' из B^p полагаем, что угол φ между B' и A находится в промежутке от 0 до 2π.

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

Ранее было доказано утверждение, которое в обозначениях настоящей статьи выглядит следующим образом.

Теорема 5. *В $\{B\}_\varphi$ существует и единственно изображение, на котором достигается минимум величин $l(B')$, рассматриваемых для всех $B' \in \{B\}_\varphi$.*

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

В построениях используется понятие характеристического изображения C для пары (A, B'), где $B' \in B^*$. Оно состоит из точки O и точек $c_1, \ldots, c_n$, называемых соответственно центром и точками ядра. В качестве O берётся произвольная точка плоскости. Затем параллельным переносом отрезка $(a_i\,b'_i)$ $(i = 1, \ldots, n)$ совмещаем точку a_i с точкой O. Точку, в которую переходит при этом b'_i, обозначаем через c_i. Обратной процедурой при заданных A и C получается, как нетрудно видеть, изображение B'. Отметим, что некоторые из точек характеристического изображения могут совпадать (сливаться), и в этом смысле C является особым изображением.

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

Пусть c_i и c_j — точки из ядра характеристического изображения ($i \neq j$, $i, j = 1, \ldots, n$). Из рассмотрений в [1—3] следует, что если параллельными переносами отрезков $(a_i\,a_j)$ и $(b'_i\,b'_j)$ совместить точку a_j с c_i и точку b'_j с c_j, то точки a_i и b'_i совместятся в одну точку (обозначим её через c_{ij}). При заданных изображениях A и B тем самым стороны $(c_i\,c_{ij})$, $(c_j\,c_{ij})$ и угол между этими сторонами в треугольнике $c_ic_jc_{ij}$ (обозначим его через α_{ij}) можно считать известными. Следовательно, может быть получена и длина отрезка $(c_i\,c_j)$:

$$(c_ic_j)^2 = (a_ia_j)^2 + (b'_ib'_j)^2 - 2(a_ia_j)(b'_ib'_j)\cos\alpha_{ij}.$$

Поскольку угол α_{ij} зависит от угла φ между изображениями, то длина отрезка $(c_i\,c_j)$ есть функция от угла φ.

Из формулы для отрезка $(c_i\,c_j)$ следует, что он равен нулю, т. е. точки c_i и c_j слиты в одну, только при равенстве длин отрезков $(a_i\,a_j)$ и $(b'_i\,b'_j)$, их параллельности и однонаправленности. Угол α_{ij} при этом равен нулю. Существует лишь конечное число углов φ, при которых это может иметь место. Для каждого такого φ из множества $\{B\}_\varphi$ выбираем то изображение B', для которого центр характеристического изображения пары (A, B') совпадает с центром ключевой окружности и при этом слившиеся точки c_i и c_j лежат на ключевой окружности. Множество всех таких изображений обозначим через U_0^ψ. При дальнейших рассмотрениях полагаем, что уже никакие две точки ядра на ключевой окружности не слиты в одну.

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

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

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

Доказано, что прямые L_1, L_2, L_3 должны пересекаться в одной точке. Точку пересечения обозначим через O_L и назовём центром трёхосника, состоящего из осей L_1, L_2, L_3. На каждой оси часть от центра в направлении оси будем называть положительной, оставшуюся часть — отрицательной. Отрезки $(O_L\,a_{i_1})$, $(O_L\,a_{i_2})$, $(O_L\,a_{i_3})$ обозначим через x, y, z соответственно. Длину отрезка $(O_L\,a_{i_1})$ полагаем равной $|x|$, x полагаем положительным, если a_{i_1} находится в положительной части оси L_1, и отрицательным в противном случае. То же положим относительно 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 α, β, γ соответственно.

Ясно, что задание конкретных значений для x, y, z, α, β, γ, R определяет конкретное положение точек $b^0_{i_1}$, $b^0_{i_2}$, $b^0_{i_3}$ относительно точек a_{i_1}, a_{i_2}, a_{i_3}, а значит, и конкретное изображение из B^p.

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

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

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

Случай 3. На ключевой окружности находятся четыре или более точек. Рассмотрим четыре из них: c_{i_1}, c_{i_2}, c_{i_3}, c_{i_4}. Из теоремы Птолемея следует, что для того чтобы эти четыре точки находились на одной окружности, необходимо и достаточно, чтобы выполнялось одно из трёх равенств:

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

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

Далее меняем изображение B преобразованием симметрии относительно прямой на изображение $\bar{B}$ и проделываем всё аналогичное описанному выше для изображения $\bar{B}$ и множества $\bar{B}^p$. Получаем множества $\bar{U}_0^\psi$, $\bar{U}_1^\psi$, $\bar{U}_2^\psi$, $\bar{U}_3^\psi$. Объединением всех множеств U_0^ψ, U_1^ψ, U_2^ψ, U_3^ψ, $\bar{U}_0^\psi$, $\bar{U}_1^\psi$, $\bar{U}_2^\psi$, $\bar{U}_3^\psi$ и для всех биекций ψ получаем множество U.

Следующее утверждение можно считать некоторой перефразировкой теоремы из [1,2].

Теорема 6. *Искомое изображение B_0 может находиться только среди изображений множества U.*

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

Частью или фрагментом изображения A назовём 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 хорошее взаиморасположение (см. описание первого случая выше). С учётом возможности по-разному ставить в соответствие друг другу точки во фрагментах $\tilde{a}$ и $\tilde{b}$ общее число вариантов соответствия фрагментов друг другу будет $(2!)(C_n^2)^2$. Каждый вариант определяет два (с учётом преобразования симметрии) изображения из B^*. Множество всех таких изображений обозначим через U_1^{ψ}.

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

Наконец, рассмотрение аналогичным образом $(4!)(C_n^4)^2$ вариантов соответствия фрагментов из четырёх точек позволит сформировать множество изображений из B^*, которое мы 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 такими изображениями A и B'. Такая же проблема возникала и выше, при формировании множества U_0^{ψ}, но там была задана в целом биекция ψ, что позволяло отобрать в множество U_0^{ψ} только те изображения 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 ещё одна точка c_{i_3}, либо ещё две точки c_{i_3} и c_{i_4}, не слившиеся с c_{i_1} и c_{i_2}. В качестве a_{i_3} и b_{i_3} (в варианте с тремя точками) могут быть взяты каждая из $n-2$ точек изображения A и каждая из $n-2$ точек изображения B, т. е. имеется $(n-2)^2$ вариантов. Аналогично вариантов с четырьмя точками может быть не более $2(C_{n-2}^2)^2$. Для каждого из вариантов берём из B^* для включения в $\tilde{U}_0$ те изображения, для которых центр характеристического изображения совпадает с центром окружности, на которой лежат точки c_{i_1}, c_{i_2}, c_{i_3} или c_{i_1}, c_{i_2}, c_{i_3}, c_{i_4}. Отметим, что формируемое таким образом множество $\tilde{U}_0$ включает, очевидно, каждое из множеств U_0^{ψ}, $\bar{U}_0^{\psi}$ 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 $\tilde{U}$ вытекает, как нетрудно видеть, следующее утверждение.

Теорема 7. *Изображение B_0 принадлежит множеству $\tilde{U}$.*

Формирование $\tilde{U}$, в отличие от U, не требует перебора $n!$ взаимно-однозначных соответствий между точками изображений A и B. При построении $\tilde{U}$ друг другу в соответствие ставятся только фрагменты $\tilde{a}$ и $\tilde{b}$ и точки в них. Число вариантов таких соответствий, как это видно из рассмотренного выше, зависит от n полиномиально.

Однако формирование множества $\tilde{U}$ ещё не решает задачу. Надо определить, каким образом из $\tilde{U}$ будет выделяться искомое изображение B_0 и как будут ставиться друг другу в соответствие 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 мы 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 вытекает процедура определения приемлемого изображения из $\tilde{U}$. Для каждой точки a_i из A определим множество Q_{a_i} всех тех точек из B', расстояние до которых от точки a_i не больше $R_{\tilde{a}\tilde{b}}$. Аналогично через $Q_{b'_j}$ обозначим множество тех точек из B', расстояние до которых от b'_j не превышает $R_{\tilde{a}\tilde{b}}$. Если все множества Q_{a_i} и $Q_{b'_j}$ $(i, j = 1, \ldots, n)$ непусты, то изображение B' приемлемое.

Назовём 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, как 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 выбирать изображение B_0 с наименьшей величиной $R_{\tilde{a}\tilde{b}}$ среди приемлемых изображений, а не среди узкоприемлемых. Теперь положим, что точке a_i соответствует не одна точка, а все точки из множества Q_{a_i}, а точке b_j^0 соответствуют все точки из множества $Q_{b_j^0}$. Кроме того, теперь изображения A и B могут состоять из разного числа точек, т. е. $i = 1, \ldots, n$, $j = 1, \ldots, m$.

К изменениям в определении искомого изображения B_0 и искомого соответствия между точками A и B_0 мы пришли последовательностью шагов. Однако эти изменения можно сформулировать и независимо от этой последовательности. Пару изображений (A, B') охарактеризуем величиной $l(B')$, которая есть наименьшее такое число, что каждая точка одного изображения имеет на расстоянии, не большем $l(B')$, хотя бы одну точку другого изображения. Расстоянием $R(A, B)$ между изображениями A и B назовём наименьшую из величин $l(B')$, полученных для всех B' из B^*. Изображение B_0, на котором этот минимум достигается, назовём искомым. Взаиморасположение A и B_0 искомое. Каждой точке a_i из A (каждой точке b_j^0 из B_0) поставим в соответствие все те точки Q_{a_i} из B_0 (все те точки $Q_{b_j^0}$ из A), которые находятся от неё 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 точками и при новом определении должно быть строго биективным. Однако это ограничение нетрудно снять рассмотрением довольно очевидных вырожденных подслучаев в рамках рассмотренных первого, второго и третьего случаев.

Однако когда точек в изображениях много (а реальное изображение, например изображение на телевизионном экране, — это сотни тысяч точек), даже полиномиальная зависимость от числа точек неприемлема или, как минимум, нежелательна. В ещё большей степени это обстоятельство применимо к глазу. Одна из особенностей зрительной информации — её огромные объёмы. Глаз человека содержит около 130 миллионов светочувствительных элементов (палочек и колбочек). Вместе с тем распознавание изображений осуществляется иногда лишь за доли секунды. Вряд ли это можно сделать, используя всю информацию на сетчатке глаза. Должно существовать некоторое «сито» и для изображений из среды, и для визуальной информации из памяти при оперировании с ними. Парадокс, однако, состоит в том, что попытки до собственно распознавания выделить на изображениях, как это иногда делается, «опорные точки», «важные детали» и пр. есть тоже распознавание, т. е. возникает замкнутый круг. Ясно, что нужно избегать этого. Поэтому сравнение изображений как таковых заменяется рассмотрением того, что можно назвать их набросками, эскизами или приближениями. Эти эскизы могут состоять из гораздо меньшего числа точек, чем исходные изображения, и вместе с тем оценка схожести между эскизами определённым образом связывается с оценкой схожести между исходными изображениями. Распознавание при таком подходе перестаёт прямо зависеть от количества точек в изображениях, и появляется возможность рассматривать изображения, бесконечные по числу составляющих их точек («непрерывные» изображения).

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

Изображения могут быть весьма схожими и даже тождественными по форме, но существенно разными по размерам. За счёт 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 [3]. И хотя в этом случае есть свои особенности, в целом данные определения и построения не слишком меняются при расширении на преобразования подобия.

Положим теперь целью расширить рассматриваемые преобразования до аффинных. У этого случая есть специфика, которая состоит содержательно в том, что, в отличие от подобных и тем более изометрических преобразований, в целом аффинные преобразования не сохраняют форму в изображениях. Очевидно, что, например, чрезмерными сжатиями и растяжениями фигуру можно сделать неузнаваемой в сравнении с оригиналом. Кроме того, априори неясно, что понимать под наиболее точным совмещением двух фигур, когда каждая из них рассматривается с точностью до аффинных преобразований.

Пусть изображение A состоит из точек $a_1, \ldots, a_n$, изображение B — из точек $b_1, \ldots, b_n$, 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 $(i = 1, \ldots, n)$. Обозначим через B^* множество всех изображений, получаемых из B аффинными преобразованиями. Полагаем, что на произвольном B' из B^* сохраняется нумерация, порождённая изображением B, т. е. 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 $(b_i\, b'_i)$ $(i = 1, \ldots, n)$ не больше ε. Преобразования, переводящие изображения из $\{B\}^\varepsilon$ друг в друга, назовём ε-аффинными. Содержательно их можно интерпретировать как некоторые ограниченные, локальные аффинные преобразования для B.

Через $l_A^\psi(B')$ обозначим длину наибольшего из отрезков $a_i\, b'_{\psi(i)}$ $(i = 1, \ldots, n)$. 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 назовём искомым изображением для изображения A, биекцию ψ_0 — искомым соответствием между точками в A и B. Величину $l_A^{\psi_0}(B_0)$ обозначим через $R_A(B)$ и назовём расстоянием от исходного изображения A до B.

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

Назовём изображение B' из B^* согласованным с A, если в B' имеется два отрезка $(b'_{j_1}\, b'_{j_2})$ и $(b'_{j_3}\, b'_{j_4})$, равные, параллельные и однонаправленные соответствующим отрезкам $(a_{i_1}\, a_{i_2})$ 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 идёт точка a_{i_1}, а затем a_{i_2}, в отрезке $(b'_{j_1}\, b'_{j_2})$ слева направо идёт сначала точка b'_{j_1}, затем b'_{j_2}.

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

Теорема 8. *Если B_0 — искомое изображение для A, то B_0 согласовано с A.*

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

Эти построения легко распространяются на трёхмерный случай, когда трёхмерное изображение (тело), представлено конечным множеством точек в трёхмерном евклидовом пространстве. Аналогичным двумерному случаю образом вводится понятие искомого изображения B_0 для заданного изображения A. И в трёхмерном случае справедлив аналог последней теоремы, только согласование изображений A и B, в отличие от плоского случая, проводится не по двум, а по трём отрезкам.

К настоящему времени созданы компьютерные реализации описанного подхода для распознавания произвольных двумерных фигур и для стереореконструкции.

Литература

[1] Козлов В. Н. О распознавании аффинно разных дискретных изображений // Интеллект. сист. — 1998. — Т. 3, № 3-4. — С. 95—122.

[2] Козлов В. Н. О зрительном образе, математических подходах к определению этого понятия и о распознавании изображений // ЖВМ и МФ. — 1999. — Т. 39, № 11. — С. 1929—1946.

[3] Козлов В. Н. Элементы математической теории зрительного восприятия. — М.: Изд-во Центра прикладных исследований при механико-математическом факультете МГУ, 2001.

[4] Kozlov V. N. Image coding and recognition and some problems of stereovision // Pattern Recognition Image Analysis. — 1997. — Vol. 7, no. 4. — P. 448—466.