Сегментация изображений и преобразования, сохраняющие форму фигур

В.Н. Козлов

Изображением в работе называем конечное (непустое) множество точек в евклидовых пространствах разной размерности. Кратко работа состоит в том, чтобы имея изображение, принадлежащую некоторому образу, извлечь из этого изображения то, что можно было бы назвать описанием этого образа.

Ключевые слова: математическое определение изображения, описание зрительного образа, распознавание изображений.

Кратко и на содержательном уровне настоящая работа состоит в том, чтобы имея изображение (фигуру), принадлежащую некоторому образу, извлечь из этого изображения то, что можно было бы назвать описанием этого образа.

Изображением называем конечное (непустое) множество точек в евклидовых пространствах разной размерности. В частности, двумерное изображение – конечное множество точек на плоскости.

Обосновываем это тем, что любую фигуру можно «аппроксимировать» конечным множеством точек , которые уже сами по себе делают фигуру вполне узнаваемой. При этом если точек много, то такая совокупность точек практически неотличима от исходной фигуры. Так же можно представлять и полутоновые, черно-бело-серые изображения, при этом разная плотность точек в разных частях изображения дает разные оттенки «серого цвета». Как известно, цветное изображение можно представлять как наложение трех монохроматических (аналогов черно-бело-серых) изображений. Это означает, что совокупностями точек можно представлять и цветные изображения. Трехмерные изображения – точки в трехмерном евклидовом пространстве. Соотнесение между трехмерным изображением и двумерными, являющимися их проекциями , приводит к задачам восстановления тел по плоским проекциям и смежным задачам. Наконец, трехмерный мир в динамике можно рассматривать как четырехмерное изображение (последовательность трехмерных сцен).

Далее рассматриваются двумерные изображения, но сказанное несложно обобщается и на случаи большей размерности.

Ранее [?, ?, ?] было введено понятие кода изображения, который можно трактовать как описание изображения с точностью до аффинных его преобразований. Это значит, что имея изображение – конкретное множество точек на плоскости – мы получаем его же описание при всех его возможных параллельных переносах на плоскости, вращениях, изменениях в размерах, сжатиях и растяжениях. Это можно рассматривать как первый шаг к формированию на основе данного изображения некоторого его обобщения. Однако шаг явно недостаточный, поскольку различия изображений в рамках одного образа явно не сводятся к только аффинным преобразованиям.

Пусть X - некоторое изображение . Точку v из X назовем внутренней, если существуют такие точки p,q,s из X, что v лежит внутри треугольника, образованного точками p,q,s (или на его стороне). Остальные точки из X называем внешними или контурными. В совокупности они есть то, что назовем капсулой $K(X)$. Случаи капсул из одной и двух точек полагаем вырожденными. Ясно, что применительно к $K(X)$ (невырожденной) можно говорить о выпуклом многоугольнике. и , соответственно, о сторонах капсулы как о сторонах в этом многоугольнике.

Обозначим через $W(X)$ выпуклую оболочку для X. Ясно, что $W(X)$ и $W(K(X))$ совпадают . Любое конечное (непустое) множество точек из $W(K(X))$ называем наполнением капсулы $K(X)$. Наполнением второго плана называем любое изображение , аффинно эквивалентное какому-либо из изображений наполнения. Ясно, что само изображение X является одним из наполнений капсулы $K(X)$.

Две капсулы A и B называем непересекающимися, если $W(A)$ и $W(B)$ не имеют общих точек.

Пусть даны капсулы A и L. Обозначим через L' капсулу L после параллельного переноса. Если капсула L такова, что ее параллельным переносом можно поместить на A так, что все точки из A принадлежат W(L’), то говорим , что L накрывает A и называем L чехлом для A.

Назовем произвольную капсулу Z (невырожденную) опорой. Определим понятие размера капсулы A по опоре Z. Пусть задано направление – прямая α . Пусть a_α длина наибольшего отрезка , вмещающегося в A и параллельного α , аналогично и соответственно обозначаем z_α в Z.

Тогда отношение a_α/z_α называем размером A по опоре Z и по направлению α. Строим величины a_α/z_α по всем направлениям. Максимум на этом множестве называем размером A (по опоре Z) и обозначаем через $r_Z(A)$.

Частный случай введенного понятия - размер вырожденной капсулы из точек x и y по опоре Z - можно трактовать как расстояние между точками x и y по опоре Z .

Пусть A – капсула, и $G(A)$ - множество всех изображений из наполнения капсулы. Расстоянием между двумя изображениями X и Y из $G(A)$ называем расстояние Хаусдорфа между ними. Для каждого X из $G(A)$ определяем множество Y изображений из $G(A)$, имеющих максимальное расстояние до X . Это расстояние обозначим через $r(X)$. Этим определена тройка $< X, \{Y\}, r(X) >$. Далее рассматриваем множество $\{< X, \{Y\}, r(X) >\}$ таких троек для всех X из $G(A)$. На этом множестве определяем минимум по $r(X)$, т.е. такую тройку $< X, \{Y\}, r(X) >$, на которой значение $r(X)$ минимально. Соответствующее X называем центром капсулы. Ясно, что X - точка и представляет собой центр описанной окружности для A. Содержательно центр капсулы можно трактовать как изображение , максимальное различие которого с остальными изображениями из наполнения капсулы минимально.

Однако капсулу и ее наполнение мы рассматриваем с точностью до аффинных преобразований. При сжатиях-растяжениях описанная окружность превратится в эллипс, и таких эллипсов можно построить бесконечное множество. Нужно выбрать какой-то один из них. Воспользуемся имеющимся результатом [5, 6] о том , что для выпуклого многоугольника существует и единственен описанный эллипс с наименьшей площадью. Центр этого эллипса и будем называть центром капсулы. При аффинных преобразованиях капсулы A с описанным эллипсом E отношение площадей сохраняется. Это значит, что для преобразованной капсулы A' преобразованный эллипс E' будет по прежнему минимальным из возможных по площади, а центр у E' - центром капсулы A' .

Пусть теперь A^+ есть набор $A_1, \ldots, A_k$, $k \geq 1$, попарно непересекающихся капсул. Капсулу L называем чехлом набора A^+, если L является чехлом для каждой капсулы в A^+. Размер набора A^+ (по опоре Z) – наибольший из размеров капсул в A^+. Наполнением набора A^+ называем совокупность наполнений каждой из капсул в A^+. Наполнение второго плана - любое изображение , аффинно эквивалентное какому-либо из изображений наполнения набора A^+.

Ясно, что набор капсул тоже есть изображение, и, с другой стороны, каждое изображение может рассматриваться как набор вырожденных (до точки) капсул.

Пусть дан набор капсул B^+, состоящий из капсул $K_1, \ldots, K_l$. Пусть выбраны капсулы с номерами i и j $(i, j = 1, \ldots, l, i \neq j)$. Склейкой называем объединение $U(K_i, K_j)$ капсул K_i и K_j и замена их на капсулу $K(U(K_i, K_j))$. Если эта капсула не пересекается с остальными капсулами , то получившийся набор капсул называем укрупнением исходного набора, а саму операцию – операцией укрупнения. Ясно , что к получившемуся укрупнению можно снова применять операцию укрупнения, и т.д. Через $B^{+\times}$ обозначим множество всех укрупнений для B^+ (включая само изображение B^+). Количество t капсул в них может меняться от 1 до l . Через $B_t^{+\times}$ обозначим подмножество множества всех укрупнений, состоящее из наборов с t капсулами.

По сути , футляр A^+ - это своеобразный прогноз, означающий то, что на месте A^+ может оказаться любое изображение из наполнения этого футляра (в том числе и некоторые другие футляры, и исходное изображение A). Футляр $_2^+$ назовем сужением футляра $_1^+$, если каждое изображение из наполнения $_2^+$ аффинно эквивалентно с некоторым изображением из наполнения футляра $_1^+$.Ясно, что для этого принадлежать наполнению футляра $_1^+$ должно либо изображение $_2^+$, либо некоторое его укрупнение.

Пусть H - множество всех изображений из наполнения набора капсул A^+. Максимум различий между изображениями из H будет определяться в целом тем , насколько большим может быть расстояние между точками изображений в одной капсуле, то есть, по сути, размером набора A^+ - чем больше размер A^+, тем в целом более различающимися могут быть изображения в наполнении набора A^+. Это можно трактовать так, что «обобщенная форма» фигуры таким набором капсул задана менее точно, в сравнении с набором, у которого размер меньше.

Для дальнейшего нам понадобятся некоторые понятия и ранее опубликованные результаты [**?**, **?**, **?**]. Содержательный смысл их в том, чтобы обеспечить наложение одного изображения на другое аффинными преобразованиями так, чтобы они "почти совпадали т.е. чтобы различие между изображениями было бы минимальным из возможных.

Пусть изображение A состоит из точек $a_1, \ldots, a_n$, изображение – из точек $b_1, \ldots, b_n$, ψ - одно из возможных взаимно однозначных соответствий между точками изображений A и , которым точке a_i из A сопоставляется точка $b_{\psi(i)}$ из $(i = 1, \ldots, n)$. Обозначим через B^* множество всех изображений, получаемых из аффинными преобразованиями. Полагаем, что на B' из B^* сохраняется нумерация, порожденная изображением , т.е. через b'_i на B' обозначается точка, в которую переходит при соответствующем преобразовании точка b_i из . 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 , назовем ε -аффинными. Содержательно их можно трактовать как ограниченные, локальные аффинные преобразования для B.

Дадим важное определение искомого (или оптимального) взаиморасположения: через $l_A(B')$ 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 $l_A(B')$ 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-аффинными преобразованиями пары (A, B_0) как целого, выполняется аналогичное свойство: для всех B'' из $\{B'_0\}^{\varepsilon_1}$ и при всех биекциях ψ минимум величин $l_{A'}(B'')$ достигается на изображении B'_0 и при биекции ψ_0 . Тогда B_0 называем искомым для изображения A (и взаиморасположение A и B_0 искомым), биекцию ψ_0 - искомым соответствием между точками в A и .

Нетрудно видеть, понятие оптимального взаиморасположения можно рассматривать, как строящееся, в некотором приближении, на основе понятия расстояния Хаусдорфа между множествами.

Из полученных ранее результатов [**?**] следуют некоторые необходимые условия , которым должны удовлетворять изображения в паре (A, B_0) . Они коротко состоят в следующем . Пусть биекцией ψ точке a_i из A сопоставляется точка $b'_{\psi(i)}$ из $(i = 1, \ldots, n)$. Возьмем на плоскости произвольную точку O и параллельными переносами отрезка $a_i b'_{\psi(i)}$ совместим точку a_i с точкой O. Точку, в которую перейдет при этом $b'_{\psi(i)}$, обозначим через $c_{i\psi(i)}$ и назовем порожденной парой соответствующих точек a_i и $b'_{\psi(i)}$. Изображение из точек $c_{i\psi(i)}$ $(k = 1, \ldots, n)$ называем характеристическим , O – центр характеристического изображения, $c_{i\psi(i)}$ - точки ядра. Окружность наименьшего по радиусу круга, включающего все точки ядра, называем ключевой. Доказано [**?**], что для пары (A, B_0)

центр характеристического изображения с необходимостью должен совпадать с центром ключевой окружности.

Назовем изображение B' из B^* согласованным с A, если существуют в B' два непараллельных отрезка $(b'_1 b'_2)$ и $(b'_3 b'_4)$, равные , параллельные и однонаправленные с соответствующими отрезками $(a_1 a_2)$ и $(a_3 a_4)$ в A. Параллельные отрезки , например , $(a_1 a_2)$ и $(b'_1 b'_2)$ называем однонаправленными , если, при условии, что $(a_1 a_2)$ слева направо сначала идет точка a_1 , а затем a_2 , то и в отрезке $(b'_1 b'_2)$ слева направо сначала идет точка b'_1 , затем b'_2 .

Доказано , что если B_0 искомое изображение для A ,то B_0 согласовано с A.

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 позволяет вычленить из $*$ конечное подмножество изображений, среди которых только и может находиться искомое изображение B_0 .

Действительно, согласовывать B' с A можно по разным парам отрезков. Пусть , например, задана пара отрезков $(a_1 a_2)$ и $(a_3 a_4)$ и эти отрезки равны параллельны и однонаправленны с соответствующими отрезками в B' , т.е. с отрезками $(b'_1 b'_2)$ и $(b'_3 b'_4)$. 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 [**?**] , среди них существует и единственно такое – обозначим его через B_{1234} – для которого в паре с A центр ключевой окружности совпадает с центром характеристического изображения.

Отрезки $(a_1 a_2)$ и $(a_3 a_4)$ можно выбрать в A не более чем $(C_n^2)^2$способами. Соответственно не больше будет и изображений B_{1234}. Множество их обозначим через $(A \mid \{B\})_\psi$. Объединим множества $(A \mid \{B\})_\psi$ для всех ψ. Это и будет множеством, в котором находится искомое изображение, обозначим его как $(A \mid \{B\})$, и назовем множеством изображений B, потенциально искомых для A . Заметим, что так определенное $(A \mid \{B\})$ выглядит требующим для своего построения конечного, но довольно большого перебора (главным образом за счет большого числа биекций ψ). Однако этот перебор можно существенно сократить [**?**]. И здесь уместно сделать некоторое общее замечание: построения тут и далее по тексту предполагают иногда конечный, но большой перебор. Почти всегда его можно кардинально уменьшить, однако здесь мы этим не занимаемся, оставляя на будущее, и довольствуясь здесь конечностью.

Отметим , что в частном случае A может быть, конечно, аффинно эквивалентным с B.

Далее пошагово опишем процедуру, которая для конечной совокупности наборов капсул устраивает на этой совокупности расширения наборов капсул , повышаюшие их «вместимость», а затем «чистку» совокупности, оставляющую минимально необходимую совокупность наборов капсул .

I. Пусть даны два набора капсул A^+ из капсул $A_1, \ldots, A_n$ и B_+ из капсул $B_1, \ldots, B_n$, и чехол L для A^+ . Называем здесь A^+ подложкой, а изображение B_+ трактуем как предназначенное для наложения аффинными преобразованиями на подложку A^+ . Полагаем, что в изображения A^+ и B_+ включены центры их капсул. Это точки соответственно $a_1, \ldots, a_n$ – обозначаем их совокупность через a_+ - и $b_1, \ldots, b_n$ – обозначаем через b_+ . Далее строим множество $(a^+ \mid \{b^+\})$ преобразованных изображений b_+ , потенциально искомых для a_+ . Изображение a_+ является частью изображения A^+ . Каждому из изображений в $(a^+ \mid \{b^+\})$ соответствует некоторым образом преобразованное изображение B^+ , частью которого оно является. Тем самым , множеству $(a^+ \mid \{b^+\})$ можно сопоставить множество $V(A^+ \mid \{B^+\})$ преобразованных изображений B^+ , по разному расположенных на изображении A^+ .

Возьмем теперь произвольный набор капсул B'^+ из $V(A^+ \mid \{B^+\})$, состоящий из капсул $B'_1, \ldots, B'_n$ причем капсула $B'_{\psi(i)}$ сопоставлена капсуле A_i $(i = 1, \ldots, n)$. Для каждой пары капсул A_i и $B'_{\psi(i)}$ рассматриваем их склейку $U(A_i, B'_{\psi(i)})$ и капсулу $K(U(A_i, B'_{\psi(i)}))$, которую обозначаем через K_i . Если каждая из капсул K_i накрывается чехлом L для A^+ и все капсулы K_i попарно не пересекаются, то набор K_i капсул называем приемлемым расширением подложки A^+ (за счет объединения с набором B'^+ и обозначаем через $(A^+ + B'^+)$ (где B'^+ из $V(A^+ \mid \{B^+\})$.

Трактовка и особенность построенного набора капсул $(A^+ + B'^+)$ в том, что и A^+ , и B'^+ принадлежат наполнению этого набора. Если рассматриваем два любых изображения, являющиеся наполнениями для A^+ и B^+ , то эти A^+ и B^+ можно считать определяющими некоторое деление на «куски» этих изображений, а биекция ψ - соответствие между кусками.

Далее строим приемлемое расширения вида $(A^+ + B'^+)$ подложки A^+ для каждого набора B'^+ из $V(A^+ \mid \{B^+\})$ (если такое расширение есть). Результат процедуры - множество $\{(A^+ + B'^+))\}$ приемлемых расширений подложки A^+ , и множество ψ соответствий между капсулами в A^+ и B_+ .

Если множество $\{(A^+ + B'^+))\}$ не пустое, то B_+ называем вложимым в A^+ , в противном случае – не вложимым.

II. Мы рассмотрели пару изображений A^+ и B_+. Теперь – некоторое обобщение: пусть вместо одного изображения B_+ имеем некоторую их совокупность $B_1^+, \ldots, B_k^+$, где $k \geq 1$. Для каждого B_i^+ $(i = 1, \ldots, k)$ строим множество $\{(A^+ + B_i'^+)\}$, и пусть Q – объединение всех этих множеств. Далее рассматриваем подмножество q множества Q , обладающее следующими свойствами:

1) каждая расширенная подложка из q состоит из капсул $K_1, \ldots, K_n$. Для каждого s $(s = 1, \ldots, n)$ объединяем (склеиваем) капсулы K_s из всех подложек из q и обозначаем их объединение через K'_s , через K_s^q обозначаем капсулу $K(K'_s)$. Пусть при этом каждая из капсул K_s^q накрывается чехлом L подложки A и капсулы попарно не пересекаются. Тогда подмножество q называем правильным.

2) пусть для любого q' такого, что q является его собственным подмножеством, это q' правильным уже не является. Тогда q называем правильным и полным. Набор капсул $K_1^q, \ldots, K_n^q$ называем расширением исходного набора A^+ , порожденным множеством q, и обозначаем через A_q^+ , а те из $B_1^+, \ldots, B_k^+$, которые в виде соответственно преобразованных $B_i'^+$ вложимы в соответствующие расширенные подложки A'^+ из q, называем вложимыми в A^{+q} .

Итогом описанных процедур для A^+ и $B_1^+, \ldots, B_k^+$ являются расширения $A_1^{+q}, \ldots, A_p^{+q}$ исходного набора капсул A^+ , построенные для каждого возможного правильного и полного подмножества q, и множества $\{B^+\}_1, \ldots, \{B^+\}_p$ вложимых в эти расширения изображений из совокупности $B_1^+, \ldots, B_k^+$. Подчеркнем, что , по построению, для каждого расширения A_i^{+q} ,все наборы капсул из $\{B^+\}_i$ $(i = 1, \ldots, p)$ принадлежат наполнению этого расширения.

Напомним, что последовательность рассмотрений была такой: сначала рассмотрели подложку A^+ и изображение B_+ , которое разными способами пытаемся вложить, «вместить» в A^+ . Если «точного» вложения нет , то расширяем исходный набор капсул A^+ по некоторым правилам, как бы «склеивая» B_+ с A^+ . Итогом этой процедуры была совокупность по разному расширенных A^+ . Затем, следующим шагом, для подложки A^+ мы рассмотрели уже не одно изображение B_+ , а некоторую их совокупность $B_1^+, \ldots, B_k^+$, и которые тоже , как в предшествующей процедуре, мы пытаемся по разному вместить в A^+ . Промежуточным итогом этой процедуры является множество Q по разному расширенных исходных подложек A^+ , с указанием того, с помощью каких из изображений $B_1^+, \ldots, B_k^+$ это расширение получено. Затем берем подмножество q множества Q со свойствами правильности и полноты: правильность означает, что все наборы в q можно «покапсульно склеить», и в получившемся наборе капсулы будут попарно непересекающимися и вмещающимися (каждая) в чехол для исходного изображения A^+ ; полнота означает, что q «максимален», т.е. его нельзя расширить, и при этом сохранить правильность. В итоге все возможные q дают совокупность m расширенных наборов капсул $A_1^{+q}, \ldots, A_p^{+q}$, с информацией о том, какие из изображений $B_1^+, \ldots, B_k^+$ вложимы в каждый из этих расширенных наборов капсул.

III. Предыдущий шаг – это рассмотрение подложки A^+ и совокупности изображений $B_1^+, \ldots, B_k^+$. Следующий шаг – рассмотрение совокупности наборов капсул $A_1^+, \ldots, A_k^+$. Поочередно берем каждый из этих наборов в качестве подложки , а все $A_1^+, \ldots, A_k^+$, включая выбранную подложку, в качестве $B_1^+, \ldots, B_k^+$. Для каждого такого случая строим совокупность m расширенных наборов капсул с информацией о том, какие из изображений $A_1^+, \ldots, A_k^+$ вложимы в каждый из этих расширенных наборов капсул. Далее рассматриваем множество M – объединение всех совокупностей m. Для каждого A^+ из M через A^{++} обозначаем множество тех из $A_1^+, \ldots, A_k^+$, что вложимы в A^+ .

IV. Заключительный шаг можно условно назвать чисткой множества M. Дело в том, что в M может быть много «лишних» изображений, например, за счет возможной повторяемости изображений в исходном наборе $A_1^+, \ldots, A_k^+$.

Рассматриваем такие совокупности изображений A^+ из M, у которых объединение их множеств A^{++} содержит исходный набор изображений $A_1^+, \ldots, A_k^+$. Среди этих совокупностей выбираем совокупность с наименьшей мощностью (если таких совокупностей не одна, то любую из них). Обозначим эту совокупность через M^*, она и есть искомая. Это есть минимальное множество расширенных наборов капсул таких, что в них вкладываются все изображения исходного набора.

Описанную процедуру обозначим как PROC , исходными для нее являются совокупность наборов капсул $A_1^+, \ldots, A_k^+$ и их чехлов $L_1, \ldots, L_k$ Результат процедуры – 1) совокупность расширенных наборов капсул $A_1^{+q}, \ldots, A_u^{+q}$, 2) для каждого из A_i^{+q} $(i = 1, \ldots, u)$ - множество A_i^{++} тех из исходных изображений $A_1^+, \ldots, A_k^+$, что вложимы в A_i^{+q} , и 3) чехлы для каждого из $A_1^{+q}, \ldots, A_u^{+q}$, обозначим их как $L_1^{+q}, \ldots, L_u^{+q}$.

Пусть теперь дано изображение A (например, некоторая фигура). Это набор вполне конкретных точек на плоскости. Наша цель – создать на базе изображения A такое его описание, которое поднималось бы до описания «образа» A . Первыми шагами в этом направлении было, как отмечалось выше, создание кода изображения A , который определяет изображение с точностью до аффинных преобразований, т.е. независимо от конкретного места на плоскости, от вращений, изменений в размерах, сжатий и растяжений. Однако этого недостаточно, ибо, ясно, к образу A надо относить и изображения, у которых в сравнении с A , например , сделаны некоторые локальные трансформации, и изображения, отличающиеся от A числом точек, и т.д. Мы используем для описания образа A набор капсул A^+ . Капсулы в A^+ полагаем образованными точками из A и включающими как наполнение все точки из A, т.е. это покрытие. Тогда относить (на этом шаге) к образу A можно все изображения из наполнения A^+ (и наполнение второго плана, т.е. все изображения , аффинно эквивалентные наполнению). Сразу вырисовываются ограничения, которые разумно наложить на A^+ : нужна как можно большая одинаковость капсул по размерам. Действительно, капсулы соответствуют разбиению исходного A на «куски», и это разбиение представляет форму изображения A, если «куски» равновелики. В противном случае, например , если есть один огромный «кусок», накрывающий практически все изображение, а остальные «куски» – какие-то точки на периферии, то такое разбиение трудно назвать представляющим форму исходного изображения. Кроме того, как отмечено в замечании выше, чем больше размер наибольшей капсулы в наборе, тем , при том же количестве капсул , больше будут разниться изображения в наполнении, т.е. тем менее «определенную» форму они в совокупности, можно считать, задают.

В качестве опоры для измерения размеров капсул в A^+ используем капсулу $K(A)$ (т.е. «накрывающую» все точки изображения A). Это своеобразная граница изображения A. У этой капсулы есть и та особенность, что какой бы набор капсул A^+ для изображения A мы не использовали, в этом A^+ есть все точки, составляющие $K(A)$.

Будем полагать, что форма чехлов для капсул есть форма капсулы $K(A)$, т.е. каждый чехол представляет собой уменьшенную с надлежащим коэффициентом и с сохранением подобия капсулу $K(A)$. Тогда минимальным возможным коэффициентом для чехла произвольной капсулы , является, нетрудно видеть, ее размер.

Капсул для A – конечное множество . Тем самым возможных размеров капсул – тоже конечное множество . Выстроим эти числа (размеры) в порядке возрастания : $r_0, r_1, \ldots, r_t$. Здесь, очевидно, всегда $r_0 = 0$, и $r_t = 1$. Из этих капсул выстраиваем все возможные наборы A^+ капсул для изображения A (т.е. из попарно непересекающихся капсул и накрывающих изображение A). Обозначим это множество через $W(A)$. Поскольку размер набора капсул – это размер наибольшей капсулы в нем, то все возможные размеры наборов капсул для изображения A характеризуются той же цепочкой чисел $r_0, r_1, \ldots, r_t$, обозначим ее через r^*.

Множество $W(A)$ – все возможные наборы капсул для A, с разным количеством составляющих набор капсул , и с разными по размерам капсулами в одном наборе. Эти две характеристики – число капсул в наборе и степень разброса капсул по размерам – можно связать в рамках некоторой процедуры, которую можно трактовать, как балансировку. Обозначим через $(W(A), r_i)$ подмножество множества $W(A)$, состоящее из тех наборов капсул, размеры которых не превышают r_i , где r_i из r^*. Далее на $(W(A), r_i)$ рассмотрим все наборы с минимальным числом капсул. Если это число обозначить через n_i , то соответствующее подмножество обозначим через $(W(A), r_i, n_i)$. Такие наборы капсул трактуем как сбалансированные по количеству (капсул) . Ясно, что n_i может принимать значения от 1 до n , где n – число точек в изображении A. Объединим теперь множества $(W(A), r_i, n_i)$ для всех n_i $(n_i = 1, \ldots, n)$ и обозначим через $\{(W(A), r_i, n_i)\}$. Это множество всех сбалансированных по количеству капсул наборов капсул изображения A. Разобьем его на подмножества $F_1, \ldots, F_n$, в каждом F_i содержатся наборы капсул с i капсулами. Далее в каждом F_i оставляем только наборы с наименьшим размером – называем это балансировкой по размеру. Оставшиеся в $F_1, \ldots, F_n$ наборы капсул сбалансированы по количеству и по размерам .

Поскольку в F_i могут быть футляры мало отличающиеся и вложимые в общее для них расширение, то к F_i применим процедуру чистки Proc , введенную выше, и обозначим результат через F_i^* . Изображения из F_i^* (i=1,…,n) называем футлярами для A. Ясно что в F_1^* всего одно изображение, и это капсула $K(A)$, в F_n^* тоже одно изображение, и это изображение A, в F_2^* изображения состоят из двух примерно равных половинок, и т.д

По сути , футляр A^+ - это своеобразный прогноз, означающий то, что на месте A^+ может оказаться любое изображение из наполнения этого футляра (в том числе и некоторые другие футляры, и исходное изображение A). Футляр A_2^+ назовем сужением футляра A_1^+ , если каждое изображение из наполнения A_2^+ аффинно эквивалентно с некоторым изображением из наполнения футляра A_1^+ .

Далее с использованием $F_1^*, \ldots, F_n^*$ построим ярусный граф G_A , который в целом трактуем как очередное приближение к понятию образа , которому принадлежит исходное изображение A. Ярусу с номером i соответствует F_i^* , вершины яруса – футляры из F_i^* . Нижний ярус в графе (с номером 1) состоит всего из одной вершины, соответствующей футляру с одной капсулой $K(A)$, верхний ярус (с номером n) состоит тоже из одной вершины, соответствующей футляру с n капсулами, то есть исходному изображению A . В ярусе с номером i (i=1, ... ,n) вершинам соответствуют футляры, состоящие из i капсул. От каждой вершины в $n-1$ ярусе проводим ребро к вершине в n-ом ярусе. Это обозначает то, что футляр (единственный) из F_n с некоторым укрупнением вкладывается в каждый из футляров в $F_n - 1$. Действительно, исходное изображение A является очевидным сужением для футляров в $F_n - 1$, как, впрочем , и для любых футляров в $F_1^*, \ldots, F_n - 1^*$.

Ребра в графе G_A строим по индукции. Базис индукции – уже построенные ребра между вершинами в ярусах с номерами $n-1$ и n . Ребра считаем направленными, с направлением от вершин в ярусах с меньшими номерами к вершинам в ярусах с большими номерами. Между вершинами в одном ярусе ребер нет.

Пусть теперь уже есть ребра между вершинами в ярусе с номером i (i=2, ... , n-1) и вершинами во всех ярусах с номерами большими i . Строим ребра между вершинами в ярусе с номером $i-1$ и вершинами в ярусах с большими номерами. Это тоже индукция, по k , $k = i, \ldots, n$. Пусть x – одна из вершин яруса $i-1$ и ей соответствует футляр X ,y – одна из вершин яруса k и ей соответствует футляр Y. Если имеется путь из вершины x в вершину y, то ребро между ними не проводим. Если нет , то для Y определяем множество $\{Y^\times\}_i - 1$. Напоминаем – это множество всех укрупнений футляра Y, состоящих из $i-1$ капсул. Футляр X рассматриваем как подложку, наборы капсул из $\{Y^\times\}_i{-}1$ - как налагаемые изображения. Если хотя бы один Y' из $\{Y^\times\}_i - 1$ вложим в X , то проводим ребро между вершинами x и y. Описанное проделываем для всех x из яруса i , и всех y из яруса k.

Граф G_A определен. Он и есть, в некотором приближении, описание (полное) образа, извлекаемое из всего лишь одного изображения A, принадлежащего этому образу. При этом граф в целом скорее можно трактовать как набор описаний, а одним таким описанием считать путь из условного корня графа – вершины, соответствующей футляру из одной капсулы – в вершину с футляром, представляющим собой исходное изображение A. Чем больше вершин на этом пути, тем более полным можно полагать такое описание образа.

Трактовку такому пути, и соответственно, цепочке футляров , можно дать следующую. Начало пути – футляр из всего одной капсулы. Это максимальное обобщение исходного изображения A, настолько максимальное , что в наполнении этого футляра, нетрудно видеть , находятся любые изображения. Конец пути – максимальная конкретика , т.е. само изображение A (напоминаем, рассматриваемое с точностью до аффинных преобразований), Это значит , что в этот «футляр», т.е. в изображение A, «вкладываются» только изображения, аффинно эквивалентные с A. Пусть теперь x и у из ярусов с номерами i и j $(i < j)$– любые две соседние вершины на этом пути. Им соответствуют футляры X и Y. При этом футляр Y с j капсулами, укрупненный до i капсул , вкладывается в подложку X (возможно, в некоторое ее расширение, ограничиваемое чехлом подложки). Это значит, что футляр Y и все его наполнения можно трактовать как получаемые из наполнения футляра X с последующим его дроблением. В этом смысле Y не противоречит X, и есть его сужение, его «конкретизация».

Отметим, что многое из процедур , описанных выше, можно делать, основываясь не на собственно изображении A , а только на его коде [**?**], определяющем изображение с точностью до аффинных преобразований: это и определение внутренних и контурных точек, и построение капсул и наборов капсул , и пр.

Список литературы

[1] В.Н. Козлов, *Введение в математическую теорию зрительного восприятия*, Издательство Центра прикладных исследований при механико-математическом факультете МГУ, М., 2007.

[2] В.Н. Козлов, "О распознавании аффинно разных дискретных изображений", *Интеллектуальные системы*, **2**, 1998, 95-122.

[3] В.Н. Козлов, "О зрительном образе, математических подходах к определению этого понятия и о распознавании изображений", *Журнал вычислительной математики и математической физики*, **39**, 1999, 1929-1946.

[4] V.N. Kozlov, "Visual Pattern and Geometric Transformation of Images", *Pattern Recognition and Image Analysis*, **10**, 2000, 321-342.

[5] Л. Данцер , Б. Грюнбаум , В. Кли, *Теорема Хелли и ее применения*, Издательство «Мир», М., 1968.

[6] В.Л. Загускин, "Об описанных и вписанных эллипсоидах экстремального объема", *Успехи математических наук*, **13**, 1958, 89-93.

Image segmentation and shape-preserving transformations

V.N. Kozlov

The image in the paper is a finite (non-empty) set of points in Euclidean spaces of different dimensions. Briefly, the work consists in having an image belonging to a certain image to extract from this image what could be called a description of this image.

Keywords: mathematical definition of the image, image recognition.