О трех начальных приближениях к формальному определению визуального образа в произвольной визуальной среде

В. Н. Козлов[1]

Распознавание визуальных образов — одна из центральных задач для интеллектуальных систем. Продвижению здесь математических, теоремных методов исследования в немалой степени мешает отсутствие полного и приемлемого формального определения понятия визуального образа в произвольной визуальной среде. В работе представлена идея подхода к такому определению последовательными приближениями, и описаны три приближения.

Ключевые слова: визуальный образ, распознавание изображений, аффинные преобразования.

1. Введение

Подавляющая часть решений (успешных) в рамках компьютерного зрения сделана на основе эвристики — здравого смысла и изобретательности применительно к частным особенностям конкретной значимой для практики задачи. Приемлемой общей, глубокой и математизированной теории при этом не возникает, и во многом потому, что нет принятого в полной мере определения зрительного образа. Интуитивное понимание того, что есть зрительный образ, существует, а формального, математического определения нет. На том же содержательном, интуитивном уровне ясно, что зрительный образ — понятие сложное, многоплановое и многоуровневое. Это не одиночная, конкретная картинка, хотя и в этой картинке присутствуют черты образа, которому принадлежит картинка, и, в общем случае, может быть, и не единственного. Образы могут пересекаться, включать один другой, содержать в себе то, что можно назвать подобразами, а это выстраивает их в сложную систему взаимозависимостей. К тому же возникают образы из взаимодействия с окружающей средой, и алгоритмы такого возникновения — во многом загадка. Переход к доказательному, теоремному уровню исследования образов упирается, тем самым, в отсутствие приемлемого формального определения образа. Вряд ли можно решить задачу построения такого определения некоторым еди-

[1] *Козлов Вадим Никитович* — профессор каф. математической теории интеллектуальных систем мех.-мат. ф-та МГУ, e-mail: vnkozlov@mail.ru.

Kozlov Vadim Nikitovich — professor, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Mathematical Theory of Intellectual Systems.

новременным усилием. Здесь предприняты некоторые начальные шаги в этой задаче.

2. Об определении изображения и об идее подхода к определению образа

Изображением называем конечное (непустое) множество точек на плоскости. Обосновываем это тем, что любую фигуру можно «аппроксимировать» конечным множеством точек (рис. 1), которые уже сами по себе делают фигуру вполне узнаваемой.

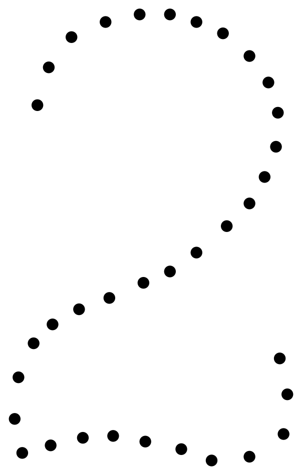

Рис. 1.

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. 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.

Частью изображения A называем любое непустое подмножество B его точек. Любое изображение B', аффинно эквивалентное изображению B, называем подизображением изображения A.

Среда S у нас — произвольное изображение, состоящее из $N(N \geq 1)$ точек. Число частей среды — $\left(2^N - 1\right)$. Множество частей (т.е. изображений) обозначим через S'. Образ из среды S будем трактовать как некоторую группу изображений из S', т.е. задавать его перечислением всех возможных примеров изображений образа в данной среде S. Обозначим множество всех возможных групп через $S^\star$. Разумеется, далеко не все из этих групп — образы в содержательном понимании. Тем самым, задача состоит в том, чтобы «ужать» множество $S^\star$ до совокупности групп таких, которые уже можно с приемлемой степенью убедительности трактовать как образы. В этом и состоит подход к определению понятия образа в данной работе. Приближений к понятию образа предполагается несколько, в этой статье описано первое, второе и третье (отдельный образ, полный образ и замкнутый образ).

Ясно, что при таком подходе визуальный образ (и система образов) зависят от среды. Содержательная интерпретация этому была дана еще в давней работе [3]. В дальнейшем могут возникнуть интересные вопросы, связанные как с вариациями в системах образов, возникающих при средах разного типа, так и с константными особенностями такого рода систем.

3. Отдельная группа (отдельный образ)

Введем сквозной для последующего изложения пример: представим для наглядности среду S как «хаос» на плоскости разных фигур — цифр, букв и пр. — по разному расположенных, разных по размерам, ориентации, пересекающихся, имеющих общие части и пр. Тогда среди групп множества $S^\star$ будут как «осмысленные», т. е. группы, например, «двоек», или «троек», так и «бессмысленные» — сочетания одновременно и «двоек», и «троек», их частей, и многого другого. Вот эти бессмысленные сочетания и надо отсеять, основываясь на некоторых далее вводимых принципах.

Перенумеруем точки среды S с единственным условием: разные точки — разные номера. Полагаем, что других точек изображений, кроме точек среды S, на плоскости нет, т.е., в частности, копий изображений из S' не создается. Каждое изображение из множества S' может быть подвергнуто аффинным преобразованиям, но при этом всегда на плоскости присутствуют только N точек (возможно, частью преобразованных) исходного изображения S.

Два изображения назовем непересекающимися или отдельными, если у них нет общих точек (т.е. пересечение множеств номеров точек этих двух изображений пусто).

Пусть G — произвольная группа из $S^\star$, и в ней есть пара изображений A и B с общей непустой частью x. Тогда при аффинном преобразовании изображения A либо аффинно преобразуется и его часть x (отдельно от остальных точек изображения B, тем самым B как таковое исчезает), либо A и B преобразуются (исключая частные вырожденные случаи) как совокупное единое изображение. И то, и другое с содержательной точки зрения не приемлемо. Отсюда возникает требование попарной непересекаемости изображений в группе, такие группы назовем отдельными. Множество всех отдельных групп обозначим через S^+, это и есть первое сокращение множества $S^\star$ всех групп.

В последующем будем рассматривать аффинные преобразования изображений из среды S. При таких преобразованиях часть точек изображения A может совпасть с частью точек изображения B. Такие точки называем кратными, сохраняем приписанными кратной точке номера слившихся точек, и при аффинных преобразованиях изображений A и B считаем их преобразующимися независимо друг от друга. На исходном изображении S кратных точек нет, после некоторых преобразований изображений среды они могут появиться. Пусть при этом исходная S преобразована в некоторую S''. Считаем, что у нас есть операция, которая, при необходимости, восстанавливает по S'' обратными преобразованиями исходную среду S.

4. Содержательное (неформальное) построение понятия футляра для группы изображений

Дано изображение X. Обозначим через $W(X)$ выпуклую оболочку для X. Известно [5], что для выпуклого множества $W(X)$ существует единственный наименьший по площади эллипс, его вмещающий. Центр этого эллипса назовем центром изображения X, а длину большей оси — размером (для вырожденного случая, когда все точки из X расположены на прямой, эллипс с очевидностью превращается в отрезок прямой, центр — середина отрезка).

Два изображения A и B называем независимыми, если $W(A)$ и $W(B)$ не имеют общих точек.

Множество всех частей изображения X обозначим через X'. Рассматриваем покрытие P_X изображения X попарно независимыми частями из X' (такое всегда есть). Части называем также кусками изображения X. Размером покрытия называем размер наибольшего куска в нем. Ясно, что возможных покрытий — конечное множество.

Уместно сделать следующее пояснение к дальнейшему. Построение модели — это всегда нечто такое, что первоначально возникает главным образом на основе интуиции, правдоподобных рассуждений. А уже затем на этой первоначальной основе строится совокупность определений (формальных, или, для начала, полуформальных), позволяющих проводить доказательные рассуждения, т.е. получать утверждения (теоремы) о свойствах модели. Эвристика, в немалой степени присутствующая в распознающих системах — это тоже правдоподобные рассуждения. Но она, как правило, только ими и ограничивается.

Далее представлены два пункта содержательных соображений, на которых в значительной мере основывается модель.

Понятие остова для отдельного изображения и для группы изображений. Предположим, в качестве гипотезы, что для группы похожих изображений во всех них есть нечто общее, объединяющее эти изображения, и что мы назовем их остовом. Опишем возникновение этого понятия на содержательном уровне в нескольких шагах-приближениях, начиная с простого и довольно очевидного. Представим «двойку» в виде совокупности точек с кругами (рис. 2). Это основа, базовое изображение A. Трактуем его так: исходная

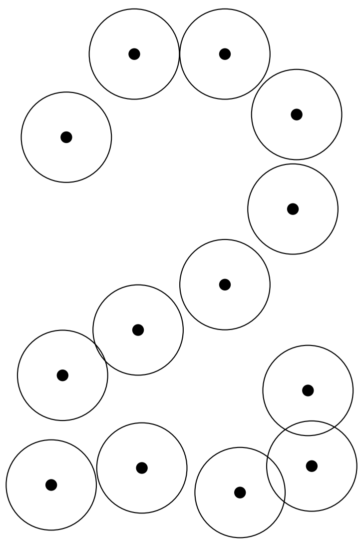

Рис. 2.

совокупность точек разбита на куски, причем такие, что тот единственный и наименьший по площади эллипс, который вмещает кусок, есть круг, причем все круги — одного радиуса, и их центры и образуют «двойку», которую мы видим на рисунке. Это несколько искусственный пример, потому что если реальную «двойку» разбить на куски, то минимальные по площади эллипсы этих кусков вовсе не обязательно будут именно кругами. Круги могут — немного — пересекаться. Вписанные же в круги куски (они на рисунке не представлены), как и следует из определений выше, пересекаться не могут. Предполагается, что при фиксированном радиусе количество кругов наименьшее из возможных. Заменим теперь содержащееся в каждом круге множество точек (на рис. 2, напомним, они не показаны) на другое, но с тем же условием: выпуклая оболочка на этом множестве порождает кусок, вписанный в круг, причем центр куска совпадает с центром круга, куски из разных кругов, естественно, не пересекаются. Это второе изображение, ясно, отличается от первого. Аналогично можно породить целый класс изображений, обозначим их $A_1, ..., A_k$ (потенциально бесконечный), они разные (и разнообразные), поскольку в круги можно «вставлять» разные изображения, даже, например, те же «двойки», только маленькие по размерам. Изображение из центров кругов для каждого из изображений можно назвать остовом, он общий для всех изображений класса, и, полагаем для наглядности на данном этапе, присутствует явно в виде точек в каждом изображении, хотя в более общем случае точки остовов можно считать особыми, вспомогательными точками, не обязательно совпадающими с какими-то собственно точками изображений. Весь класс порожденных таким образом изображений можно рассматривать как определяющий преобразования для этих конкретных изображений более широкие, чем аффинные преобразования: каждая точка в пределах круга может быть преобразована в любую другую точку, но в пределах того же круга. Если в остове m точек, то его можно рассматривать как приближение m-го уровня для изображений класса (соответственно называть m-остовом). Остов есть «выжимка» из каждого такого изображения, и эта выжимка у них одинаковая (точнее: остовы аффинно эквивалентны, а в таком представлении, как на рис. 2, просто совпадают). Полагаем, что описанные изображения класса можно, с некоторой степенью убедительности, считать относящимися к одному образу. Сразу отметим, что если в исходном изображении n точек, то можно потенциально рассматривать классы изображений, порождаемых уровнями, начиная от 1 и до n.

Описанное было первым шагом. В следующем приближении можно расширить класс изображений m-го уровня, считая, что в любой из кругов базового изображения (рис. 2) может помещаться любое множество точек, не обязательно с центром именно в центре круга, но, конечно, внутри круга и без пересечений с другими кусками того же изображения. Если исходный рис. 2 рассматривать как футляр, то такие новые изображения можно трактовать как наполнение футляра, и точки наполнения, по сути, ограничены только тем условием, что они находятся внутри кругов. Полагаем, однако, что если есть набор кусков A с m-остовом a (базовое изображение, рис. 2), и набор кусков B с m-остовом b, то a и b находятся в оптимальном взаиморасположении, то есть точки изображений a и b максимально «придвинуты» друг к другу (в предыдущем, более частном случае, эти остовы просто совпадали). Этим исключается, например, такой случай, когда точки из a и b не совпадают, но только потому, что сдвинуты по отношению друг к другу параллельным переносом. Итак, куски изображения B из наполнения у нас уже не обязательно равновелики по размерам (ранее равновеликость имела место по причине вписанности кусков в круги одинакового радиуса на рис. 2). Но для кусков исходного изображения A это условие пока сохраняется. Равновеликость важна, ибо в противном случае разбиение трудно считать представляющим форму исходного изображения. Отойти от условия равенства по размерам кусков (для A), сохраняя смысл равновеликости, можно так: пусть разбиение на куски с центрами $a_1, ..., a_m$ таково, что существует R такое, что каждый кусок находится внутри круга радиуса R (и с центрами в $a_1, ..., a_m$), и круги эти не пересекаются (это условие далее уточняется). Смысл изменения в том, что теперь неважно, какого размера собственно кусок исходного разбиения (он может быть даже точкой), но кусок находится внутри круга, одинакового по радиусу с другими кругами, и эти круги не пересекаются. Такой набор кругов с центрами в точках остова называем также футляром (порожденным данным изображением). Уточним условия на радиус кругов и возможное их пересечение. Полагаем, что ни один круг не включает центр другого круга в качестве внутренней точки (т.е. максимальный радиус равен минимальному расстоянию между точками остова). Содержательный смысл этого условия (ограничения) очевиден. (Пример: пусть точка a — центр первого круга, и этот круг включает точку b — центр второго круга.Тогда a можно преобразовать в точку b, и затем в любую точку второго круга, в том числе, и за пределами первого круга, что, по смыслу, недопустимо). Круги могут соприкасаться, и даже пересекаться, но к внутренним точкам круга (усеченного) относим только большую часть круга по соответствующую сторону от отрезка — границы пересечения. Точки наполнения футляра — только внутренние точки (возможно усеченных) кругов.

Как совместить привязку футляров к конкретным изображениям в среде с тем, что мы декларируем рассмотрение с точностью до аффинных преобразований. Мы намереваемся рассматривать изображения при распознавании с точностью до аффинных преобразований, т.е. безотносительно к параллельным переносам, преобразованиям симметрии, вращениям, изменениям в размерах, сжатиям, растяжениям и любым их комбинациям. Это в какой-то мере соответствует тому, что мы рассматриваем именно форму фигур, если говорить о фигурах, а не какие-то не имеющие отношения к форме обстоятельства, связанные с внешними системами координат [1, 2, 4]. Вместе с тем известно, что в целом аффинные преобразования не сохраняют форму фигур: чрезмерными сжатиями или растяжениями можно сделать фигуру неузнаваемой в сравнении с оригиналом. Тем самым, эта чрезмерность должна быть ограничена. Но чем и как, в каких пределах, и чем эти пределы должны определяться? Содержательные рассмотрения выше были явно привязаны к «месту», т.е. к конкретному изображению, которое мы называли базовым. Конкретными были и размеры кусков, и радиусы кругов. Как это совместить с рассмотрением с точностью до аффинных преобразований? Положим в нашей среде S есть группа G изображений «двойки», безусловно вложимых аффинными преобразованиями в футляр A на рис. 2, но разные по размерам, положению на плоскости, с локальными изменениями, и пр., в том числе, для примера, может присутствовать и изображение X, аффинно эквивалентное с изображением из центров кругов на рис. 2, однако сжатое к некоторой прямой с таким коэффициентом, что все его точки выстроились практически в одну прямую. Ясно, что узнать в этом множестве точек X остов из рис. 2 невозможно, хотя они и аффинно эквивалентны. Так вот в наших модельных построениях должно быть учтено, с одной стороны, то, что этот X с остовом на рис. 2 аффинно «совпадают», с другой — то, что уж примером формы «двойки» X явно служить не может. Изображения группы G — это конкретные наборы точек в разных частях среды S. С содержательной точки зрения они в разной степени соответствуют тому, чтобы называться типичной двойкой, эталоном. Мы должны выделить изображение, наиболее достойное того, чтобы считаться эталоном в группе, причем, напомним, безотносительно к размерам, ориентации и пр., к сжатию-растяжению в определенных пределах, притом что пределы эти должны возникать «внутри» модели, а не задаваться извне. Мы будем пробовать на роль эталона группы поочередно все изображения группы. Возьмем, для начала, изображение A на рис. 2. Положим, к виду рис. 2 оно приводится разбиением исходного множества точек на m равновеликих кусков, где m — число кругов на изображении рис. 2, и центры кусков — точки изображения рис. 2. Изображение из центров кусков обозначим через a^+. Далее для произвольного изображения A^0 из группы G осуществляем все возможные его разбиения на m кусков (вообще говоря, уже не заботясь о равновеликости), строим (для каждого разбиения) точечное изображение a_0^+ из центров этих кусков, и укладываем аффинными преобразованиями изображение a_0^+ на изображение a^+. Если укладка (оптимальное взаиморасположение) такова, что при этом и каждый соответствующий кусок оказывается внутри соответствующего круга, то изображение A_0 называем приемлемым для эталона A (при уровне дробности m эталона). Поочередно проверяем на приемлемость все изображения группы, и поочередно при каждом изображении, рассматриваемым в качестве эталона. Если приемлемости нет в каждом из этих рассмотрений, то уровень m дробности называем не адекватным группе. Наибольший уровень адекватности группы называем ее дробностью, он и служит основной характеристикой «однородности» группы, близости по форме составляющих ее изображений. Чем ближе по форме изображения, тем, полагаем, больше дробность. Предельный случай — когда все изображения аффинно эквивалентны — даст максимально возможную в этом случае дробность, равную числу точек в каждом изображении. Минимальная дробность, и она есть всегда, для любой группы, равна единице. Разнородность изображений в группе понижает дробность: ясно, то если, например, в группу «двоек» добавить «четверку», то дробность понизится. Нетрудно видеть, что в этих построениях хоть и используются в целом аффинные преобразования, но они все же ограничены, в известной степени, примерами, т.е. изображениями группы G. И это разумно, ибо возможных аффинно преобразованных изображений континуум (их можно назвать своеобразными «фантазиями»), но их приемлемость мы связываем с конечным «опорным» множеством, т.е. множеством конкретных примеров из данной среды S, с изображениями группы G (это есть представленная нам «реальность», в отличие от фантазий).

5. Искомое (оптимальное) взаиморасположение изображений

Пусть изображение A состоит из точек $a_1, ..., a_n$, изображение B — из точек $b_1, ..., b_n$, ψ — одно из возможных взаимно однозначных соответствий между точками изображений A и B, которым точке a_i из A сопоставляется точка $b_{\psi(i)}$ из B $(i = 1, ..., n)$. Обозначим через $B^\star$ множество всех изображений, получаемых из B аффинными преобразованиями. Полагаем, что на B' из $B^\star$ сохраняется нумерация, порожденная изображением B, т.е. через b'_i на B' обозначается точка, в которую переходит при соответствующем преобразовании точка b_i из B.

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

Дадим определение искомого (или оптимального) взаиморасположения: через $l_A(B')$ обозначим длину наибольшего из отрезков $\left(a_i b'_{\psi(i)}\right)$ $(i = 1, ..., n)$. Рассмотрим B_0 — некоторое изображение из $B^\star$, и ψ_0 — одно из взаимно однозначных соответствий между точками изображений A и B. Пусть существует такое ϵ_1, что для всех B' из $\{B_0\}^{\epsilon_1}$ 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 $\{B'_0\}^{\epsilon_1}$ 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 B.

Что есть в содержательной интерпретации искомое (оптимальное) расположение изображения B на A, например, в случае, если A «эталонное» изображение «двойки», а B — тоже «двойка», но несколько искаженная по форме? Тогда B_0 — расположенная аффинными преобразованиями на A искаженная «двойка», причем расположенная так, чтобы, несмотря на исходные искажения, максимально повторять своей формой форму неискаженной A, при этом — безотносительно к размерам, ориентациям и сжатиям-растяжениям исходной фигуры B. Параметр $l_A(B')$ и служит мерой несовпадения форм фигур. Для двух двоек он, предполагается, будет существенно меньше, чем, скажем, для «двойки» и «четверки». В [1] представлены теоремы, на основе которых можно находить оптимальное расположение конечной процедурой.

6. Футляры для групп изображений

Выше было дано содержательное описание футляра для изображения A (или порождаемого изображением A). Уточним его некоторыми определениями и ссылками на ранее полученные теоремы. Зададимся некоторым m из промежутка от 1 до n (n — число точек в A). Рассмотрим покрытие A' изображения A независимыми кусками $A_1, \ldots, A_m$ (таких покрытий — конечное множество). Центры кусков обозначим через $a_1, \ldots, a_m$, в целом они составляют изображение a^+, его называем остовом (m-остовом). Пусть R — наименьшее расстояние между точками $a_1, \ldots, a_m$, полагаем каждую из этих точек центром круга радиуса R, и каждую из точек куска A_i ближе к a_i, чем к другим центрам. Последнее значит, что если круги пересекаются (по отрезку прямой), то кусок должен быть внутри усеченного круга. Если R меньше половины размера покрытия, то A' называем неправильным. Далее, когда имеются ввиду футляры, рассматриваем правильные покрытия для A. Футляр изображения есть пара $<$изображение A, остов $O>$. Этой пары достаточно, чтобы по ней на изображении A построить систему кругов радиуса R с центрами в точках из O, определить границы-отрезки пересечения кругов (если пересечения есть), куски изображения A в каждом из кругов. Пример (вырожденный): изображение A состоит из точек $a_1, \ldots, a_n$, рассматриваем разбиение на n кусков, т.е. каждый кусок — точка, R равен минимальному расстоянию между точками.

Итак, построен m-футляр изображения A. Если такой футляр не единственный, то каждый рассматривается независимо.

Пусть теперь B' есть одно из возможных разбиений изображения B на m непересекающихся кусков (не обязательно правильное), с центрами $b_1, ..., b_m$ (обозначение в целом: b^+). Если мы теперь аффинными преобразованиями трансформируем изображение B' вместе с точками $b_1, ..., b_m$, то новые положения точек $b_1, ..., b_m$ будут по прежнему центрами преобразованных кусков. Это следует из теоремы 1.

Теорема 1. *Пусть дано изображение X из точек $x_1, ..., x_n$ и точка y — его центр. Пусть изображение X' из точек $x'_1, ..., x'_n$ есть аффинно преобразованное изображение X, и точка y' — его центр. Тогда изображения из точек $x_1, ..., x_n, y$ и точек $x'_1, ..., x'_n, y'$ аффинно эквивалентны.*

Доказательство. Утверждение означает, что при аффинных преобразованиях изображения X из точек $x_1, ..., x_n, y$ в изображение X' из точек $x'_1, ..., x'_n, y'$ центр y изображения X переводится в центр y' изображения X'. Действительно, допустим, что центр преобразованного изображения X' есть точка y'', отличная от y'. Точка y есть центр эллипса E минимального по площади, в который вписано изображение X, т.е. если обозначить площадь эллипса через $P(E)$, площадь выпуклой оболочки изображения X через $P(X)$, то отношение $P(E)/P(X)$ минимальное из возможных. Такой эллипс согласно [5] существует и единственен. Аналогично точка y'' есть центр наименьшего по площади эллипса E'' для изображения X'. При аффинном преобразовании эллипс E преобразуется в эллипс E' с центром в y'. Поскольку y' и y'' предполагаются разными, то и эллипсы E' и E'' тоже разные. Но при аффинном преобразовании отношение $P(E)/P(X)$ сохранится равным $P(E')/P(X')$, а это значит, что эллипс E' тоже минимален по площади для X' и единственен — пришли к противоречию. Итак, E' и E'' должны совпадать, а, значит, совпадают и y'' с y'. Теорема доказана.

□

Определим понятие вместимости для B' в m-футляр изображения A. Расположим изображение b^+ на изображении O искомым образом. Это преобразует и изображение B' в целом. Если при этом каждый кусок изображения B' окажется внутри соответствующего круга (возможно, усеченного) футляра изображения A, то говорим, что данное m-разбиение изображения B вместимо в этот футляр. Изображение B называем m-вместимым в изображение A, если хотя бы одно из его m-разбиений вместимо в хотя бы один из m-футляров изображения A.

Пусть группа G из S^+ состоит из изображений $A_1, ..., A_k$, и пусть u — наименьшее число точек в изображениях группы. Зададимся некоторым

m из промежутка от 1 до u и пусть существует хотя бы одно из изображений группы такое, что все $A_1, ..., A_k$ m-вместимы в некоторый футляр этого изображения. Тогда группу называем m-совместимой. Наибольшее значение m, для которого группа m-совместима, обозначаем через M и называем дробностью группы. Те из изображений группы, M-футляры которых вмещают все изображения группы, называем ее эталонами. Теперь каждая группа из S^+ снабжена характеризующим ее параметром — дробностью, набором эталонов, и, будем полагать, M-футлярами на эталонах.

7. Полные группы (полные образы)

Пусть G_1 и G_2 — группы из S^+ с одинаковой дробностью, и пусть множество изображений G_2 есть подмножество (собственное) группы G_1. Тогда говорим, что группа G_1 полнее группы G_2. Группу называем полной, если нет группы полнее, чем она. Таким образом, здесь понятие полной группы основывается на задании группы перечислением входящих в нее изображений.

Пусть G — группа, M' есть M-футляр одного из эталонов группы. По построению, все изображения группы вместимы в M'. Однако, в M' могут быть вместимы и другие изображения среды. Назовем совокупность всех вместимых в M' изображений максимальным наполнением футляра M'. Нетрудно видеть, что максимальное наполнение футляра M' есть полная группа. Итак, ранее для полной группы было возможно ее задание, как и для всех групп, перечислением всех входящих в группу изображений. Не очень удобный способ, если изображений много. Однако теперь для полной группы возможно ее задание каким-либо футляром эталона группы, и указанием, что группу составляют все вместимые в этот футляр изображения.

Как можно содержательно трактовать описанное выше, т.е. наличие полной группы G и ее подгрупп, неполных, но максимальное наполнение которых совпадает с G? Задание примеров изображений группами (множество S^+) означает, что эти подгруппы могут быть разными: где-то примеров больше, где-то меньше, причем для одного и того же образа. В среде, в разных ее частях, может быть много разных примеров одной и той же фигуры (возможно, с вариациями в контурах). Скажем, некто познакомился с изображением «двойки» на сравнительно небольшом множестве примеров (т.е. это одна из возможных групп в S^+. Он «выработал» футляр для этой группы и это позволит ему при появлении неизвестного изображения определить, вместимо оно в этот футляр или нет, то есть распознать изображение. Другой субъект «выработает» схожий футляр (иными словами, схожее понятие «двойки») на другом множестве ее примеров (другая группа), но в целом, возможно, изображения первой группы вместимы в футляр второй и наоборот, что и означает, что их можно объединить в одну группу. В целом это представляет в модели то содержательно очевидное обстоятельство, что одному и тому же образу можно «обучиться» на разных примерах и из разных частей среды. Множество всех полных групп обозначаем через S^{++} и трактуем как множество образов во втором приближении.

Предшествующее определение футляра основывалось на делении на куски изображений группы и рассмотрении центров этих кусков. Это определение использует ясные алгоритмы построения футляра, его и эти алгоритмы можно назвать первичными. Далее – несколько более общие определения ранее введенных понятий.

Пусть A — изображение состоящее из точек $a_1, \ldots, a_u$, изображение O (далее называем его остовом) состоит из точек $o_1, \ldots, o_n$, причем n не больше u, и дано некоторое число r не большее наименьшего расстояния между точками остова. Пусть для a_i (i=1, … ,u) ближайшая к ней точка из O единственна, обозначим ее через o_j, и расстояние между точками o_j и a_i меньше r. Точку a_i называем прилежащей к точке o_j. Совокупность точек из A, прилежащих к точке o_j, называем куском изображения A, прилежащим к точке o_j, обозначаем через A_j (j=1, … ,n). Если теперь каждая из точек из A попадает в некоторый кусок, и все куски изображения A непустые, то A и O называем приемлемо взаиморасположенными. Совокупность кусков A_j ($j = 1, \ldots, n$) называем разложением изображения A по остову O. Пару <A,O> называем футляром (n-футляром), порожденным изображениями A и O. Пусть существует изображение B', являющееся аффинно преобразованным изображением B, и приемлемо взаиморасположеное с O. Тогда B называем вложимым в футляр <A,O>, а изображение B' – его вложением. Куски B'_1, … ,B'_n называем разложением вложения B' на куски по футляру.

Нетрудно видеть, что ранее введенное понятие футляра, остова и вложимости изображения в футляр есть частный случай последних определений. В отличие от первичного футляра <A,O>, для которого остов O строится по изображению A, в обобщении остов O в паре <A,O> рассматривается как данный, данным является и взаиморасположение изображений A и O.

Пусть задана полная группа G, состоящая из изображений $g_1, g_2, \ldots, g_k$, задан футляр $< g_1, O >$, в который вмещены все соответственно преобразованные изображения группы G, которые обозначим через $g_1, g'_2, \ldots, g'_k$. Пусть имеется изображение X, про которое известно, что оно вмещено в футляр $< g_1, O >$. Тогда X совпадает с одним из изображений $g_1, g'_2, \ldots, g'_k$. Действительно, группа G полная, значит никаких изображений, отличных от $g_1, g_2, \ldots, g_k$, в футляр $< g_1, O >$

вместить нельзя. Не может изображение X быть и копией какого-либо из изображений $g_1, g'_2, \ldots, g'_k$, несколько по иному расположенной в футляре, поскольку каждое из исходных изображений $g_1, g_2, \ldots, g_k$ преобразуется, но не копируется.

Пусть G и G' — полные группы, причем G' — подгруппа (собственное подмножество) группы G. Говорим, что G' максимальная подгруппа, если не существует среди множества всех полных подгрупп группы G такой группы G'', что G' есть ее подгруппа. Такую максимальную подгруппу G' называем непосредственным подобразом (по вложению) образа G. Набор всех непосредственных подобразов образа G называем спектром (по вложению) этого образа.

8. Продолжаемые и замкнутые группы изображений

Полный образ может содержать своеобразные «внутренние», «продолжаемые» образы. Дадим соответствующие определения. Пусть O — остов футляра, состоящий из точек $o_1, \ldots, o_n$. Пусть задана полная группа G изображений, и $B_1, \ldots, B_t$ есть вложения изображений группы G в футляр. Тем самым $B_1, \ldots, B_t$ есть аффинно преобразованные изображения из группы G и каждое из них приемлемо взаиморасположено с O. Отметим, что каждое изображение из G может быть потенциально не единственным образом приемлемо взаиморасположено с O, но реализуется только одно такое взаиморасположение, т.е. каждое изображение из G представлено среди $B_1, \ldots, B_t$ только одним изображением. Через b_{ij} $(i = 1, \ldots, t; j = 1, \ldots, n)$ обозначим разложение на куски вложения B_i по футляру.

Пусть $o_{j1}, \ldots, o_{jh}$ — непустое подмножество точек остова O, и ${b_{j1}}^i, \ldots, {b_{jh}}^i$ есть куски разложения изображения B_i, прилежащие к точкам соответственно $o_{j1}, \ldots, o_{jh}$. Обозначим изображение из точек кусков ${b_{j1}}^i, \ldots, {b_{jh}}^i$ через g_i $(i = 1, \ldots, t)$.

Еще раз отметим, что каждое из вложенных в футляр изображений $B_1, \ldots, B_t$ есть аффинно преобразованное изображение из исходной полной группы G, соответственно каждое из g_i $(i = 1, \ldots, t)$ есть аффинно преобразованная часть некоторого изображения из G. Обозначим совокупность всех этих частей через Q и говорим, что группа Q есть часть образа G, порожденная частью $o_{j1}, \ldots, o_{jh}$ остова O. Если группа Q полная, то образ Q называем существенной частью образа G, порождаемой частью $o_{j1}, \ldots, o_{jh}$ остова O.

Тривиальный пример части (и существенной части) образа можно представить на основе рис. 2. Если считать, что наполнение этого футляра есть все фигуры «двойки» из рассматриваемой среды, то удаление одной точки из футляра приведет к тому, что оставшиеся точки футляра породят образ, являющийся частью образа «двойки» (причем почти неотличимой от «двойки»).

Трактовка групп G и Q состоит в том, что если в среде обнаружено изображение X, принадлежащее образу G, то с необходимостью имеется в среде изображение x, являющееся частью X, и принадлежащее образу Q. И наоборот, если обнаружено x из Q, то существует и X из G, частью которого это x является. Это можно трактовать и как прогнозирование X по x, и, наоборот — x по X. Существенная часть Q образа G может быть одновременно и существенной частью другого образа P. В этом случае образ Q называем перекрестком образов G и P. Перекресток может быть общей частью не только двух, но и большего числа образов Образ Z, не являющийся существенной частью никакого другого образа, назовем замкнутым.

Пусть O_1 и O_2 есть непересекающиеся части остова O замкнутого образа G из соответственно точек $o_{i1}, \ldots, o_{im}$ и $o_{j1}, \ldots, o_{jk}$. Пусть O_3 есть объединение множеств O_1 и O_2, и при этом O_3 есть собственное подмножество множества точек остова O. Пусть V_1, V_2 и V_3 есть группы изображений, порождаемых соответственно частями O_1, O_2, и O_3 остова O и группы G. Задано, что V_1 есть существенная часть образа G. Вопрос: является ли V_3 существенной частью образа G? Вопрос можно трактовать так: O_1 порождает существенную часть V_1 образа G, к O_1 добавляем точки множества O_2, возникает часть O_3 остова O. Будет ли порождаемая им группа V_3 изображений тоже существенной частью образа G? Это можно трактовать и как вопрос о расширении, продолжении существенной части V_1 образа G до существенной части V_3 того же образа.

Теорема 2. *Группа V_3 изображений является существенной частью образа G.*

Доказательство. Надо показать, что группа V_3 изображений является полной. Предположим, что это не так, то есть помимо изображений из V_3 есть еще изображение X, приемлемо расположенное в отношении остова O_3, и не входящее в V_3. Оно будет состоять из двух непересекающихся частей : первая есть X_1, вмещенное в O_1, и вторая есть X_2, вмещенное в O_2. Поскольку — V_1 полная группа, то X_1 должно совпадать с одним из изображений этой группы. Рассмотрим изображение Z из G, частью которого является X_1. Оно содержит часть — обозначим ее через Y — вмещенную в O_3, и состоящую из двух частей — обозначим их через Y_1 и Y_2, вмещенные соответственно в O_1 и в O_2. При этом Y не совпадает с X, но Y_1 совпадает с X_1. Рассмотрим теперь изображение Z' с частью X вместо части Y. Изображения Z и Z' имеют общую часть (это часть

X_1, совпадающая с Y_1). Изображение Z' по построению принадлежит исходной группе G, чего не может быть, ибо группа G — полная. Это первое противоречие, к которому мы приходим. А второе состоит в том, что у двух изображений в получившейся группе G — у Z и Z' — имеется непустое пересечение, что противоречит отдельности рассматриваемых образов. $\square$

Теорему 2 можно рассматривать как обоснование для названия существенной части V_1 образа G образом, продолжаемым в замкнутый образ G.

Из определения замкнутой группы следует, что каждая полная группа либо замкнута, либо нет. В последнем случае группа является продолжаемой (причем в общем случае, не в единственную) замкнутую группу. Совокупность всех групп, продолжаемых в данную замкнутую группу, называем ее шлейфом. Таким образом, все множество полных групп делится на два подмножества: замкнутых групп и групп из шлейфов замкнутых групп.

Пусть существенная часть Q замкнутого образа G порождена частью $o_{i1}, \ldots, o_{im}$ остова, и любое (собственное) подмножество множества $o_{i1}, \ldots, o_{im}$ уже не порождает существенную часть. Тогда часть Q называем минимальной или признаком образа G. Тем самым, применительно к образу G можно говорить о множестве признаков для него. Частный случай признака, когда m равно единице, называем меткой образа G. При поиске в среде какого-либо изображения X образа G задача может быть, очевидно, заменена на поиск признака, или даже метки образа G.

Множество всех замкнутых групп считаем очередным приближением к понятию визуального образа.

Список литературы

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

[2] Козлов В. Н., “Conclusiveness and Heuristics in Visual Recognition”, *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications*, **24**:4 (2014), 1 - 7.

[3] Крушинский Л. В., Козлов В. Н., Кудрвцев В. Б., *О некоторых результатах применения математики к моделированию в биологии*, в сборнике «Математические вопросы кибернетики», 1988.

[4] Кудрявцев В. Б., Гасанов Э. Э., Подколзин А. С., *Введение в теорию интеллектуальных систем*, М.: Издательство «МАКС Пресс», 2006.

[5] Загускин В. Л., “Об описанных и вписанных эллипсоидах экстремального объема”, *Успехи математических наук*, **13**:6 (1958), 89–93.

On three initial approximations to the formal definition of a visual image in an arbitrary visual environment

Kozlov V.N.

Visual image recognition is one of the central task for intelligent systems. The advancement of mathematical and theorem-based research methods in this field is significantly hampered by the lack of a complete and acceptable formal definition of the concept of a visual image in any visual environment. This paper presents an approach to such a definition using successive approximations and describes three approximations.

Keywords: visual image, image recognition, affine transformations.

References

[1] Kozlov V. N., *Introduction to the Mathematical Theory of Visual Perception*, Publishing House of the Center for Applied Research at the Faculty of Mechanics and Mathematics of Moscow State University, Moscow, 2007 (in Russian).

[2] Kozlov V. N., “Conclusiveness and Heuristics in Visual Recognition”, *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications*, **24**:4 (2014), 1–7 (in Russian).

[3] Krushinsky L. V., Kozlov V. N., Kudryavtsev V. B, *On Some Results of Applying Mathematics to Modeling in Biology*, Mathematical Issues of Cybernetics, 1988.

[4] Kudryavtsev V. B., Gasanov E. E., Podkolzin A. S., *Introduction to the Theory of Intelligent Systems*, Publishing House “MAKS” Press, Moscow, 2006.

[5] Zaguskin V. L., “On circumscribed and inscribed ellipsoids of extremal volume”, *Uspekhi Matematicheskikh Nauk*, **13**:6 (1958), 89–93 (in Russian).