Алгоритмы формирования системы взаимосвязанных образов

В. Н. Козлов

В работе показывается, как плоская среда «порождает» образы, и как в этом случае представляется распознавание образов. Доказываются некоторые утверждения о свойствах кодов изображений.

Ключевые слова: распознавание образов, зрительные образы, математическое моделирование в биологии, зрение роботов.

В этой работе показывается, как плоская среда «порождает» образы, и как в этом случае представляется распознавание образов. Доказываются некоторые утверждения о свойствах кодов изображений. Истоки работы — в моделировании опытов Л. В. Крушинского по изучению так называемой элементарной рассудочной деятельности у животных [1].

Традиционно распознавание понимается, в общих чертах, так: есть некоторый набор классов $X_1, \ldots, X_m$ объектов и есть неизвестный объект x. Его нужно отнести (приписать) к одному из этих классов. Классы интерпретируются как образы, отнесение x к одному из классов — распознавание образов [2, 3].

Вопрос о том, как возникли классы $X_1, \ldots, X_m$, рассматривается обычно как отдельный и второго плана. Как правило, полагают, что уже есть некое «неструктурированное» множество всех возможных объектов и ставится задача лишь разбить его на классы $X_1, \ldots, X_m$. То, как возникли объекты, составляющие неструктурированное множество — обычно вопрос третьего плана.

Здесь мы несколько отойдем от традиционной постановки и будем полагать, что объекты встроены в некоторую объемлющую среду, явным образом не выделены, могут «непрерывно» переходить друг в друга и иметь общие части. Кроме того, как объекты могут рассматриваться и совокупности объектов, и среда в целом.

Объектами будут изображения, классами объектов — классы изображений, и, тем самым, распознавание здесь — это распознавание визуальных (зрительных) образов.

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

Любое (непустое) подмножество точек изображения — часть изображения, часть называем собственной, если она не содержит всех точек изображения.

Среда S — некоторое изображение, любая его часть — изображение из среды.

Память P — аффинная копия среды S, то есть изображение, аффинно эквивалентное среде S. Любая часть памяти — изображение из памяти.

Поле зрения x — произвольное изображение из среды.

Все вместе — среда S, память P и поле зрения x, вместе с некоторой совокупностью правил по оперированию с ними, и составляют модель, которая здесь рассматривается.

Одна из возможных наглядных интерпретаций того, чем может быть среда, следующая: это совокупное изображение, представляющее собой разбросанные по плоскости символы цифр и букв, другие фигуры, разные по размерам, ориентациям, по разному сжатые и растянутые. Эти фигуры могут образовывать конгломераты, в свою очередь, тоже рассматриваемые как подлежащие распознаванию изображения. Все эти объекты на плоскости могут пересекаться, «непрерывно» переходить друг в друга, и т. п.

Другая возможная интерпретация для S — план, карта некоторой местности, например, аллей в парке, с возможно повторяющимися частями, с условными значками, изображающими объекты на местности, и пр. В этом случае может стоять задача так сориентироваться на местности, чтобы найти переход к каким-то объектам на ней.

Память P — аффинная копия среды, причем то, что это не часть, а именно полная копия среды — трактуется как максимально полное знание о среде, полная ее модель.

Поле зрения x — это то, что мы видим в среде. Это произвольное (непустое) подмножество точек среды. Его можно трактовать как некоторый участок среды, доступный восприятию. Но это не вполне «окошечко» в среду. Ибо в окошечке мы, конечно, ограничены в восприятии среды размерами окошка, но уж то, что попало в окошко, обычно видим полностью. Здесь же часть точек, попавших в окошко, может быть недоступна восприятию («загорожена», и т. п.). Поэтому поле зрения — именно подмножество точек среды.

Перенумеруем точки среды S попарно различными числами. Теми же числами пометим и соответствующие точки памяти P. Под соответствующими точками понимаются те, которые совпадают при аффинном совмещении P и S. Если такое совмещение не единственно (например, вследствие симметрии), то фиксируем какое-либо из них.

Положим теперь, что точки поля зрения x заданы вместе с их номерами в S, а остальные точки из S не заданы (условно говоря, не видны), то есть их расположение относительно точек, вошедших в x, априори неизвестно. Тогда мы можем восстановить знание о среде, если наложим аффинными преобразованиями память P на поле зрения x так, чтобы совместились одинаково помеченные точки. Если в x не менее трех точек, не лежащих на одной прямой (такие изображения будем называть невырожденными), то такое наложение единственно. В целом это означает, что используя знание о среде, заложенное в памяти, и видя только часть среды в виде x, мы восстанавливаем полное знание о среде, то есть знаем все «скрытые» точки среды в их расположении относительно «видимых» точек из x. При этом восстановление знания о среде осуществляется как «вширь», то есть мы узнаем о точках среды за пределами участка, ограниченного полем зрения — это можно трактовать как экстраполяцию, прогнозирование, так и «вглубь», поскольку мы узнаем о точках в пределах участка, занимаемого полем зрения, но отсутствующих в x потому, например, что они «не видны». Это последнее можно трактовать как распознавание, поскольку мы восстанавливаем, например, фигуру по ее сильно «прореженному» эскизу. Впрочем, в рамках модели принципиальной разницы между так понимаемыми прогнозированием и распознаванием, как это видно, нет.

Нечто подобное описываемой схеме действий присутствовало в работах по моделированию опытов по изучению элементарной рассудочной деятельности и экстраполяции у животных [1]. В той модели налагаемое на поле зрения изображение из памяти именовалось «руководством к действию», поскольку содержало, как предполагалось, необходимые для организации целенаправленного поведения сведения о среде.

Совокупность точек памяти, соответствующих точкам из x (то есть с теми же номерами), обозначим через p. Если в x не меньше трех точек, лежащих не на одной прямой, то возникает только один вариант прогноза, и обеспечивается это однозначностью совмещения точек из x с точками из p. Положим теперь, что номера точек в x и p несущественны, то есть x «обезличено», и прогноз обеспечивается аффинным совпадением части p памяти P и поля зрения x. Тогда в памяти может найтись набор разных ее частей $p_1, \dots, p_d$, аффинно эквивалентных с x. При этом часть p_i будем считать разной с p_j ($i, j = 1, \dots, d$, $i \neq j$), если они разнятся хотя бы одной (по номеру) точкой. Кроме того, каждая из частей p_i в некоторых случаях (например, вследствие симметрии) может по разному совмещаться с x. Все это дает разные варианты $P_1, \dots, P_k$ ($k \geqslant d$) наложения памяти P на поле зрения x (через P_i, $i = 1, \dots, k$, обозначена память P после соответствующего преобразования). Каждый из вариантов наложения P_i имеет часть p_i, точки которой совмещены с точками поля зрения x. Все варианты $P_1, \dots, P_k$ наложения памяти P на x попарно разные. При этом варианты P_i и P_j ($i, j = 1, \dots, k$, $i \neq j$) называем разными, если в их частях p_i и p_j есть хотя бы по одной разной точке (разной по номеру в памяти P), совмещаемой с одной и той же точкой в x.

Итак, есть среда S с наложенными на нее изображениями $P_1, \dots, P_k$, представляющими собой преобразованную память P. Каждая из $P_1, \dots, P_k$ своими частями $p_1, \dots, p_k$ совмещена с частью x среды. Точки из x можно назвать k-кратными, имея ввиду то, что с каждой из них совмещена точка из P_i ($i = 1, \dots, k$) и таких точек, тем самым, k. Однако точки из x, возможно, не единственные k-кратные точки на S. Обозначим через X всю совокупность точек из S, являющихся k-кратными. Так же, как точкам из x соответствуют на P_i точки ее части p_i, так и точкам из X соответствуют на P_i точки ее части, которую обозначим через X_i $(i = 1, \ldots, k)$.

Итак, теперь есть среда S с наложенными на нее изображениями $P_1, \ldots, P_k$. И теперь каждое из $P_1, \ldots, P_k$ своими частями $X_1, \ldots, X_k$ совмещено с частью X среды. Поле зрения x есть часть (не обязательно собственная) изображения X, все точки в X k-кратные и других k-кратных точек в S нет. Изображение X можно назвать замыканием поля зрения x, и в случае, когда x совпадает с X, поле зрения можно назвать замкнутым относительно среды S. Изображения совокупности $X_1, \ldots, X_k$ называем представителями изображения X в памяти P.

Интерпретация описаному следующая. При данном поле зрения x этому полю зрения могут соответствовать в памяти разные участки и соответственно получаться разные прогнозы. Однако прогноз нужен один, но достоверный. Как мы выяснили, при всех вариантах возможного прогноза у них есть общая часть, представленная, можно полагать, k-кратными точками среды, то есть изображением X. Это и есть достоверный прогноз. Точки из X гарантированно будут присутствовать в среде, безотносительно к тому, с какой частью памяти P может аффинно совпасть поле зрения x.

Теперь то, что мы получили как изображения $X_1, \ldots, X_k$ памяти, «спроецируем» на среду. Имеется ввиду то, что изображение X_i с частью p_i $(i = 1, \ldots, k)$ состоит из точек памяти со вполне определенными номерами. Обозначим в среде через S_i с частью s_i $(i = 1, \ldots, k)$ изображение из точек с теми же номерами. Совокупность $(S_1, \ldots, S_k)$ назовем псевдообразом, изображение X (замкнутое, по построению) — представителем псевдообраза. Нетрудно видеть, что, по построению, изображение X присутствует среди изображений $(S_1, \ldots, S_k)$ псевдообраза.

По причине симметрии в псевдообразе $(S_1, \ldots, S_k)$ могут быть повторяющиеся изображения, то есть состоящие из точек с одними и теми же номерами. Устраним, однако, такие повторения и так сокращенную и заново перенумерованную совокупность $(S_1, \ldots, S_t)$, где $t \leqslant k$, назовем образом. В ней все изображения попарно разные, все аффинно эквивалентны изображению X, и других частей среды, аффинно эквивалентных изображению X, нет.

Исходное поле зрения x однозначно определило свое замыкание X и образ $(S_1, \ldots, S_t)$. В каждом изображении S_i $(i = 1, \ldots, t)$ есть часть s_i, соответствующая полю зрения x. Очевидно, что если рассматривать s_i в качестве поля зрения, то в рамках описанной процедуры его замыканием будет изображение S_i и возникнет тот же образ $(S_1, \ldots, S_t)$. На этом основании каждое из изображений в образе может рассматриваться как его представитель.

Вместе с тем изображение x в общем случае не единственное, замыканием которого будет изображение X. Из этого следует, что в рамках описанной процедуры данный образ $(S_1, \ldots, S_t)$ может возникать при разных аффинно не эквивалентных друг другу полях зрения.

Итак, каждое невырожденное изображение x из среды, рассматриваемое в качестве поля зрения, либо замкнуто относительно среды, либо однозначно порождает свое замыкание X. Изображение X является представителем однозначно определяемого образа $(S_1, \ldots, S_t)$. Образ состоит из всех разных изображений среды, аффинно эквивалентных с X.

Как уже было сказано, X может быть, в общем случае, замыканием разных полей зрения. Назовем все изображения (незамкнутые), замыканием которых являются представители образа $(S_1, \ldots, S_t)$, бассейном притяжения этого образа (используя некоторую аналогию с термином, введенным Хопфилдом в нейрокомпьютерной модели [4]). Тогда множество всех невырожденных изображений среды распадается на непересекающиеся подмножества $Q_1, \ldots, Q_d$. Каждое Q_i $(i = 1, \ldots, d)$ состоит из представителей одного образа, и из бассейна притяжения этого образа. Каждое изображение из Q_i, рассматриваемое в качестве поля зрения в рамках описанной процедуры, порождает этот образ.

Примером образа является среда S. Если рассматривать ее в качестве поля зрения, то она порождает совокупность представителей образа, состоящую из единственного изображения, то есть изображения S.

Другой пример — образ, порожденный любыми тремя точками, не лежащими на одной прямой. Замыканием такого поля зрения будет тот же трехточечник. В этом случае в образе $(S_1, \ldots, S_t)$ — все симплексы с точками из S (симплекс-образ).

Не столь тривиальные примеры образов возникают, очевидно, если в среде S есть повторяющиеся, то есть аффинно эквивалентные части.

Все изложенное можно считать исходным (и, в значительной мере, содержательным) описанием модели. Дополним его некоторыми формальными определениями. Это позволит доказать некоторые теоремы и в целом привнесет доказательность в модель.

Пусть A — произвольное (невырожденное) изображение. Перенумеруем точки в A попарно различными числами и обозначим множество номеров через M_A. Рассмотрим биекцию $\psi : M_A \to M_A$ такую, что существует аффинное преобразование, переводящее A в A так, что совмещаются точки i и $\psi(i)$ (для всех i из M_A). Такую биекцию назовем правильной. Число правильных биекций назовем индексом симметрии k_A изображения A. Отметим, что одна правильная биекция, тривиальная, при которой $i = \psi(i)$, всегда существует.

Пусть B есть невырожденная часть изображения A. Пусть для любой части B' изображения A, аффинно эквивалентной с B, выполнено следующее: если $\varphi : M_B \to M_{B'}$ есть такая биекция, что B и B' совместимы аффинным преобразованием так, что любая точка b из B совмещается с точкой $\varphi(b)$ в B', то биекция φ есть сужение некоторой правильной биекции $\psi : M_A \to M_A$. Тогда изображение B называем симметрично сопряженным с A с индексом симметрии k_{AB}, равным числу биекций φ.

Отметим, что в последнем определении часть B изображения A не обязательно собственная, и, кроме того, B может совпадать с B'. В целом это определение можно рассматривать как расширение предшествующего.

Лемма 1. *Пусть изображение B симметрично сопряжено с изображением A. Тогда в A содержится k_{AB}/k_B разных частей, аффинно эквивалентных изображению B.*

Доказательство. Рассмотрим, из чего складывается величина k_{AB}. Во-первых, из числа биекций, отображающих B в B, таких биекций k_B. Далее, если B' — часть A, отличная от B, то будет еще k_B биекций, отображающих B в B'. Затем рассмотрим B'', отличное от B и B' и аффинно эквивалентное изображению B, и, соответственно, еще k_B биекций, отображающих B в B''. И так далее. Если k — число разных частей в A, аффинно эквивалентных изображению B, то получаем, что $k_{AB} = k\dot{k}_B$. Лемма доказана.

Лемма 2. *Пусть изображение B симметрично сопряжено с изображением A. Тогда $k_{AB} = k_A$.*

Доказательство. Каждой биекции φ (из определения симметричной сопряженности B с A) соответствует некоторая правильная биекция ψ, сужением которой эта φ является. Поэтому двум разным φ не может соответствовать одна и та же ψ.

Каждой правильной биекции $\psi : M_A \to M_A$ соответствует некоторое преобразование A в A, значит и преобразование B в B', и тем самым, некоторая биекция $\varphi : M_B \to M_{B'}$. При этом двум разным биекциям ψ_1 и ψ_2 не может соответствовать одна и та же биекция φ (то есть эти две ψ не могут иметь сужением одну и ту же φ). Действительно, пусть это не так. Тогда это значит, что по ψ_1 мы можем аффинными преобразованиями перевести A в A_1, и по ψ_2 — A в A_2. В целом у A_1 и A_2 номера совпавших точек не совпадают, но в части B' — номера у всех совпавших точек одинаковые. Поскольку в B' не менее трех точек, лежащих на одной прямой, то совпадение по ним при аффинном преобразовании должно было бы привести к совпадению по номерам и у всех остальных точек.

Итак, показано, что каждой φ соответствует одна ψ, и наоборот — каждой ψ соответствует одна φ. Отсюда $k_{AB} = k_A$. Лемма доказана.

Лемма 3. *Пусть A' и A'' — два разных аффинно эквивалентных изображения, и B — их общая (невырожденная) часть. Тогда B не может быть симметрично сопряженным с A' или с A''.*

Доказательство. Действительно, если у разных A' и A'' есть общая часть B, то это значит, что индекс симметрии у B по меньшей мере равен двум. Но биекция $\varphi : M_B \to M_B$ и соответствующее аффинное преобразование дает в одном случае A', в другом — A'', причем A' отлично, по условию, от A''. Это значит, что B не сопряжено симметрично с A' (или с A''). Лемма доказана.

Лемма 4. (Критерий симметричной сопряженности изображения B с изображением A) *Пусть B является частью изображения A и k — число разных частей в A, аффинно эквивалентных* *с B. Тогда B симметрично сопряжено с A в том и только в том случае, когда $k_A = k \cdot k_B$.*

Доказательство. В одну сторону утверждение следует из лемм 1 и 2.

Положим теперь, что $k_A = k \cdot k_B$. Нужно показать, что любая биекция $\varphi : M_B \to M_{B'}$, по которой B совмещается с B', есть сужение правильной биекции $\psi : M_A \to M_A$.

Число разных биекций φ равно $k \cdot k_B$, число разных биекций ψ равно k_A. Дано $k_A = k \cdot k_B$. Если какой либо биекции φ не соответствует биекция ψ (то есть преобразование по $\varphi : M_B \to M_{B'}$ не переводит A в A), то это значит, что какие-либо две правильные биекции ψ должны при сужении давать одну биекцию φ, чего, как показано при доказательстве леммы 2, не может быть. Лемма доказана.

Пусть A — произвольное изображение. Множество всех его частей, являющихся невырожденными изображеними, обозначим через A^*.

Рассмотрим на S^* бинарное отношение аффинной эквивалентности, порождающее разбиение H на классы. Из рассмотрений первой части следует, нетрудно видеть, что каждый класс из H есть либо образ, либо входит в бассейн притяжения ровно одного образа. Как выяснить, что именно имеет место?

Назовем класс $\widetilde{B} = (B_1, \ldots, B_k)$ из H симметрично сопряженным с классом $\widetilde{A} = (A_1, \ldots, A_m)$ из H, если каждое изображение из $\widetilde{B}$ является частью некоторого изображения из $\widetilde{A}$, и, при условии, что B_i есть часть A_u, B_j есть часть A_q, то совмещение аффинными преобразованиями B_i с B_j совмещает A_u с A_q.

Ясно, что последнее определение есть расширение определения симметричного сопряжения изображения B с A.

Естественно, каждый класс тривиальным образом симметрично сопряжен с собой же. Однако если не иметь ввиду этот тривиальный случай, то, нетрудно видеть, то, что мы ранее называли образом, есть класс $\widetilde{X}$ из H такой, что не существует другого класса в H, с которым $\widetilde{X}$ был бы симметрично сопряжен. С другой стороны, все изображения из всех классов, симметрично сопряженных с образом $\widetilde{X}$, будут составлять бассейн притяжения этого образа. Отсюда следует важность формулируемого далее критерия симметричной сопряженности класса $\widetilde{B}$ с классом $\widetilde{A}$.

Назовем класс $\widetilde{B}$ потенциально симметрично сопряженным с классом $\widetilde{A}$, если всякая часть B изображения A из $\widetilde{A}$, аффинно эквивалентная с изображениями из $\widetilde{B}$, симметрично сопряжена с изображением A.

Отметим, что определение класса $\widetilde{B}$, симметрично сопряженного с классом $\widetilde{A}$, существенно отличается от определения класса $\widetilde{B}$, потенциально симметрично сопряженного с классом $\widetilde{A}$, тем, что во втором случае в классе $\widetilde{B}$ могут быть изображения, не являющиеся частями изображений из $\widetilde{A}$, и на них никаких условий не наложено. Определить, является ли класс $\widetilde{B}$ потенциально симметрично сопряженным с классом $\widetilde{A}$, нетрудно на основании критерия симметричной сопряженности изображения B с изображением A.

Назовем весом произвольного класса из H произведение числа изображений в классе на индекс симметрии изображений класса.

Теорема 1. *Пусть класс $\widetilde{B}$ потенциально симметрично сопряжен с классом $\widetilde{A}$. Тогда $\widetilde{B}$ симметрично сопряжен с классом $\widetilde{A}$ в том и только в том случае, когда веса классов $\widetilde{B}$ и $\widetilde{A}$ равны.*

Доказательство. Пусть класс $\widetilde{A}$ состоит из n изображений, индекс симметрии изображений класса равен k_A, вес класса $v_A = n \cdot k_A$. Класс $\widetilde{B}$ состоит из m изображений, индекс симметрии изображений класса равен k_B, вес класса $v_B = m \cdot k_B$ и, очевидно, $m \geqslant n$.

Положим, что класс $\widetilde{B}$ симметрично сопряжен с классом $\widetilde{A}$, и покажем, что $v_A = v_B$. Действительно, из лемм 1 и 2 следует, что число разных частей, аффинно эквивалентных изображениям из $\widetilde{B}$, на каждом из изображений из $\widetilde{A}$, равно k_A/k_B. Из леммы 3 следует, что у любой пары изображений из $\widetilde{A}$ нет общих частей, аффинно эквивалентных изображениям из $\widetilde{B}$. Следовательно, общее число разных частей, аффинно эквивалентных изображениям из $\widetilde{B}$, во всех изображениях из $\widetilde{A}$, равно $(k_A/k_B)n$. Все эти части входят в класс $\widetilde{B}$ и других изображений в $\widetilde{B}$ нет. Следовательно, $(k_A/k_B)n = m$, или $v_A = v_B$.

Пусть теперь класс $\widetilde{B}$ потенциально симметрично сопряжен с классом $\widetilde{A}$ и $v_A = v_B$. Покажем, что $\widetilde{B}$ симметрично сопряжен с $\widetilde{A}$. Действительно, на всех изображениях из $\widetilde{A}$ содержится $(k_A/k_B)n$ разных частей, аффинно эквивалентных изображениям из $\widetilde{B}$. Поскольку $v_A = v_B$, то $(k_A/k_B)n = m$, то есть других изображений в $\widetilde{B}$, кроме частей из $\widetilde{A}$, нет.

Если B_i и B_j есть части изображения A из $\widetilde{A}$, то совмещение аффинными преобразованиями B_i с B_j переводит A в A, поскольку B_i (и B_j) симметричны относительно A. Если теперь B_i есть часть A_u, B_j есть часть A_q, и при этом A_u и A_q из одного класса $\widetilde{A}$, то есть аффинно эквивалентны, то, ясно, совмещение аффинными преобразованиями B_i с B_j совместит и A_u с A_q. Теорема доказана.

Оставим далее в рассмотрении только те классы из H, которые являются образами.

Пусть $\widetilde{A}$ и $\widetilde{B}$ — образы. Говорим, что $\widetilde{B}$ есть подобраз образа $\widetilde{A}$, если в каждом изображении из $\widetilde{A}$ есть части, принадлежащие $\widetilde{B}$. Если $\widetilde{B}$ не совпадает с $\widetilde{A}$, то $\widetilde{B}$ — собственный подобраз для $\widetilde{A}$. Далее, если нет оговорок имеются ввиду собственные подобразы.

Если $\widetilde{B}$ — подобраз для $\widetilde{A}$, и $\widetilde{C}$ — подобраз для $\widetilde{B}$, то, очевидно, $\widetilde{C}$ есть подобраз для $\widetilde{A}$.

Если $\widetilde{C}$ — подобраз для $\widetilde{A}$, и нет такого образа $\widetilde{B}$, что $\widetilde{B}$ есть подобраз для $\widetilde{A}$, и $\widetilde{C}$ есть подобраз для $\widetilde{B}$, то $\widetilde{C}$ называем непосредственным подобразом для $\widetilde{A}$.

Пусть $\widetilde{A}$ и $\widetilde{B}$ не есть подобразы один другого, пусть a и b есть части некоторых представителей образов $\widetilde{A}$ и $\widetilde{B}$ и пусть a и b аффинно эквивалентны. Замыканиями a и b не могут быть образы $\widetilde{A}$ и $\widetilde{B}$, поскольку они разные. Значит, существует образ $\widetilde{C}$ такой, что изображения a и b либо его представители, либо входят в его бассейн притяжения. Такой образ $\widetilde{C}$ — общий для $\widetilde{A}$ и $\widetilde{B}$ подобраз. Очевидно, общий подобраз у $\widetilde{A}$ и $\widetilde{B}$ есть всегда, например, симплекс-образ.

Все сказанное позволяет представить множество образов, порождаемых средой S, в виде некоторой графоподобной ярусной конструкции W. Проведем $(n-2)$ горизонтальные прямые и каждую, снизу вверх, пометим числами от 3 до n. На прямой (ярусе) с номером k $(k = 3, \ldots, n)$ отметим столько точек (вершин), сколько имеется образов с представителями из k точек. Если вершине y соответствует образ, являющийся непосредственным подобразом образа, сопоставленного вершине x, то проводим ребро от y к x. Все ребра, тем самым, будут от вершин на ярусах с меньшими номерами, к вершинам на ярусах с большими номерами. На нижней и верхней прямой будет, очевидно, по одной вершине, соответствующих симплекс-образу (обозначим ее через Z), и среде S. В получившемся графе W есть «источник» — симплекс-образ Z, содержащийся во всех остальных образах, и «универсальный» образ S, содержащий все остальные образы.

Под распознаванием произвольного (невырожденного) изображения x из среды понимаем определение его замыкания X, и, соответственно, образа $\widetilde{X}$, представителем которого X является. Эта процедура, с использованием ярусного графа W, может быть организована следующим образом.

Выбираем на x произвольную невырожденную тройку t точек. Ей, соответствует вершина, обозначим ее через Z, симплекс-образа на самом нижнем ярусе. Рассматриваем все ребра, исходящие из вершины Z, и вершины, обозначим их через $\widetilde{V}_1, \ldots, \widetilde{V}_h$, на их концах. Каждый из образов $\widetilde{V}_1, \ldots, \widetilde{V}_h$ содержит Z в качестве непосредственного подобраза. Это значит, что образы $\widetilde{V}_1, \ldots, \widetilde{V}_h$ могут рассматриваться как своеобразные «гипотезы» при дальнейшем расширении исходного трехточечника t. Нужно выяснить, какая из гипотез реализуется при данном x и данном t. Возьмем, для определенности, образ $\widetilde{V}_1$ и какого-либо представителя этого образа, обозначим его через v. Наложением последовательно совмещаем трехточечники из v с трехточечником t, причем с учетом индекса симметрии трехточечников. Каждое наложение дает, в общем случае, свой вариант расположения точек из v относительно точек из t. Остается только проверить в среде, какой из вариантов прогноза имеет место. При этом в целом реализоваться в среде может один и только один из образов $\widetilde{V}_1, \ldots, \widetilde{V}_h$. Мало того, при построении гипотетических точек в их расположении относительно точек из t ни один из вариантов прогноза для одного образа не может совпасть ни с одним из вариантов прогноза для другого образа. Ибо такое совпадение означало бы, что у двух образов есть в их представителях аффинно эквивалентные части, а значит и общий подобраз, отличный от Z, а это противоречит тому, что образ Z для $\widetilde{V}_1, \ldots, \widetilde{V}_h$ — непосредственный подобраз.

Пусть в рамках этой процедуры из $\widetilde{V}_1, \ldots, \widetilde{V}_h$ был выбран один из образов, обозначим его через $\widetilde{V}'$. Если x есть представитель образа $\widetilde{V}'$ или аффинно эквивалентен части одного из представителей, то образ $\widetilde{V}'$ — искомый. Если нет, то повторяем описанную процедуру причем роль симплекс-образа теперь будет играть образ $\widetilde{V}'$: через v' обозначим представителя образа $\widetilde{V}'$, который является частью изображения x; определяем по ярусному графу W совокупность образов $\widetilde{K}_1, \ldots, \widetilde{K}_u$ образов, для которых образ $\widetilde{V}'$ является непосредственным подобразом. Далее опять совокупность $\widetilde{K}_1, \ldots, \widetilde{K}_u$ образов будет играть роль гипотез при расширении v' в рамках имеющегося изображения x. Опять наложением представителей образов $\widetilde{K}_1, \ldots, \widetilde{K}_u$ на v' получаем гипотетические места для точек и сравнением с точками из x выбираем один из образов в $\widetilde{K}_1, \ldots, \widetilde{K}_u$. Он либо искомый, либо описанное повторяется. В конечном счете мы находим искомый образ, обозначим его через $\widetilde{X}$, представителем или частью представителя которого является изображение x. Это процессу соответствует путь по графу W от начальной вершины, соответствующей симплекс-образу, до искомой вершины $\widetilde{X}$.

Распознавание в рамках описанной процедуры можно интерпретировать, как это видно, как исследование, анализ доступной части x среды S с использованием цепочки гипотез. Причем этот анализ и его результаты не зависят от того, в какой части среды S изображение x, или изображение, аффинно эквивалентное ему, находится. Анализ, можно сказать, ориентирован на работу с образами, безотносительно к тому, с каким конкретно представителем образа приходится иметь дело.

Что дает для оперирования в среде распознавание, то есть знание образа $\widetilde{X}$, представителем которого (или из бассейна притяжения которого) изображение x является? Во-первых, это возможность экстраполяционного прогноза разных уровней. Под экстраполяционным прогнозом первого уровня понимаем все те образы на W, непосредственным подобразом для которых является $\widetilde{X}$, то есть те образы на графе, расстояние до которых по графу равно единице. Образы с расстоянием два от $\widetilde{X}$ — второй уровень экстраполяционного прогноза, и т. д.

Во-вторых — возможность интерполяционного прогноза, то есть анализа того, какие подобразы содержатся в $\widetilde{X}$. Прогноз первого уровня — это все образы, которые непосредственно содержатся в $\widetilde{X}$, или, по другому, расстояние от которых по графу до $\widetilde{X}$ равно единице. Прогноз второго уровня — все те образы, расстояние от которых до $\widetilde{X}$ равно двум, и т. д.

Ранее в [3] говорилось на содержательном уровне о формировании средой системы образов живого существа, и о том, что разные среды формируют разные такие системы. В связи с этим интересно было бы понять, чем схожи и чем разнятся графы W, формируемые разными средами и классами сред.

Рассмотрим теперь вопрос о кодах среды и системы образов в ней.

Кодом изображения A называем пару $\langle M_A, T_A\rangle$. Здесь M_A — множество номеров точек изображения, T_A — множество всех чисел вида $\rho_{mnu,ksp} = S_{mnu}/S_{ksp}$, где S_{mnu} и S_{ksp} — площади треугольников в вершинах с точками соответственно m, n, u и k, s, p. При $S_{ksp} = 0$ полагаем $\rho_{mnu,ksp}$ не определенным. Изображения A и B с кодами $\langle M_A, T_A\rangle$ и $\langle M_B, T_B\rangle$ называем эквивалентными, если существует такая биекция $\psi : M_A \to M_B$, что $\rho_{mnu,ksp} = \rho_{\psi(m)\psi(n)\psi(u),\psi(k)\psi(s)\psi(p)}$. Если все точки изображения не лежат на одной прямой или двух параллельных прямых, то изображение называем плоским.

Теорема 2. [5, 6]. *Два плоских изображения эквивалентны тогда и только тогда, когда они аффинно эквивалентны.*

Содержательно теорема 2 означает, в частности, что код изображения задает его с точностью до аффинных преобразований.

Пусть B есть часть изображения A. Если код для A есть $\langle M_A, T_A\rangle$, то, очевидно, код $\langle M_B, T_B\rangle$ можно получить, если собрать в M_B номера из M_A всех точек, вошедших в B, и собрав в T_B все те $\rho_{mnu,ksp}$ из T_A, для которых m, n, u, k, s, p вошли в M_B. Говорим в этом случае, что код $\langle M_B, T_B\rangle$ есть часть кода $\langle M_A, T_A\rangle$.

Коды всех изображений всех образов — части кода среды S. Все изображения одного образа аффинно эквивалентны, и, соответственно, эквивалентны по кодам.

Известно, что для построения изображения A по коду $\langle M_A, T_A\rangle$ достаточно таких элементов $\rho_{mnu,ksp}$ из T_A, у которых тройки mnu и ksp разнятся только одним номером. Возникает вопрос: какова может быть роль других элементов $\rho_{mnu,ksp}$ в коде?

Назовем изображения A и B эквидистантными, если существует такая биекция $\psi : M_A \to M_B$, при которой для любых точек с номерами m, n, u из M_A (не лежащих на одной прямой), число $\rho_{mnu,\psi(m)\psi(n)\psi(u)}$ есть константа, не зависящая от выбора точек m, n, u.

Название «эквидистантные изображения» объясняется следующей аналогией с более простым случаем. Пусть A и B есть изображения, совместимые параллельным переносом. Тогда, очевидно, существует биекция $\psi : M_A \to M_B$ такая, что все расстояния $r(a, \psi(a))$ между соответствующими точками двух изображений есть константа. Отрезки $r(a, \psi(a))$ для всех a из M_A в этом случае не только равны, но и параллельны. Очевидно, имеет место и обратное: если все эти отрезки равны и параллельны, то A и B совместимы параллельным переносом.

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

Доказательство. Пусть A и B аффинно эквивалентны. Тогда существует такая биекция $\psi : M_A \to M_B$, что B переводится в B', совмещенное с A, то есть при этом каждая точка a из A совмещена с точкой $\psi(a)$. Ясно, что при этом для каждой тройки m, n, u из A (не лежащих на одной прямой) и соответствующей тройки k', s', p' из B' (здесь $k' = \psi(m), s' = \psi(n), p' = \psi(u)$) имеем $\rho_{mnu,k's'p'} = S_{mnu}/S_{k's'p'} = 1$. Вернем теперь обратным преобразованием B' в B. При этом площадь каждого треугольника с вершинами $k's'p'$ из B' при переводе в треугольник ksp с вершинами из B умножится на одну и ту же для всех треугольников величину k. Следовательно $\rho_{mnu,ksp} = S_{mnu}/S_{ksp} = S_{mnu}/kS_{k's'p'} = 1/k$.

Пусть теперь A и B эквидистантны. Если B' получено из B аффинным преобразованием, то, нетрудно видеть, A и B' тоже эквидистантны. Выберем некоторые три точки (не на одной прямой) m, n, u на A, и пусть преобразованное B' таково, что его точки k', s', p', соответствующие точкам m, n, u, совпали с ними, то есть площади треугольников mnu и $k's'p'$ равны. Но тогда, с учетом эквидистантности, должны быть равны площади и всех остальных соответствующих друг другу треугольников из A и B'. Однако это значит, что коды $\langle M_A, T_A \rangle$ и $\langle M_{B'}, T_{B'} \rangle$ эквивалентны, и, значит, в силу теоремы 1, изображения A и B' аффинно эквивалентны. Но тогда аффинно эквивалентны и изображения A и B. Теорема доказана.

Итак, прояснена роль элементов $\rho_{mnu,ksp}$ кода с полностью различными тройками mnu и ksp. Роль элементов с двумя различиями в этих тройках пока не ясна.

С содержательной точки зрения настоящая модель — лишь упрощенный, идеализированный вариант того, что нужно бы получить далее. В первую очередь потому, что в один образ мало включать изображения среды, отличающиеся только аффинными преобразованиями.

Список литературы

[1] Крушинский Л. В., Кудрявцев В. Б., Козлов В. Н. О некоторых результатах применения математики к моделированию в биологии // Математические вопросы кибернетики. — 1988. — Т. 1. — С. 52–88.

[2] Кудрявцев В. Б., Гасанов Э. Э., Подколзин А. С. Введение в теорию интеллектуальных систем. — М.: Издательский отдел факультета ВМиК МГУ; МАКС Пресс, 2006.

[3] Журавлев Ю. И. Об алгебраическом подходе к решению задач распознавания или классификации // Проблемы кибернетики. — 1978. — Т. 33. — С. 5–68.

[4] Hopfield J. J. Neural networks and physical systems with emergent collective computational abilities // Proceedings of the National Academy of Science. — 1982. 79.

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

[6] Kozlov V. N. Image Coding and Recognition and Some Problems of Stereovision // Pattern Recognition and Image Analysis. — 1997. Vol. 7. N 4. — P. 448–466.