

УДК 51:153:681.3.06

КВАЗИПСИХОЛОГИЧЕСКИЙ ПОДХОД  
К ПРОБЛЕМЕ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

А.С.Нудельман

1. Актуальность решения проблем, связанных с построением автоматов, достаточно хорошо имитирующих восприятие человеком зрительной информации, не вызывает сомнения и будет, по-видимому, возрастать в будущем. Для решения этих проблем к настоящему времени в распознавании зрительных образов сформировались два самостоятельных и достаточно независимых подхода. Один из них, называемый геометрическим, является, по существу, перенесением идей и методов распознавания, не ориентированных на какую-либо конкретную природу образов, на задачи, связанные с распознаванием изображений. Другой подход, называемый лингвистическим (структурным), связан только с задачами распознавания зрительных образов. Существующие методы распознавания, разработанные в рамках этих двух подходов, не могут еще обеспечить создания автомата, "восприятие" которого было бы сравнимо по качеству с восприятием человека.

В данной работе формулируется новый подход, названный квазипсихологическим, который представляет собой явную "психологизацию" лингвистического подхода и рассчитан на решение только таких задач распознавания, которые успешно решает зрительный анализатор человека.

2. На чем может быть основана вера в успешность квазипсихологического подхода к решению задачи распознавания изображений по сравнению с другими решениями?

Рассмотрим для начала изображения различных символов (букв, цифр и т.д.), порожденных самим человеком. Здесь прежде

всего поражает огромное разнообразие конкретных реализаций каждого символа (образа), если эти реализации рассматривать как физические объекты, т.е. как двумерные функции яркости. Но при всем физическом разнообразии у реализаций каждого образа должно быть что-то общее, и это общее связано скорее всего с психологией человеческого восприятия, поскольку все эти образы создавались с учетом удобства и легкости восприятия их человеком (как взрослым, так и ребенком).

Рассмотрим теперь зрительные образы, которые сформировались у человека при непосредственном наблюдении физического мира. Здесь тоже имеется большое физическое разнообразие элементов образа (примером могут служить такие образы, как: "кучевые облака", "перистые облака", "дерево" и т.д.). Образы и в этом случае не могут быть психологически сложными, поскольку в повседневную человеческую практику скорее всего могли войти те образы, которые наиболее просты в восприятии.

Рассмотрим, наконец, изображения, создаваемые человеком и отражающие свойства физического мира (осциллограммы, спектры, графики и т.д.). Здесь человек, используя физические преобразования сигналов, делает все возможное, чтобы изображения классифицировались самим человеком достаточно легко. Таким образом, разумно допустить, что в задачах, на решение которых рассчитан квазипсихологический подход, образы и, следовательно, границы между образами являются простыми не в физическом, а в психологическом плане.

3. Содержание понятия "психологическая простота" границ образов раскрывается из рассмотрения тех образов, которыми оперирует человек при естественном зрительном восприятии. Естественность восприятия проявляется, прежде всего, в отсутствии каких-либо вспомогательных (измерительных, вычислительных и т.п.) средств анализа изображения. Для реализации естественного восприятия достаточно, с точки зрения внешнего поведения человека, только наблюдения воспринимаемого изображения. Основной тезис квазипсихологического подхода состоит в том, что человек может воспринимать изображения (относить их к тому или иному образу) таким способом, что при этом восприятии анализ изображений происходит исключительно в терминах тех вещей, которые непосредственно наблюдаемы в изображении. В случае линейчатых

изображений такими вещами являются линии, точки, направления линий, углы. Если описание образа составлено в терминах непосредственно наблюдаемых вещей, то этот образ (значит, и его границу) можно считать психологически естественным, поскольку такой образ человек будет способен выделить при естественном восприятии. Подчеркнем, что основной тезис является не гипотезой о действительной психологии зрительного восприятия человека, а утверждением только того, что некий "психологический" механизм восприятия можно реализовать. Психологическая естественность и простота образов, с которыми человек имеет дело в повседневной практике и восприятие которых можно согласовать с основным тезисом, легко усматриваются на одном примере. Допустим, что учитель хочет научить кого-то читать цифру 8. Пусть ученик (возможно, ребенок!) уже знает, что такое "кружок". Процесс обучения может выглядеть так: учитель говорит, что цифра 8 — это два приблизительно одинаковых кружка, сцепленных (но не сильно) и расположенных один над другим; затем учитель рисует цифру 8 и приводит примеры, когда что-то в описании цифры 8 нарушено: ○ — один кружок, ○○ — два кружка, но не сцепленные, ⊙ — два кружка, но сцепленные сильно, ⊖ — два кружка, но совсем неодинаковые, ⊙ — два кружка, но расположенные не правильно. Понятно, что при восприятии цифры 8 ученик будет пользоваться описанием, данным ему учителем (граница образа "цифра 8" будет выражена через такие вещи, как: "кружок", "приблизительно одинаковые", "сцепленные" и т.д.). И здесь характерно то, что установление соответствия описания какому-либо изображению является психологически простым и будет происходить при естественном восприятии.

4. Примером использования "физически простых" границ между образами является геометрический подход [1] к распознаванию изображений. Суть этого подхода состоит в том, что формируется пространство всех (допустимых) изображений. Изображение представляется точкой такого пространства, образ — область в пространстве. Границы образов выражаются либо непосредственно, в виде некоторых физически простых гиперповерхностей, либо косвенно, через эталонные изображения и физически простую меру сходства изображений. Неадекватность таких границ приводит к необходимости усложнения физических границ образов. Это выра-

жается в том, что используются либо композиции исходных разделяющих гиперповерхностей, либо несколько эталонов на один образ. Происходит как бы кусочно-физическая аппроксимация естественных границ образов.

Отметим два момента. Первый: при естественном восприятии человек не оперирует образами, имеющими кусочно-физические границы. Действительно, для любого геометрического алгоритма распознавания, область допустимых изображений которого не является крайне узкой (например, фиксированный печатный шрифт), можно найти допустимое изображение, которое человеком воспринимается вполне уверенно и на котором алгоритм совершает ошибку, а это говорит о том, что образ, которым пользуется человек, отличается от образа, реализуемого алгоритмом. Второй момент: с образами, имеющими кусочно-физические границы, человек оперировать не может. Это проявляется в том, что человек не способен выразить интуитивно понимаемую им границу образа в кусочно-физической форме. Именно поэтому такую границу всегда строит автомат, который по необходимости должен быть самообучающимся. Здесь возникает ситуация: к достаточно сложной и еще не решенной проблеме выбора вида границ образов "привязывается" не менее сложная проблема самообучения автомата. По всякой же разумной методологии, прежде чем заниматься проблемой обучения чему-то, необходимо выяснить, нужно ли обучать этому.

Возникновение лингвистического подхода [2] к решению задач распознавания зрительных образов вызвано тем, что исследователь обратил внимание на психологию зрительного восприятия человека. В этом подходе уже учитывается специфика зрительных образов. Однако психология зрительного восприятия представлена здесь недостаточно, и сохранен по меньшей мере один психологически неоправданный прием геометрического подхода. Этот прием состоит в том, что образуется пространство описаний (каждому изображению сопоставляется точно одно описание) и в этом пространстве образы определяются (как и в геометрическом подходе) с помощью эталонов-описаний и некоторой меры сходства описаний.

5. Основная задача квазипсихологического подхода состоит в разработке совокупности формальных понятий, настолько развитой, чтобы в терминах этих понятий можно было бы строить логику "видящих" автоматов, и настолько психологически приемлемой,

чтобы в терминах этих понятий можно было бы выражать психологические аспекты зрительного восприятия. Именно психологическая направленность понятий позволяет надеяться на эффективность применения этих понятий для построения распознающих автоматов.

Методологическую базу квазипсихологического подхода можно сформулировать так: все понятия, методы и принципы, на основе которых предполагается строить распознающий автомат, должны быть такими, чтобы сам человек мог естественно воспринимать зрительную информацию, строго следуя алгоритму этого распознающего автомата. Достаточно развитая совокупность понятий, методов и принципов, построенная в соответствии с этой методологической концепцией, позволяет, по мнению автора, достичь внешнефункциональную адекватность естественного восприятия человека и поведения распознающего автомата. Такой адекватности вполне достаточно при решении задач, на которые рассчитан квазипсихологический подход.

6. Полезность и эффективность понятий, вводимых в соответствии с квазипсихологическим (и любым другим) подходом к распознаванию зрительных образов, может быть установлена только экспериментально — постройкой автоматов, достаточно хорошо имитирующих зрительное восприятие человека. Автору представляется необходимым оценивать качество "восприятия" автомата по крайней мере двумя способами: определением типичности и определением показателя маскировки [5]. Понятие типичности (на множестве изображений  $R$  среди множества людей  $T$ ) является обобщением понятия надежности распознавания на случай произвольного (по качеству) множества изображений. Показатель маскировки (на множестве изображений  $R$  среди множества людей  $T$ ) характеризует внешнефункциональную "человечность" распознающего автомата. Показатель маскировки ( $\mu$ ) всегда больше или равен нулю. Если  $\mu < 1$ , то это означает, что автомат выделяется среди людей из  $T$  именно как нечеловек. Если  $\mu > 1$ , то это означает, что автомат чаще принимается за человека, а не за машину. Если  $\mu = 1$ , то это означает, что на множестве изображений  $R$  среди множества людей  $T$  автомат "ведет" себя так же, как и человек.

7. Решение основной задачи распознавания, т.е. построение алгоритма распознавания  $n$  образов на классе изображений  $K$ ,

проведенное в рамках квазипсихологического подхода, складывается из трех стадий. На первой стадии исследователю необходимо самому научиться уверенно распознавать заданные образы в заданном классе изображений. Предполагается, что на этой стадии происходит интуитивное восприятие (распознавание). Вторая стадия состоит в выявлении совокупности понятий, в терминах которых исследователь может описать на идейном неформальном уровне как свои представления об исходных образах, так и некоторую схему алгоритма распознавания. Схема алгоритма должна быть такой, чтобы сам исследователь, наблюдая конкретное изображение и "воспринимая" это изображение по правилам схемы, вынужден был принимать решение (о принадлежности изображения к определенному образу), согласующееся с решением, полученным в результате естественного, интуитивного восприятия. На этой же стадии составляется неформальная характеристика класса изображений  $K$ . На третьей, заключительной, стадии строятся формальные аналоги всех понятий, выявленных на предыдущей стадии. Формальный и конструктивный аналог схемы алгоритма представляет собой искомый алгоритм распознавания. Поскольку алгоритм распознавания "перерабатывает" формальное представление изображения, необходимо предусматривать техническую реализуемость устройства, осуществляющего перевод реального изображения в его формальный аналог (т.е. техническую реализуемость "глаза"). Решение более широкой задачи распознавания (самообучающиеся алгоритмы, алгоритмы с автоматическим нахождением признаков и т.д.) происходит аналогично решению основной задачи.

8. Рассмотрим три особенности квазипсихологического подхода, отличающие его от лингвистического (и геометрического) подхода, на примере решения одной задачи. Задача состоит в разработке понятий, методов и принципов, предназначенных для построения распознающих автоматов, имитирующих восприятие человеком плоских линейчатых изображений. Квазипсихологическое решение такой задачи описано в [3,4,6] и представляет собой совокупность формальных понятий, которую будем для краткости называть  $KB$  - системой (системой понятий для конструктивного описания зрительного восприятия).

Основным понятием в  $KB$  - системе является понятие языка описания, представляющего собой множество предложений, вы-

полняющих функцию описаний соответствующих образов (классов изображений). Попутно заметим, что язык описания - формальный аналог множества представлений реального зрительного анализа - тора о тех образах, которые этот анализатор способен выделять в классе всех линейчатых изображений. Правила образования предложений равносильны правилам образования формул прикладного исчисления предикатов. Каждое предложение языка описания есть утверждение о существовании (в наблюдаемом изображении) некоторой конструкции из линий. В каждом предложении говорится только о линиях, точках, направлениях и углах (поскольку формализуется непосредственное восприятие), т.е. о таких объектах, которые обычно используются в работах по распознаванию зрительных образов. Множество всех предложений (описаний) в  $KB$  - системе является бесконечным.

Фундаментальное отличие квазипсихологического подхода от лингвистического связано с отношением между описанием и изображением. В лингвистическом подходе каждому изображению сопоставляется точно одно описание. Вследствие этого можно сказать, что описание - предложение утверждает все свойства изображения (выразимые на языке, в терминах которого составляются описания) или, другими словами, что при сопоставлении изображению его описания происходит простая перекодировка изображения на абстрактный язык (абстрактный в том смысле, что одно и то же описание может быть сопоставлено физически различным изображениям). В квазипсихологическом подходе между изображением и предложением используется не процедура перекодировки, а отношение истинности. Для любой пары <предложение, изображение> всегда можно установить, истинно или ложно данное предложение на данном изображении. Если предложение истинно на изображении, то это означает, что изображение обладает всеми свойствами, которые утверждаются предложением и, следовательно, принадлежит образу, формальным описанием которого является это предложение.

Таким образом, в квазипсихологическом подходе рассмотрение изображения не как физического объекта, а как элемента образа возможно только "с точки зрения" уже имеющегося описания этого образа (предложения). Другими словами, характер "психологической обработки" изображения существенно зависит от образа, принадлежность к которому должна быть выяснена в результа-

те такой "обработки". Такая ситуация, известная в психологии восприятия как "установка" [8] на восприятие, более полно отражает отношение человека к воспринимаемому им изображению, поскольку в этом изображении человек может "искать" в один момент одну конструкцию из одних линий (выяснить принадлежность одному образу), в другой момент — другую конструкцию, состоящую из других линий (выяснить принадлежность другому образу), и т.д.

Покажем это на примере. Пусть изображение есть отрезок прямой линии, наклоненный под углом  $45^\circ$  к вертикали и направленный вправо — вниз (отрезок прямой направления  $\vec{E}$ ). В этом изображении отрезок прямой (физический объект) может восприниматься в виде самых различных психологических образов: 1) горизонтальной линии, направленной вправо и имеющей "допуск" по направлению  $\pm 90^\circ$  (1 — прямая горизонтальная [3, стр. 52; 4; стр. 81]); 2) вертикальной линии, направленной вниз и имеющей "допуск" по направлению  $\pm 45^\circ$  (2 — прямая вертикальная); 3) линии направления  $\vec{E}$  с "допуском" по направлению  $\pm 22,5^\circ$  (3 — прямая направления  $\vec{E}$ ) и т.д.

Второе отличие квазипсихологического подхода связано с характером физических преобразований изображений, используемых для улучшения качества распознавания. В лингвистическом подходе физические преобразования (заполнение разрывов, стирание "шума") выполняются до лингвистического (психологического) анализа изображения, и после выполнения этих преобразований информация о том, что изменено в изображении, теряется. Такая процедура физических преобразований никогда не исключает того, что в изображение добавится нечто существенно искажающее или из изображения исключится нечто существенно необходимое. Кроме того, такая процедура не кажется психологически естественной. Действительно, если человек улучшает качество изображения, то он всегда знает, какое содержание изображения (образ) необходимо оставить (если что-то стирается) или подчеркнуть (если что-то дорисовывается). При этом всегда фиксируется, что и "сколько" изменено. В квазипсихологическом подходе основной процедурой алгоритма распознавания является процедура установления значения истинности предложения на "воспринимаемом" изображении. Физические преобразования изображения непосредственно входят в

эту процедуру, и от смысла исходного предложения зависит то, какой конкретной перестройке подвергается изображение. "Величина" физической перестройки (психологическое "усилие" для ее выполнения) формализуется понятием интенсивности перестройки. В КВ-системе имеются три вида физических преобразований изображения: спрямление линий (осуществляется неявно использованием понятия условной истинности), достройка существенных линий (осуществляется непосредственно алгоритмами достройки) и стирание несущественных линий (осуществляется неявно тем, что при установлении значения истинности в изображении "выделяются" только те линии, о существовании которых говорится в предложении, — остальные линии в изображении как бы не "видятся"). Тот факт, что общая достройка (дополнение) изображения [4] выполняется до "психологической обработки" и независимо от нее, не противоречит вышесказанному, так как в идее общей достройки выражена часть наиболее простых физических преобразований, рассчитанная на все виды "психологической обработки" (на все предложения). Одноразовая же достройка всех допустимых отрезков значительно сокращает время, затрачиваемое на распознавание.

Наконец, третья особенность квазипсихологического подхода связана с вопросом самообучения автоматов. Пусть задача распознавания состоит в разработке автомата (алгоритма), назначение которого — распознавать фиксированный класс образов. Режим самообучения такого автомата нельзя назвать, вообще говоря, необходимым, поскольку этот режим работы не является эксплуатационным. Выше (п.4) было отмечено, что геометрические (часто и лингвистические) алгоритмы используют методы самообучения не в силу специфики задачи, а вследствие специфики подхода к решению задачи, что не может выглядеть достаточно естественным. Кроме того, корректная разработка принципов самообучения автомата не может быть проведена при отсутствии подходящего критерия оценки качества этих принципов. Сейчас такого критерия нет, оценка же качества собственно самообучения подменяется оценкой качества уже обученного состояния автомата, куда органически входят и качество решающих функций (границ образов).

В квазипсихологическом подходе решение любой задачи распознавания может быть осуществлено исключительно в пределах требований к задаче без привлечения дополнительных принципов. Действительно, автору [5,7] удалось построить два автомата,

"восприятие" которых сравнимо (по показателю маскировки) с восприятием человека и которые не являются самообучающимися. Особенности квазипсихологического подхода приводят к необходимости разработки новых принципов самообучения автоматов, поскольку существующие базируются на сопоставлении каждому изображению только одного описания (в геометрическом подходе - точки), имеющего к тому же сравнительно простую логическую структуру.

В [3,4,6] вопросы самообучения не рассматривались, это - предмет будущих исследований. Здесь же с точки зрения исходной методологической позиции можно сказать только одно: прежде всего необходимо выработать конструктивное представление о том, каким образом может происходить переход от необученного состояния к обученному у человека, и проверить полезность такого представления психологическим экспериментом. В классе задач, охватываемых квазипсихологическим подходом, самообучаемость автомата необходимо оценивать в сравнении со способностью обучаться, присущей человеку. Такое сравнение можно проводить, например, в следующем эксперименте: человек и автомат обучаются на некотором множестве изображений, причем образы должны быть неизвестны ни человеку, ни автомату; затем с помощью новых изображений проверяется адекватность представлений об образах, сформированных человеком и автоматом (здесь следует особо подчеркнуть, что обучающая выборка является одной и той же во всех отношениях для человека и для автомата).

9. В заключение отметим, что, как показывают результаты экспериментов, описанных в [5,7], квазипсихологический подход к проблеме распознавания изображений позволяет строить автоматы, достаточно хорошо имитирующие восприятие человеком зрительной информации.

#### Л и т е р а т у р а

1. ВАСИЛЬЕВ В.И. Распознающие системы. (Справочник). Киев, "Наукова думка", 1969.

2. ЗАВАЛИНИН Н.В., МУЧНИК И.Б. Лингвистический (структурный) подход к проблеме распознавания образов. (Обзор). - "Автоматика и телемеханика", 1969, № 8, с. 86-118.

3. НУДЕЛЬМАН А.С. Об одном формализованном подходе к изучению зрительного восприятия. - В кн.: Вычислительные системы. Вып. 55. Новосибирск, 1973, с. 36-78.

4. НУДЕЛЬМАН А.С. Активное восприятие зрительной информации. - Там же, с. 79-83.

5. НУДЕЛЬМАН А.С. Эксперимент по сравнению зрительного "восприятия" машины с восприятием человека. - Там же, с. 84-89.

6. НУДЕЛЬМАН А.С. Принцип целевой достройки в активном восприятии зрительной информации. - Настоящий сборник, с. 69-72.

7. НУДЕЛЬМАН А.С. Эксперимент "цифровое письмо". - Настоящий сборник, с. 73-82.

8. УЗНАДЗЕ Д.Н. Психологические исследования. М., "Наука", 1966.

Поступила в ред.-изд.отд.  
8 мая 1974 года