

А. Г. Аркадьев, Э. М. Браверман

**Обучение машины
распознаванию образов**

**Москва
«Книга по Требованию»**

УДК 621.39
ББК 32
А11

- A11 **А. Г. Аркадьев**
Обучение машины распознаванию образов / А. Г. Аркадьев, Э. М. Браверман – М.: Книга по Требованию, 2012. – 110 с.

ISBN 978-5-458-25154-9

Вопрос о том, как нам удастся отличать, например, мужские портреты от женских или буквы «а» от букв «б», словом, вопрос о том, как мы различаем образы, гораздо интереснее и сложнее, чем это может показаться на первый взгляд. Своеобразие вопроса состоит в том, что научиться узнавать сколько угодно большое число представителей образа (например, любые женские портреты, в отличие от мужских, или какую-нибудь букву во всевозможных начертаниях) оказывается возможным на основе знакомства с ограниченным их числом (с несколькими портретами или начертаниями буквы)...

ISBN 978-5-458-25154-9

© Издание на русском языке, оформление
«YOYO Media», 2012

© Издание на русском языке, оцифровка,
«Книга по Требованию», 2012

Эта книга является репринтом оригинала, который мы создали специально для Вас, используя запатентованные технологии производства репринтных книг и печати по требованию.

Сначала мы отсканировали каждую страницу оригинала этой редкой книги на профессиональном оборудовании. Затем с помощью специально разработанных программ мы произвели очистку изображения от пятен, клякс, перегибов и попытались отбелить и выровнять каждую страницу книги. К сожалению, некоторые страницы нельзя вернуть в изначальное состояние, и если их было трудно читать в оригинале, то даже при цифровой реставрации их невозможно улучшить.

Разумеется, автоматизированная программная обработка репринтных книг – не самое лучшее решение для восстановления текста в его первозданном виде, однако, наша цель – вернуть читателю точную копию книги, которой может быть несколько веков.

Поэтому мы предупреждаем о возможных погрешностях восстановленного репринтного издания. В издании могут отсутствовать одна или несколько страниц текста, могут встретиться невыводимые пятна и кляксы, надписи на полях или подчеркивания в тексте, нечитаемые фрагменты текста или загибы страниц. Покупать или не покупать подобные издания – решать Вам, мы же делаем все возможное, чтобы редкие и ценные книги, еще недавно утраченные и несправедливо забытые, вновь стали доступными для всех читателей.

Эти машины, позволяя моделировать малоизученный психический процесс, могут оказаться одним из средств изучения функций высшей нервной деятельности.

Для обширной литературы, посвященной проблеме узнающих машин, характерно чрезвычайное разнообразие точек зрения и подходов к решению практических и теоретических вопросов; многие практические достижения в этой области кажутся результатами удачных догадок. Например, знаменитый «персептрон» американского ученого Ф. Розенблатта «научился» различать зрительные образы значительно раньше, чем стало ясно, как он это делает.

В предлагаемой читателю книге в качестве руководящей идеи использована выдвинутая одним из авторов (Э. М. Браверманом) достаточно общая концепция, так называемая *гипотеза о компактности образа*. Эта концепция позволила, с одной стороны, построить описанные ниже узнающие машины и наметить пути их совершенствования, а с другой — объяснить действие некоторых узнающих машин, построенных за рубежом. В книге приведены также результаты экспериментов с узнающими машинами, проведенных Э. М. Браверманом в сотрудничестве с О. А. Башкировым и И. Б. Мучником.

Авторы учитывали необходимость дать общее представление об узнающих машинах специалистам смежных областей — биологам, психологам, медикам, — которые с разных точек зрения интересуются этими машинами, но не всегда в состоянии одолеть специальную кибернетическую литературу. Книга доступна читателю, владеющему математикой в объеме средней школы.

Как уже говорилось, в основу книги положены работы одного определенного направления; работы других создателей узнающих машин описаны лишь в той мере, в которой они могут быть трактованы с точки зрения гипотезы

компактности и в свете основанных на этой гипотезе идей. Вследствие этого книга ни в коей мере не претендует на полный обзор проблемы узнающих машин.

Помимо изложенного в книге подхода к созданию таких машин (он может быть назван «геометрическим» подходом), разрабатываются и другие направления. Из них наиболее перспективными представляются работы, основанные на изучении вероятностных свойств подлежащих различению образов [9—13].

Попытки более детально моделировать процесс узнавания приводят к задаче выработки полезных признаков изображения. Один из возможных путей решения этой задачи рассматривается в последней главе предлагаемой книги. Из имеющихся в настоящее время работ, посвященных этой проблеме, следует отметить работы [14, 15].

Ближайшие практические результаты даст, по-видимому, направление, основанное на различении образов по заранее запрограммированным признакам [16—25].

В приложенном в конце книги списке литературы авторы не стремились дать сколько-нибудь исчерпывающие сведения о работах, посвященных проблеме узнающих машин. Список включает в себя лишь тот минимум литературы, который, по мнению авторов, может послужить исходным материалом для читателя, серьезно заинтересовавшегося рассматриваемыми в книге проблемами.

А. Г. Аркадьев, Э. М. Браверман

ГЛАВА ПЕРВАЯ

ПОНЯТИЕ ОБРАЗА

Воспринимая явления внешнего мира, мы всегда производим их классификацию, т. е. разбиваем эти явления (предметы, ситуации) на группы «похожих» явлений (именно «похожих», а не тождественных). По тем или иным причинам оказывается необходимым отнести в одну группу значительно отличающиеся друг от друга, но в чем-то сходные явления или предметы. Например, все фигуры, изображенные на рис. 1, мы называем «буквой А», несмотря на большие различия в их начертании.



Рис. 1.

Весьма существенно, что, выделив такие группы или множества* объектов, мы получаем способность «узнавать», то есть относить к уже известному множеству, новые, ранее не встречавшиеся нам объекты, например узнавать буквы, написанные новым для нас почерком. Составив понятие о том, что такое «буква А» на основе некоторого, обычно не очень большого, количества начертаний («экземпляров») этой буквы, мы оказываемся в состоянии

* Здесь и далее мы применяем термин «множество» не в его обиходном смысле, т. е. не в значении «много», а в смысле математическом, как совокупность предметов, определенную каким-либо способом. Множество, понимаемое в этом смысле, может состоять из бесконечного числа представителей (например, «атомы вселенной», «натуральные числа»), ограниченного их числа («книжки моей библиотеки») и даже не содержать представителей вообще («бессмертные люди»).

узнать сколь угодно большое, практически бесконечное число других ее экземпляров*. То же самое можно сказать и о других буквах, цифрах, а также о многих других множествах изображений, предметов и явлений.

Вместе с тем далеко не все множества объектов дают нам возможность, ознакомившись с небольшой частью множества, узнавать затем сколь угодно большое число других, ранее нам не известных его представителей. Например, фотографии читателей этой книги образуют некоторое множество, однако нет возможности, ознакомившись, скажем, с фотографиями десяти людей, читавших эту книгу, определить по внешности одиннадцатого, держал он ее в руках, или нет.

Итак, существуют множества некоторого особого типа. Эти множества обладают характерным свойством, которое проявляется в том, что люди (и животные), ознакомившись с конечным числом объектов множества, оказываются способными узнавать (т. е. выделять из разнообразных предметов или явлений) сколь угодно большое число других его представителей. Множества этого типа мы будем в дальнейшем называть *образами*. Как видно из этого определения, термину «образ» здесь придается иной смысл по сравнению с принятым в теории познания.

Примерами образов могут служить множества: «мужчины», «женские портреты», «портреты Пушкина», «картины Врубеля», «буквы а», «цифры 5». Применять термин «образ» в указанном смысле дает нам право то обстоятельство, что «ознакомление» с образом не связано с запоминанием отдельных объектов, а узнавание нового объекта происходит без его сравнения с каждым из ранее известных. По-видимому, в нашем сознании формируются некие общие представления о группах сходных явлений, создаются как бы образы этих явлений, что и позволяет нам классифицировать бесконечные множества объектов, ознакомившись с конечным их числом.

Восприятие явлений в форме образов играет важнейшую роль в процессах познания внешнего мира. Оно позволяет более экономно использовать память, так как осво-

* Конкретные экземпляры представителей множеств, например «первая буква А на пятой странице этого экземпляра книги», в дальнейшем будем называть «объектами».

бождает нас от необходимости запоминать несчетное множество конкретных явлений, и, что, вероятно, наиболее важно, оно позволяет использовать накопленный ранее опыт. Не обладая способностью группировать объекты в образы, мы становились бы втупик перед каждым новым явлением, ибо ни одно из них не есть точная копия встречавшихся ранее явлений. Мы не могли бы, например, разбираться в незнакомом почерке, читать книги, напечатанные шрифтом нового для нас размера и, строго говоря, вообще не могли бы читать, так как не только в письмах, но и в книгах не найдется и двух букв, абсолютно сходных между собой (хотя бы из-за дефектов печати и неизбежных различий между литерами набора).

Как следует из самого определения образа, «узнаванию» новых для нас объектов образа всегда предшествует процесс обучения. Во время обучения мы знакомимся с некоторым количеством объектов и, кроме того, из какого-либо источника (например, от более опытного человека, «учителя»), получаем информацию о том, к какому образу относится каждый из этих объектов. Именно так мы учимся грамоте (учитель называет нам буквы), обучаемся отличать музыку Шопена от музыки Шостаковича (афиша или концертная программа сообщают нам имя композитора), приобретаем способность, взглянув на картину, узнавать кисть Врубеля (вначале мы читаем его имя на музейных этикетках).

Характерное свойство образов объективно в том смысле, что разные люди, обучавшиеся на различных группах объектов образа, большей частью одинаково и независимо друг от друга классифицируют одни и те же новые объекты. Именно объективность этого свойства образов, его независимость от индивидуальных особенностей обучающихся позволяет, например, людям, учившимся в разных школах, у разных учителей, успешно читать один и тот же ранее не известный им почерк. Однако, обладая указанным объективным свойством, сами образы в то же время до некоторой степени расплывчаты, так как вопрос о том, принадлежит объект к данному образу или нет, не всегда может быть решен однозначно.

На рис. 2 показано, как можно последовательными небольшими изменениями превратить цифру три в пятерку.

Между «хорошей» тройкой и «хорошей» пятеркой лежит целый ряд фигур. Часть этих фигур, хотя они и «похуже», мы все уверенно примем за тройку и пятерку, а принадлежность другой части фигур к «пятеркам» или «тройкам» вызовет у нас сомнения. Если предложить разным людям разбить рис. 2 на «пятерки», «тройки» и «нецифры», резуль-

5 5 5 5

5 5 3 3

3 3 3 3

Рис. 2.

таты не у всех будут одинаковы. Такое же явление будет наблюдаться и в случае иных множеств — цифр, букв и т. д.

Чтобы исключить подобную неоднозначность, можно прибегнуть к голосованию, условившись называть тройкой (пятеркой, семеркой, буквой «а» и т. д.) объекты, которые будут признаны таковыми определенным большинством, например не менее чем семьюстами людей из каждой ты-

сячи. В дальнейшем мы будем подразумевать, что границы образа определены именно таким способом.

Помимо только что сказанного, рис. 2 иллюстрирует еще одну важную особенность объектов любого образа. Эта особенность состоит в том, что можно в более или менее широких пределах изменять объект, и он все-таки останется объектом того же образа. Тройка не перестанет быть тройкой, если немного ее искривить, или, например, пропустить в начертании часть контура (рис. 3).

3 3 3

Рис. 3.

Впрочем, объекты образов не всегда одинаково стойки по отношению к изменениям того или другого характера. Примером может служить вторая слева во втором ряду фигура на рис. 2. Достаточно небольших изменений ее верхней части, чтобы она из «очень плохой» пятерки стала либо «сносной» пятеркой, либо «очень плохой» тройкой, либо «не цифрой». В то же время при значительных деформациях нижней части фигуры она останется «пятеркой». «Хорошие» пятерка и тройка более устойчивы — нужны значительные искажения, чтобы превратить их в другие цифры или в «нецифры».

Из того, что объект может подвергаться более или менее серьезным искажениям, оставаясь при этом объектом того же образа, следует также, что логические понятия не являются образами. Не являются образами, например, треугольники или окружности в их геометрическом смысле. Достаточно слегка деформировать окружность или выкинуть хотя бы одну точку из стороны треугольника, чтобы они для геометра перестали быть таковыми. В житейском же смысле слова треугольник и круг — образы. Именно так мы назовем фигуры, изображенные на рис. 4.

Объективный характер основного свойства образов позволяет ставить вопрос о том, можно ли моделировать на машине процесс обучения распознаванию образов. Мы отвечаем на этот вопрос утвердительно. Правда, пока мы умеем строить * машины, обучающиеся различению только достаточно простых зрительных образов типа цифр или букв, иными словами, сегодня мы знаем, как построить

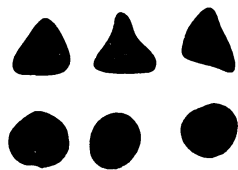


Рис. 4.

машину, способную научиться «читать». Позволительно спросить, впрочем, нужно ли обращаться к понятию образа, чтобы построить «читающую», т. е. различающую буквы или цифры машину. Казалось бы, можно обойтись более простыми средствами. Можно, например, ввести в память машины эталонные изображения подлежащих «чтению» букв (или цифр). В этом случае «узнавание» предъявленной машине буквы будет производиться путем последовательного сравнения этой буквы со всеми буквами эталонного набора.

Такие машины, действительно, могут быть построены, но они обладают существенными недостатками. Поскольку емкость машинной памяти на практике всегда ограничена, машины этого типа способны «читать» ограниченное число, скажем, шрифтов и почерков. Увеличение числа «читаемых» машиной шрифтов замедляет процесс чтения, так

* Выражение «строить машину» не обязательно подразумевает действительное ее изготовление; во многих случаях достаточно разработать программу воспроизведения специальной (например, узнающей) машины на универсальной вычислительной машине.

как приводит к необходимости сравнивать каждую подлежащую распознаванию букву или цифру с большим числом эталонных изображений. Кроме того, подобные машины весьма чувствительны к дефектам печати, т. е. к мелким, но чрезвычайно разнообразным искажениям фигур, характер которых практически невозможно предусмотреть.

Возможен и другой способ построения читающих машин, основанный на различении каких-либо признаков подлежащих распознаванию фигур. В качестве признаков

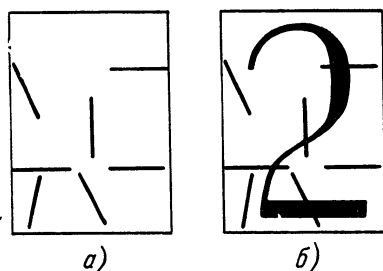


Рис. 5.

могут быть выбраны различные особенности фигур, например их геометрические свойства (характеристики составляющих фигуры кривых), топологические свойства (взаимное расположение элементов фигуры) и т. п. Известны читающие машины, в которых различение букв или цифр

производится по так называемому «методу зондов», т. е. по числу пересечений контура фигуры с несколькими особым образом расположенными прямыми. На рис. 5, а показано расположение таких прямых («зондов»), применявшееся для чтения арабских цифр. Если проектировать цифры на поле с зондами (рис. 5, б), то окажется, что каждая из цифр пересекает вполне определенные зонды, причем комбинации пересекаемых зондов различны для всех десяти цифр. Эти комбинации и используются в качестве признаков, по которым производится различение цифр.

Такие машины успешно справляются, например, с чтением машинописного текста, но их возможности ограничены тем шрифтом (или группой сходных шрифтов), для которого была разработана система признаков.

Итак, описанные методы построения читающих машин прежде всего недостаточно универсальны. Каждая такая машина может оперировать только с вполне определенным набором подлежащих различению фигур (например, чи-

тать только буквы или только цифры), и с вполне определенным шрифтом или группой сходных шрифтов. Поэтому при переходе к другому шрифту или при необходимости различать вместо цифр буквы может потребоваться полная замена содержимого машинной памяти или изменение структуры самой машины.

Еще более существенный для нас недостаток описанных машин состоит в том, что работа по созданию набора эталонных фигур или системы признаков должна производиться человеком. Качество работы машины, т. е. надежность «узнавания» предъявляемых фигур определяется качеством этой предварительной подготовки и без участия человека не может быть повышено. Иными словами, описанные машины не являются машинами обучающимися.

Нас же интересует прежде всего моделирование процесса обучения распознаванию образов. При этом в термин «моделирование процесса обучения» мы вкладываем следующее интуитивно оправданное содержание: обучению не предшествует сообщение машине каких-либо сведений о тех образах, распознаванию которых она должна научиться; само обучение заключается в предъявлении машине некоторого конечного числа объектов каждого образа. В результате обучения машина должна оказаться способной узнавать сколь угодно большое число новых объектов, относящихся к тем же образам.

Таким образом, имеется в виду следующая схема экспериментов:

а) никакие сведения о подлежащих различению образах в машину заранее не вводятся,

б) в ходе обучения машине предъявляется некоторое количество объектов каждого из подлежащих различению образов и сообщается, к какому образу относится каждый объект,

в) машина автоматически обрабатывает полученную информацию, после чего

г) с достаточной надежностью различает сколь угодно большое число новых, ранее ей не предъявлявшихся объектов каждого из образов.

Машины, работающие по такой схеме, мы будем называть *узнающими* машинами.

То обстоятельство, что подлежащие распознаванию множества в этой схеме названы образами, весьма существенно. Выше мы видели, что не все множества, а только образы, дают возможность человеку, ознакомившись с конечным числом объектов, узнавать затем бесконечное их число. Это и заставляет нас обращаться к понятию образа при создании машин, различающих простые фигуры. Если окажется возможным найти некоторое объективное свойство, присущее каждому простому зрительному образу и отсутствующее у множеств, не являющихся образами, если, далее, окажется возможным дать этому свойству формальное описание, в наших руках, быть может, окажется метод, позволяющий строить универсальные узнающие машины. Универсальными эти машины будут в том смысле, что без каких-либо изменений в структуре или программе они смогут быть обучены распознаванию любых однотипных по сложности образов. Такую машину можно будет, например, обучив различению славянских букв, «переучить» и заставить различать латинские буквы, а затем вновь «переучить» так, чтобы она различала цифры.

Безусловно, узнающие машины могут представлять интерес, значительно выходящий за рамки узких прикладных задач построения читающих автоматов; они позволяют нам глубже понять общий характер процессов обучения, поскольку процесс обучения, выполняемый узнающими машинами, по своему содержанию значительно богаче, чем у других, ранее изучавшихся типов обучающихся машин, например, у машин, моделирующих образование условных рефлексов, или у самонастраивающихся регуляторов.

Устройства, моделирующие образование условных рефлексов, отвечают конечным числом реакций на конечное число внешних воздействий. Процесс «выработки» условного рефлекса состоит в переборе всех реакций на каждое внешнее воздействие и запоминании «удачных» реакций. «Обучение» в этом случае заключается в запоминании найденной путем перебора зависимости между конечным числом заранее заданных внешних воздействий и конечным числом реакций, также заранее заданных.

Самонастраивающиеся регуляторы автоматически находят минимум (или максимум) некоторой функции, которая по условию имеет только один минимум (максимум). Простейшую процедуру такого рода можно представить следующим образом: находясь в