

Загоруйко Н. Г.

**Методы распознавания и их
применение**

**Москва
«Книга по Требованию»**

УДК 50
ББК 22
3-14

- 3-14 **Загоруйко Н. Г.**
Методы распознавания и их применение / Загоруйко Н. Г. – М.: Книга по Требованию, 2012. – 211 с.

ISBN 978-5-458-31894-5

Книга посвящена одной из наиболее актуальных проблем кибернетики — автоматическому распознаванию образов. В первой части обсуждаются вопросы методики распознавания. Описываются алгоритмы решения основных типов задач распознавания: выработка правил принятия решений, оценка возможностей распознающих устройств, выбор информативной системы признаков, установление исходной классификации (алгоритм «самообучения»). Рассматриваются и более сложные задачи. Вторая часть посвящена описанию приложений методов распознавания. Приводятся примеры решения задач в технике связи, медицине, геологии, экспериментальной физике, социологии. Книга рассчитана на специалистов и студентов, занимающихся применением кибернетических методов в различных областях знаний — в проектировании и эксплуатации высокопроизводительных средств управления, биологии и медицине, геологии и геофизике, социологии и истории.

ISBN 978-5-458-31894-5

© 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 первоизданном 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.

знания образов в таких областях, как прогнозирование погоды (в частности, для предсказания ураганов), ботаника и зоология (классификация видов по признакам), геофизика и геология (прогнозирование землетрясений, классификация горных пород, прогнозирование перспективности разведываемого района), социология (стратификация анкетированных, поиск определяющих социальных факторов, прогнозирование социальных ситуаций) и т. д.

О значении проблемы обнаружения и классификации для оборонной техники говорит такой факт. По данным американской фирмы Хонивелл, разработавшей систему ПАТТЕРН [6], (метод установления важности систем и научных проблем) для достижения целей нации, связанных с исследованием космоса, проблема обнаружения и распознавания стоит на первом месте из семнадцати наиболее важных проблем. В числе конкурентов проблемы распознавания были такие проблемы, как средства доставки, средства поражения и т. д. Несмотря на это, ее относительный вес 0,22, в то время как общий вес остальных шестнадцати — 0,78.

Признание важности и перспективности применения методов распознавания в различных областях привело в последние годы к стремительному увеличению количества публикаций на эту тему. Круг специалистов, занимающихся распознаванием, быстро пополняется инженерами, врачами, геологами, социологами, математиками. Они решают разные задачи, разными методами и при этом говорят на разных языках. Возникла необходимость в систематизации этого разнородного набора задач и методов их решения. Среди интересных попыток достижения этой цели можно отметить вышедшие в разные годы работы [7—14, 162, 163, 165]. Однако эти книги не охватывают всех методов, используемых в настоящее время для решения задач распознавания. Но даже если бы это было не так, еще одну попытку систематизации методов распознавания, построенной на некоторых новых посылах, вероятно, не следовало бы считать бессмысленной. Как отмечает Р. Фейнман [15], считать, что мы понимаем некоторое явление, можно лишь тогда, когда мы умеем объяснять его с помощью нескольких моделей. Кроме попытки изложения «еще одной модели», объясняющей природу и смысл проблемы распозна-

вания и описания современных методов решения задач распознавания, в этой книге даются примеры решения конкретных задач из различных прикладных областей исследования.

Если материал книги заинтересует специалистов, работающих над методами распознавания, и будет стимулировать их усилия построить общую теорию распознавания образов и если книга окажется полезной тем, кто решает прикладные задачи распознавания, то автор будет считать свою цель достигнутой.

На разных этапах работы над этой книгой большую помощь оказывали Евреинов Э. В., Косарев Ю. Г., Волошин Г. Я., Лбов Г. С., Елкина В. Н., Величко В. М., Самохвалов К. Ф., Курилов Б. М., Юдина Л. С. и Белак Е. В. Весьма полезными были дискуссии с Ворониным Ю. А., Ковалевским В. А., Турбовичем И. Т., Чистович Л. А., Кожевниковым В. А., Молчановым А. П. и Бондарко Л. В.

Автор считает своим приятным долгом выразить им искреннюю благодарность.

Н. ЗАГОРУЙКО

В этой главе предлагается принцип классификации, в соответствии с которым можно было бы свести все многообразие задач распознавания к задачам нескольких типов и указать алгоритмы, пригодные для решения задач того или иного типа.

В начале необходимо дать некоторые определения.

1. Определения

Под *образом* будем понимать наименование области в пространстве признаков, в которой отображается множество объектов или явлений материального мира, выделенное в соответствии с определенной целью. В зависимости от *цели* мы можем выбрать разные признаковые пространства или разные области в этих признаковых пространствах, в результате чего одна и та же пара объектов может быть отнесена как к одному и тому же, так и к различным образам. *Признаком* в определении названо количественное описание того или иного свойства (параметра) исследуемого предмета или явления. В геометрической интерпретации под *образом* понимается область в n -мерном выборочном пространстве, вдоль координатных осей которого отложены значения параметров. Описание этой области называется *эталоном*, а отдельная точка в ней — *реализацией* образа. Перечень (список наименований) фиксированных областей, на которые разделено (или необходимо разделить) выборочное пространство, называется *алфавитом* объектов распознавания (или алфавитом образов). В любой задаче распознавания число элементов алфавита конечно и не меньше двух [9].

При проектировании распознающего автомата B всегда имеется в виду автомат A , характеристики ко-

торого должны быть с той или иной степенью точности воплощены в автомате B [16]. Часто в роли автомата A (образца для подражания) рассматривается человек. Критерии для оценки правильности выбора алфавита и надежности распознавания автомата B вырабатываются в результате анализа целевого назначения и характеристик автомата A . Характеристики распределения реализаций образов в разных случаях могут быть известны с различной степенью точности. В практических задачах распознавания распределения генеральных совокупностей обычно неизвестны. Исходный материал (*обучающая последовательность*) представлен конечной (часто очень малой) выборкой.

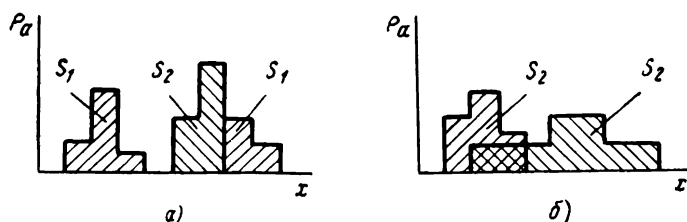


Рис. 1. Детерминистская (а) и вероятностная (б) ситуации.

Если трактовать частоту встречаемости реализаций образа как априорную вероятность (P_a) его появления, то могут иметь место два случая:

1. В любой точке выборочного пространства автомата B с ненулевой априорной вероятностью могут появляться реализации только одного образа (рис. 1, а).

2. Хотя бы в одной точке выборочного пространства автомата B с ненулевой априорной вероятностью могут отражаться реализации, относимые автоматом A к различным образам (рис. 1, б).

Ситуацию, соответствующую первому случаю, назовем *детерминистской*, а второму — *вероятностной*. Автомат B в первом случае имеет принципиальную возможность распознавать обучающую последовательность безошибочно, в то время как во втором случае решения, принимаемые автоматами B и A , будут не всегда одинаковыми.

Что касается самого важного момента — *распознавания контрольной последовательности*, т. е. реализаций,

не совпадающих с реализациями обучающей выборки, то обычно как в первом, так и во втором случае используются одни и те же методы интерпретации и экстраполяции: делаются предположения либо о характере распределения генеральной совокупности, либо о возможной структуре множества обучающих и контрольных реализаций, либо о типе допустимых правил принятия решений. Добавление гипотез такого рода к информации, содержащейся в обучающей выборке, обеспечивает единственность решения задачи распознавания и предопределяет его качество. Эвристический этап выбора гипотезы или «угадывания» закономерностей генеральной совокупности по конечной выборке имеется практически во всех алгоритмах распознавания. Строгой формальной процедуры для его выполнения пока нет. В связи с этим, последующее применение строгих решающих правил, заимствованных, например, из теории статистических решений, не может сделать строгим весь алгоритм распознавания в целом. Поэтому, пока нет оснований для разделения алгоритмов решения задач распознавания образов на «строгие» («статистические») и «эвристические».

Прежде чем перейти к классификации задач распознавания, кажущейся нам оправданной, необходимо рассмотреть структуру распознающих автоматов.

2. Структура распознающих автоматов

Всякое распознающее устройство состоит из блока измерения параметров (рецептора X), блока принятия решения (классификатора D) и блока исполнительных устройств (эффектора S) (рис. 2, *а*). Рецептор измеряет значения параметров X реализации образа. Классификатор в соответствии с определенными правилами D указывает, какой области S выборочного пространства принадлежит эта точка X . Результат решения классификатора воплощается эффектором в определенное действие (нажимается клавиша пишущей машинки, включается двигатель и т. д.).

Такая грубая модель распознающего устройства редко соответствует действительности. Обычно бывает трудно выделить в автомате один рецептор, один классифика-

тор и один эффектор. Чаще всего распознающий автомат представляет собой цепочку более простых автоматов, каждый из которых имеет некоторый эквивалент рецептора, классификатора и эффектора. Так, например, можно представить себе один из вариантов алгоритма распознавания устных слов (рис. 2, б). В процессе звучания слова рецептор измеряет некоторые параметры речевого сигнала. По истечении определенного промежутка времени, соотносимого, например, со слогом, классификатор определяет принадлежность этого участка речи к тому или иному слогу. Подобная процедура повторяется до конца звучания слов. Затем автомат, используя

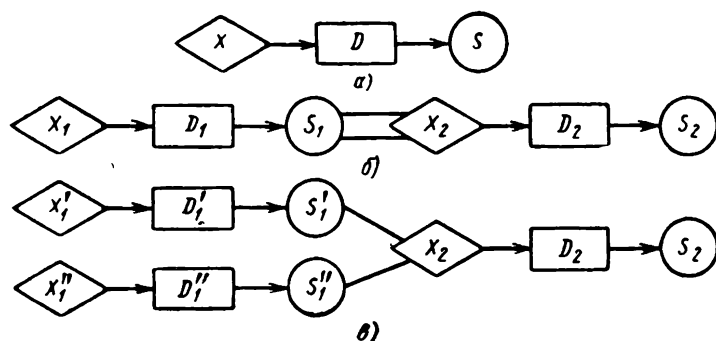


Рис. 2. Блок-схемы распознающих устройств.

промежуточные решения относительно слогов S_1 , определяет, какому из слов, имеющих в алфавите S_2 , принадлежит данная последовательность слогов. Такой алгоритм моделирует два последовательно распознающих автомата; рецептором первого является устройство, измеряющее текущие значения параметров речевого сигнала X_1 . В пространстве этих параметров классификатор D_1 выделяет области S_1 , называемые «слогами». Автомат (его эффектор) фиксирует в виде определенного сигнала решение классификатора D_1 о принадлежности реализации X_1 к одному слогу S_1 . По окончании слова вступает в действие второй распознающий автомат. Его рецептор анализирует сигналы на выходе эффектора первого автомата. Выборочным пространством X_2 второго автомата служит пространство слогов. Часто его называют пространством «вторичных» признаков. Классификатор D_2 в этом пространстве X_2 определяет область S_2 , назы-

ваемую «словом». Эффектор второго автомата сигнализирует о том, какое слово распознано.

Анализ структуры разработанных к настоящему времени распознающих устройств позволяет выделить в них достаточно сложную сеть таких простых автоматов, каждый из которых работает в своем выборочном пространстве, по своим правилам принятия решения и со своим алфавитом (рис. 2, в). В частности, используемые в устройствах нормализаторы (уровня громкости, темпа речи и т. д.) тоже могут рассматриваться в качестве элементарных распознающих автоматов. Анализ таких иерархических распознающих устройств с позиций теории сложных систем дан в гл. 5. Здесь же мы будем пока обсуждать структуру одноступенчатого элементарного автомата (рис. 2, а).

Изучение элементарного распознающего автомата или задачи, которую он решает, обычно сводится к поиску ответа на следующие вопросы:

1. Какие элементы алфавита (S) автомат должен отличать друг от друга (что распознавать).

2. На какие свойства (X) предметов или явлений следует обращать внимание (что измерять).

3. По каким правилам (D) следует принимать решение о принадлежности реализации к тому или иному элементу алфавита (образу).

4. Какие дополнительные ограничения и гипотезы (H) использованы в алгоритме распознавания.

5. Какова стоимость (N) распознающего устройства и потеря из-за ошибок в его работе.

Эти элементы — S , X , D , H и N — в том или ином виде присутствуют в любой задаче распознавания образов.

3. Основные типы задач распознавания образов

Задача, стоящая перед разработчиком распознающего устройства, сводится к минимизации затрат N , складывающихся из затрат двух видов: N_1 — стоимости процедуры и N_2 — стоимости потерь. Под стоимостью процедуры N_1 будем понимать меру, с помощью которой можно оценивать такие элементы затрат, как объем памяти, количество машинных операций, вес, габарит, стоим-

мость оборудования и т. д., необходимые для реализации, т. е. для построения и использования данного распознающего устройства или имитирующего его машинного алгоритма. В N_1 входит стоимость реализации решающего правила $N(D)$, стоимость измерения признаков $N(X)$ и стоимость описания эталонов $N(S)$, т. е.

$$N_1 = N(D) + N(X) + N(S). \quad (1)$$

Стоимость потерь N_2 при заданном виде матрицы S стоимости ошибок распознавания зависит от вероятности этих ошибок R , т. е. $N_2 = N(C, R)$. Обычно имеется в виду, что N не должно превышать некоторую наперед заданную величину N_0^*). Уменьшение затрат N при заданной стоимости процедуры ($N_1 = \text{const}$) может быть достигнуто путем уменьшения N_2 , т. е. уменьшения вероятности ошибок R . В рамках фиксированных затрат на процедуру уменьшить ошибки можно за счет использования более эффективных процедур, т. е. более эффективных решающих правил, более информативных признаков и т. д.

При заданной стоимости потерь ($N_2 = \text{const}$) уменьшение N достигается снижением стоимости процедуры N_1 — опять же за счет использования более эффективных процедур.

Таким образом, в том и другом случае проблема минимизации затрат N сводится к выбору высокоэффективных процедур. Рассмотрение способов выбора процедур и сравнения их эффективности мы начнем с наиболее простых случаев, когда величина затрат N_0 и два из трех процедурных элементов S , X и D заданы и требуется выбрать третий процедурный элемент. Из такой постановки проблемы естественно вытекает *формулировка задач трех основных типов*, словесное описание которых выглядит следующим образом:

I. Задан список образов S , указаны признаки X , в пространстве которых эти образы следует отличать друг от друга. Требуется найти такое решающее правило D , которое при максимальной простоте и дешевизне его реализации в виде машинного алгоритма или специального устройства, обеспечивало бы распознавание элемен-

*) Так называемое требование «практической реализуемости», предложенное Б. М. Куриловым.

тов S в пространстве X с затратами, не превышающими заданной величины N_0 . Этому типу задач посвящено, пожалуй, наибольшее количество работ в области распознавания образов.

II. Заданы элементы S , тип решающих правил D и допустимые затраты N_0 . Требуется найти такую систему признаков X , которая обеспечивала бы извлечение достаточного количества информации при минимальных затратах на измерение этих признаков. Постановка задач этого типа также встречается довольно часто.

III. К третьему типу относятся задачи так называемой «таксономии»: множество реализаций, т. е. векторов пространства X , с помощью решающих функций D (т. е. по определенному критерию «близости», «связанности») нужно разделить на такое количество и таких элементов алфавита S , чтобы при подобной перекодировке затраты не превышали бы заданной величины N_0 .

Таксономические проблемы постоянно возникают в биологии (ботанике и зоологии), в геологии (палеонтологии, минералогии), в социологии, в фонетике и вообще всюду, где устанавливается какая-либо систематика, классификация образов автоматом B в процессе его обучения.

Более строгая постановка основных задач распознавания образов будет сделана в следующих главах при описании алгоритмов решения этих задач. Можно было бы, как в работе [17], выделить еще один тип задач — задачи поиска экономичного способа G оценки или прогнозирования затрат N . Однако эта задача обычно является вспомогательной и носит разовый характер, поэтому, органичимся кратким описанием методов ее решения в конце гл. 2.

Гораздо больше разновидностей задач может появиться, если рассмотреть многоуровневые иерархические системы распознавания (см. гл. 5).

Проведем анализ трехступенчатого устройства распознавания устных команд. Заданными являются: исходное описание речевых сигналов X_1 , список слов S_3 и затраты, связанные с их распознаванием N_3 . Ограничений на вид решающих функций D и на промежуточные перекодировки не накладывается. Тогда может получиться следующая структура распознающего устройства: в пространстве X_1 с помощью решающих функций D_1 распознаются короткие участки речевого сигнала (назовем их условно «фонемами»). Множество та-

ких отрезков составляет алфавит S_1 . В этом пространстве «вторичных» признаков $S_1 = X_2$ с помощью правила D_2 выделяются элементы S_2 большей длины («слоги»). Затем в пространстве слогов $S_2 = X_2$ по правилам D_3 распознаются слова S_3 .

Здесь для элементарного автомата может иметь место или задача одного из описанных выше трех основных типов, или задача, в которой задан один из трех элементов S_i , X_i , D_i , либо даже ни один из них неизвестен. Так, например, можно предоставить машине «свободу выбора» всех элементов второго уровня S_2 , X_2 , D_2 , указав лишь пределы их возможных значений. Критерием правильности решения этой задачи служит оценка сложности (стоимости) N распознающего устройства, в целом удовлетворяющего условиям X_1 , S_3 , D_3 .

Как будет показано в гл. 5 и 6, решение таких задач «комбинированного типа» сводится к последовательному перебору разных вариантов, причем в рамках одного элементарного автомата в каждый момент времени решается задача одного из трех описанных выше типов.

Тот факт, что любую из известных задач распознавания можно трактовать в терминах задач основных типов и вместе с тем, ни одной задачи нельзя анализировать без использования всех этих элементов: S , D и X — служит косвенным, но достаточно убедительным доказательством того, что предлагаемая классификация задач распознавания образов является минимальной и полной.

В литературе описаны другие виды классификации задач распознавания и методов их решения.

Например, Г. С. Себастьян [8] делит все методы решения задач распознавания на три группы по степени сложности, общности правила принятия решения и способа измерения признаков. К первой группе он относит специфические методы решения конкретных задач, которые нельзя распространить ни на какие другие конкретные задачи. Признаки выбираются и измеряются человеком вручную, а решение принимает машина по простому совпадению реализаций с эталоном. В методах второй группы измерение признаков и принятие решений осуществляется ЭВМ, моделирующей некую гипотетическую модель функций человека. Несовершенство этой модели часто приводит к большим ошибкам распознавания. Третью группу представляют собой методы, в которых задачи поиска информативного описания и построения решающей функции решаются на ЭВМ одновременно, в некотором смысле оптимально и в наиболее общем виде.

По нашей терминологии это классификация комбинированных задач только I—II типов. Помимо того, что при этом не рассматриваются задачи таксономии, границы такой классификации расплывчаты. Неясно также, к какой группе отнести метод, использующий, например, простой способ измерения признаков и сложные решающие функции или наоборот.