

Р. Дуда

Распознавание образов и анализ сцен

**Москва
«Книга по Требованию»**

УДК 004.42
ББК 32.811
Р11

Р11 **Р. Дуда**
Распознавание образов и анализ сцен / Р. Дуда – М.: Книга по Требованию, 2021. – 508 с.

ISBN 978-5-458-42642-8

Тематика книги связана с исследованиями по созданию интегральных роботов, способных целенаправленным действиям в сложных условиях. В последние годы такие исследования интенсивно проводятся в разных странах. В книге полно и систематически изложены методы распознавания образов и дан анализ пространственных сцен по их плоскому изображению. Существующие в области распознавания образов методы авторы рассматривают с новых, нестандартных точек зрения. В конце каждой главы приводится список задач. Книга доступна широкому кругу читателей, интересующихся или работающих над созданием искусственного интеллекта. Она может служить хорошим вводным курсом для студентов, специализирующихся в этой области.

ISBN 978-5-458-42642-8

© Издание на русском языке, оформление
«YOYO Media», 2021
© Издание на русском языке, оцифровка,
«Книга по Требованию», 2021

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.

ПРЕДИСЛОВИЕ

При написании данной книги мы ставили своей целью дать систематическое изложение важнейших разделов распознавания образов — области науки, связанной с машинным распознаванием тех или иных закономерностей при наличии шума или в сложных условиях. Стимулированное вычислительной техникой распознавание образов вступило в пору расцвета в начале 60-х годов и более десяти лет энергично развивалось. Такому росту содействовало привлечение многих других дисциплин, таких, как математическая статистика, теория связи, теория коммутационных схем, теория управления, исследование операций, биология, психология, лингвистика и вычислительные науки. Читатели, знакомые с литературой, оценят, насколько все они оживили данную область.

Такая широта создает также и значительные трудности при написании книги. Единой теории распознавания образов, включающей все главные разделы, нет, ибо каждой области применения свойственны одной ей присущие особенности, в расчете на которые и строится соответствующий подход. Наиболее развитым разделом, не связанным областью применения, является теория классификации, составляющая предмет первой части книги. Для классификации образов, представленных абстрактно в виде векторов, предлагаются формальные математические процедуры, основанные на статистической теории принятия решений.

Попытки отыскать универсальные, независимые от области приложения процедуры для построения этих векторных представлений не дали пока общих результатов. Наоборот, каждая конкретная область задач характеризуется набором процедур, соответствующих ее специфическим свойствам. Среди различных областей, представляющих интерес, наибольшее внимание до настоящего времени привлекала область изображений. Кроме того, работы по этой тематике развивались от классификации изображений к их анализу и описанию. Систематическому изложению этих вопросов в связи с анализом зрительных сцен посвящена часть II данной книги.

Так как теории и методы распознавания образов по своему характеру являются математическими дисциплинами, то необходимо остановиться на уровне математической строгости, принятом в нашем изложении материала. Откровенно говоря, он невысокий. Нас гораздо больше беспокоила возможность разобраться и понять, нежели построить строгое математическое обоснование задачи.

Это нашло отражение в множестве поясняющих примеров, правдоподобных доводах и обсуждениях поведения решений. Наравне с этим мы избегали пользоваться теорией меры и попытались обойтись без таких тонкостей, как сходимости последовательностей случайных величин, обоснование применения дельта-функций и возможность патологических случаев. Мы все же предполагаем наличие у читателя общего представления об основных разделах прикладной математики, включая теорию вероятностей и линейную алгебру; для понимания гл. 8 также окажется полезным знакомство с преобразованиями Фурье. Требуемая математическая подготовка должна соответствовать уровню аспиранта первого года обучения по специальностям вычислительная наука, электротехника и статистика.

В связи с тем что распознавание образов представляется весьма специальной областью, возможно, имеет смысл особо отметить методологическую гибкость курса по этому предмету. Многое заимствовав из различных разделов математики, а также других ранее упомянутых дисциплин, распознавание образов служит почти идеальным средством представления различных вопросов в единых рамках. Тем студентам, для которых распознавание образов как таковое не представляет самостоятельного интереса, все же целесообразно овладеть знаниями и развить мастерство, которое сослужит им хорошую службу при других обстоятельствах.

Курсы лекций для аспирантов, основанные на материале данной книги, читались нами в Калифорнийском университете в Беркли и в Станфордском университете. Каждая из двух частей этой книги может быть сравнительно неплохо изучена в течение полусеместра и достаточно глубоко — за семестр. При более сжатых сроках мы рекомендуем работать с отдельными разделами из большинства глав: одни главы прорабатывая основательно, а другие — нет. Как и всегда, интересы преподавателя должны определять окончательный выбор материала.

Мы полагали также, что эта книга окажется полезной и для специалистов-прикладников, и с этой целью попытались сделать материал более доступным. Где возможно, мы старались пользоваться общепринятыми обозначениями и включили в книгу обширный указатель. Каждую главу завершают библиографические и исторические замечания и полезный, как мы надеемся, список литературы. Хотя большие списки литературы и создают впечатление завершенности книги, на самом деле публикации появляются столь интенсивно, что полноты достичь нельзя, и мы не притязали на это.

Весьма полезным при работе над данной книгой оказалось сотрудничество с многими организациями и отдельными лицами. Прежде всего мы хотим поблагодарить Отдел информационных систем Научно-исследовательского управления военно-морских сил за финансирование по контракту № 0014-68-C-0266. Факультеты

электротехники и вычислительных наук в Беркли и Станфорде предоставили возможность проверить материал на учебной аудитории. В центре по искусственному интеллекту при Станфордском научно-исследовательском институте по инициативе руководителей д-ра Ч. А. Розена и д-ра Б. Рафаэля нам были созданы идеальные условия для работы. Хотя мы не в состоянии перечислить всех, кто помог нам своими замечаниями, особо хотелось бы поблагодарить д-ра Н. Дж. Нильсона за его многочисленные предложения по улучшению рукописи. Считаем своим долгом поблагодарить д-ра Т. О. Бинфорда, д-ра Т. М. Ковера, К. Л. Феннема, д-ра Г. Ф. Гронера, Д. Дж. Холла, д-ра М. Е. Хелмана, д-ра М. А. Касслера и д-ра Дж. Х. Мансона за их ценные замечания. Кроме того, нам хочется поблагодарить д-ра Р. К. Синглтона за его участие в работе над примерами гл. 8. И наконец, с удовольствием выражаем признательность неутомимой К. Л. Спенс за перепечатку различных вариантов рукописи.

*Р. О. Дуда,
П. Е. Харт*

Часть I

КЛАССИФИКАЦИЯ ОБРАЗОВ

Глава 1

ВВЕДЕНИЕ

1.1. МАШИННОЕ ВОСПРИЯТИЕ

Стремление расширить область применения цифровых вычислительных машин существует со времени их появления. В известной мере это связано с требованиями практики, с поиском наиболее эффективных способов деятельности. Отчасти это вызвано также вполне понятным стремлением к усовершенствованию конструкции или способов программирования с целью придать машинам новые, ранее недоступные им функции. Так или иначе обе эти причины имеют отношение к определенному разделу науки об искусственном интеллекте, который мы будем называть машинным восприятием.

Способность машины воспринимать окружающий мир в настоящее время крайне ограничена. Для преобразования света, звука, температуры и т. п. в электрические сигналы созданы разнообразные датчики. Если окружающая среда контролируется достаточно четко, а сигналы по своему смыслу просты, что имеет, например, место при применении обычных устройств ввода в ЭВМ, то задача восприятия оказывается несложной. Но, как только вместо считывания перфокарт или магнитных лент от машины требуется чтение рукописного текста или расшифровка биомедицинских фотографий, вместо задач ввода данных приходится иметь дело с гораздо более сложными задачами их интерпретации.

Та видимая легкость, с которой животные и даже насекомые справляются с задачами восприятия, одновременно и ободряет, и обескураживает. Психологические и физиологические исследования дали ряд интересных результатов, касающихся процессов восприятия животными. Тем не менее этого пока недостаточно для воспроизведения процессов восприятия с помощью ЭВМ. Особую привлекательность этому вопросу придает и то обстоятельство, что восприятие есть нечто известное по опыту каждому, а вместе с тем на деле никем не понятное. Бесполезными оказываются и попытки

исследования сущности восприятия посредством самоанализа из-за того, что, по-видимому, большинство обычных процессов восприятия протекает подсознательно. Парадоксально, что все мы хорошо владеем восприятием, но никто из нас не знает о нем достаточно.

Неполнота теории восприятия не помешала человеку попытаться решить менее сложные задачи. Ряд таких задач связан с классификацией образов — отнесением материальных объектов или явлений к одному из нескольких predetermined классов.

В результате широкого исследования вопросов классификации была получена абстрактная математическая модель, составившая теоретическую основу для разработки устройства-классификатора. Естественно, что для каждой отдельной области применения требуется в конечном счете собрать воедино специфические характеристики конкретной задачи. Наибольшее внимание среди множества таких областей привлекли задачи, относящиеся к изображениям. Цель данной книги — систематически изложить основные принципы теории классификации образов и тех методов анализа сцен, которые представляются наиболее широко применимыми и интересными.

1.2. ПРИМЕР

Чтобы пояснить характер задач, с которыми в дальнейшем придется иметь дело, рассмотрим следующий, несколько надуманный пример.

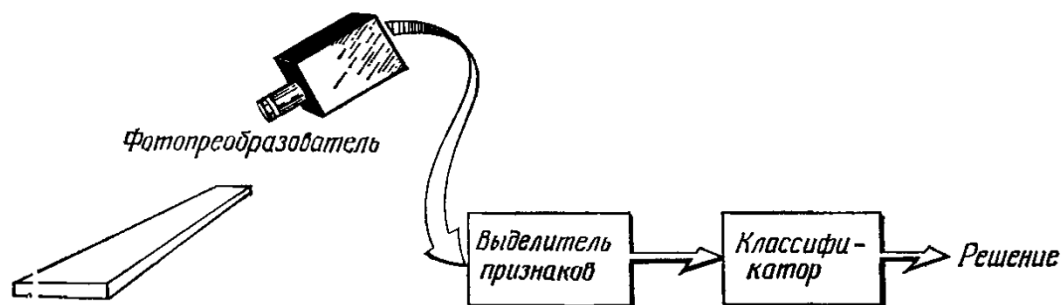


Рис. 1.1. Схема системы классификации.

Допустим, что на некоторой деревообделочной фабрике, выпускающей изделия из древесины различных сортов, требуется автоматизировать процесс сортировки обрабатываемой древесины по породе дерева. Первоначально решено попытаться разделять древесину пород березы и ясеня по ее внешнему виду. Система, предназначенная для выполнения этой задачи, может быть построена по схеме, изображенной на рис. 1.1. Полученное в камере изображение древесины передается на *выделитель признаков*, назначение которого состоит в уменьшении объема данных измерением конкретных «признаков» или «свойств», отличающих вид древесины березы и ясеня. Далее эти признаки (точнее, измеренные значения этих

признаков) подаются на *классификатор*, предназначенный для оценки представленных данных и принятия окончательного решения относительно типа древесины.

Рассмотрим, по какому принципу могут быть построены выделитель признаков и классификатор. Допустим, что на фабрике нам сообщили, что окраска березы обычно светлее, нежели ясеня. Таким образом, яркость оттенка древесины становится очевидным признаком, и можно попытаться распознавать древесину, наблюдая попросту, превышает или нет ее средняя яркость x некоторое критическое значение x_0 . Для выбора величины x_0 можно взять несколько кусков древесины той и другой породы, измерить их яркость и

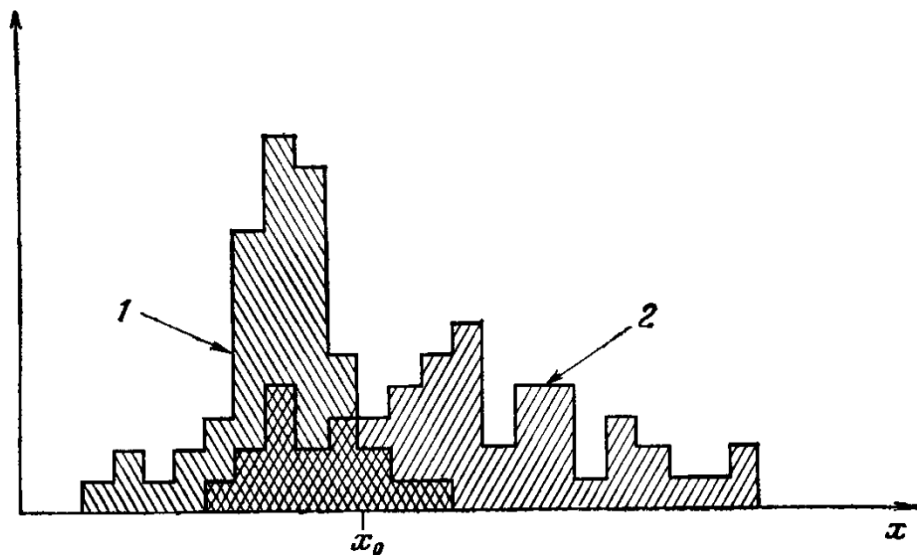


Рис. 1.2. Гистограмма яркости (по оси ординат — число кусков древесины; 1 — яшень, 2 — береза).

оценить результаты измерений. Допустим, что все это проделано и построены гистограммы, приведенные на рис. 1.2. Рассматривая гистограммы, можно установить, что береза обычно светлее ясеня, но вместе с тем очевидно, что один этот критерий не достаточно надежен. Вне зависимости от выбора x_0 отличить березу от ясеня только по яркости невозможно.

При поиске других признаков можно попытаться воспользоваться тем соображением, что волокна ясеня обычно более заметны. Хотя измерить этот признак значительно труднее, чем среднюю яркость, однако можно предположить, что заметность волокон можно оценивать по амплитуде и частоте появлений на изображении переходов от темного к светлому. Таким образом, теперь мы имеем два признака для распознавания древесины — яркость x_1 и заметность волокон x_2 . Выделитель признаков сводит, таким образом, каждую из картинок к точке или *вектору признаков* x двумерного простран-

ства, где

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Наша задача заключается теперь в разделении пространства признаков на две области, все точки одной из которых соответствуют березе, а другой — ясеню. Предположим, что мы определили значения векторов признаков для наших кусков древесины и получили разброс точек, показанный на рис. 1.3. Анализ графика

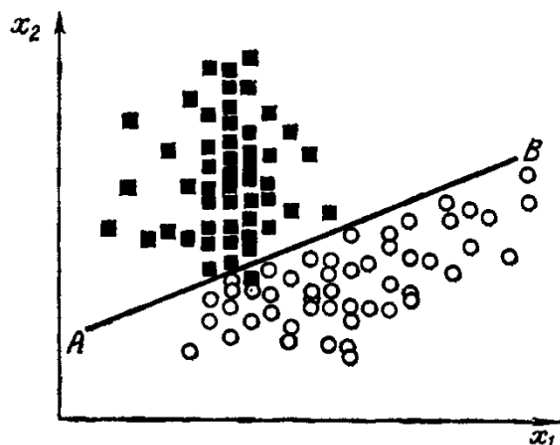


Рис. 1.3. Диаграмма разброса векторов признаков (■ — ясень, ○ — береза).

приводит к следующему правилу распознавания полученных данных: считать, что данная древесина есть ясень, если вектор признаков попадает выше прямой АВ; в противном случае считать, что это береза.

Несмотря на то что разделение взятых кусков посредством данного правила достаточно хорошее, нет никакой гарантии, что так же хорошо оно будет и по отношению к новым кускам. Безусловно, разумнее было бы взять еще несколько кусков и

посмотреть, какая часть из них будет определена правильно. Таким образом, мы приходим к заключению, что поставленная задача включает в себя и элементы статистики, в связи с чем может потребоваться поиск такой процедуры классификации, которая сделает вероятность ошибки минимальной.

Говоря об этом, следует помнить, что в качестве примера была выбрана простая задача. Значительно реальнее задача сортировки древесины на много различных классов. В частности, чтобы отличить дуб от березы и ясеня, вполне могут потребоваться и такие менее очевидные признаки, как «ровность волокон». При большем числе классов и признаков, возможно, придется отказываться от графического способа представления классификатора. Поэтому для дальнейшего изложения вопроса потребуется развить теоретические основы метода.

1.3. МОДЕЛЬ КЛАССИФИКАЦИИ

Рассмотренный пример содержит многие элементы наиболее распространенной абстрактной модели распознавания образов — модели классификации. Модель эта состоит из трех частей — датчика, выделителя признаков и классификатора. Датчик воспринимает воздействие и преобразует его к виду, удобному для машинной обработки. Выделитель признаков (называемый также рецептором,

фильтром свойств, детектором признаков или препроцессором) выделяет из входных данных предположительно относящуюся к делу информацию. Классификатор на основе этой информации относит эти данные к одной из нескольких категорий.

Обсуждение принципа действия или конструкции датчика не входит в задачи данной книги. Выделение признаков и классификация, напротив, представляет для нас интерес. С точки зрения теории провести черту между этими двумя действиями можно весьма условно. Идеальный выделитель признаков предельно упрощает работу классификатора, тогда как при наличии всемогущего классификатора не требуется помощь выделителя признаков. Различие это хотя и весьма важно, но делается оно, исходя скорее из практических, нежели теоретических соображений.

Вообще говоря, задача выделения признаков более специализирована по сравнению с задачей классификации. Хороший выделитель признаков, предназначенный для сортировки древесины, принесет, по-видимому, малую пользу при сравнении отпечатков пальцев или классификации клеток крови. Однако большое число технических приемов было развито в связи с задачей извлечения полезной информации из изображения. Часть II данной книги как раз посвящена описанию этих методов и их свойств.

Задача классификации по существу представляет собой задачу разбиения пространства признаков на области, по одной для каждого класса. Разбиение это в принципе надо производить так, чтобы не было ошибочных решений. Если этого сделать нельзя, то желательно уменьшить вероятность ошибки, или если ошибки имеют различную цену, то сделать минимальной среднюю цену ошибки. При этом задача классификации превращается в задачу статистической теории принятия решений, широко применяемую в различных областях теории распознавания образов. Этому и ряду других вопросов, относящихся к теории классификации, посвящена часть I данной книги.

1.4. ОПИСАТЕЛЬНЫЙ ПОДХОД

В области распознавания образов и машинного восприятия существует много задач, для которых модели классификации совершенно не пригодны. Например, при анализе фотографии следов частиц в пузырьковой камере требуется не просто классификация, а описание картины. Такое описание должно содержать информацию как об отдельных частях картины, так и о связи между ними. В принципе оно должно непосредственно отражать структуру, присущую исходной сцене.

Рассмотрим, например, простую сцену, показанную на рис. 1.4. Возможно, что для некоторых целей окажется вполне удовлетворительным простейшее описание вроде «на письменном столе» или

«стоит телефон». Более полный анализ должен включать перечень всех основных имеющихся объектов — телефона, бумаги, чашки, карандашей, ластика и т. д. При еще более полном анализе будут выявлены отношения между этими объектами, что может вылиться в выражение «(Два карандаша на пачке бумаги) перед (чашка слева от (ластик перед телефоном))».

Задача анализа видимой сцены и получения ее структурного описания оказывается очень трудной. Приемлемая для всех форма-

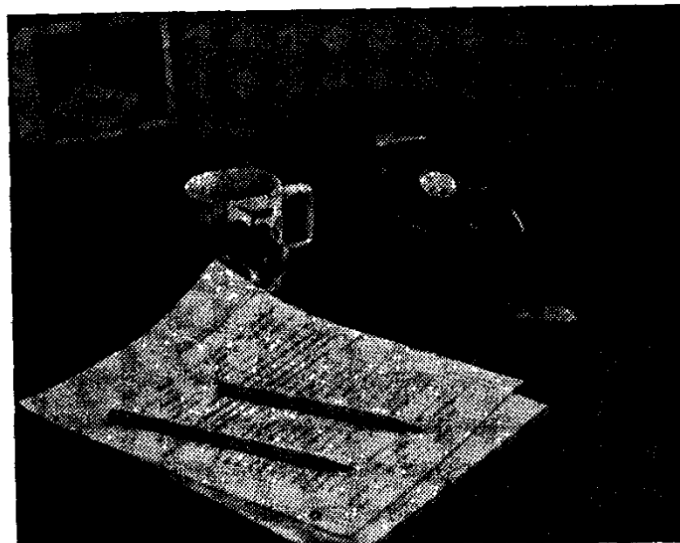


Рис. 1.4. Простая сцена.

лизация этой проблемы, аналогичная модели классификации, едва лишь выявляется из работ, выполненных к настоящему моменту. Были предприняты попытки заимствовать понятия из теории формальных языков и построить лингвистическую модель для анализа сцен. При этом сцена рассматривается как выражение на языке, грамматика которого определяется допустимыми структурными отношениями. Исходя из этой формулировки, анализ сцен можно рассматривать как процесс грамматического разбора, в результате которого описание всей сцены получается в виде композиции взаимосвязанных подсцен.

Методы использования структурных отношений между элементами картины с целью ее анализа и создания полезного описания не исчерпываются лингвистической моделью. Хотя общей идейной основы, объединяющей все эти методы, еще не создано, тем не менее был разработан ряд интересных процедур для конкретных задач, которые носят эвристический характер. Эти процедуры сами по себе можно подвергнуть рассмотрению, и изучением этого общего вопроса описательного подхода к анализу сцен заканчивается часть II.