АКАДЕМИЯ НАУК СССР

Отделение информатики,
вычислительной техники и автоматизации

Распознавание
Классификация
Прогноз

МАТЕМАТИЧЕСКИЕ МЕТОДЫ И ИХ ПРИМЕНЕНИЕ

Выпуск 2

Ежегодник основан в 1988 г.

Ответственный редактор
член-корреспондент АН СССР
Ю. И. ЖУРАВЛЕВ

Москва «Наука», 1989

I

ТЕОРИЯ РАСПОЗНАВАНИЯ

УДК 007+519.712.2

РАСПОЗНАВАНИЕ ОБРАЗОВ И РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ

Ю. И. ЖУРАВЛЕВ, И. Б. ГУРЕВИЧ

ВВЕДЕНИЕ

Одна из ключевых проблем информатики — разработка, исследование и реализация методов синтеза при помощи обучения алгоритмических процедур преобразования и анализа информации, предназначенных для решения таких информационных задач, для которых соответствующие алгоритмы неизвестны. Эти методы уже свыше 50 лет неявно или явно составляют сердцевину математической теории алгоритмов, кибернетики и ныне информатики. Задачи, требующие использования таких методов, возникают в связи с обработкой и преобразованием на ЭВМ структур, образованных из символов, т. е. структур, представляющих в программах искусственного интеллекта знания о проблемной области в целом и знания, относящиеся к конкретной задаче. Несмотря на подобную универсальность задач, однако, эти методы стали предметом интенсивных исследований, развития и в конце концов получили оформление в виде законченной математической теории лишь в рамках одного, хотя и весьма обширного, класса задач преобразования и анализа информации — задач, достаточно давно известных под названием (вероятно, не очень точным и удачным) задач распознавания образов.

Основное содержание данного обзора составляет общая характеристика проблемы распознавания (гл. 1), изложение современного состояния и важнейших результатов математической теории распознавания (главным образом для случая стандартной исходной информации) (гл. 2) и анализ проблемы распознавания изображений (в том числе концепции дескриптивного подхода к анализу изображений) (гл. 3) [1]

Глава 1
ПРОБЛЕМА РАСПОЗНАВАНИЯ

1.1. Основные понятия

В силу чисто исторических причин (термин «pattern recognition», переведенный как «распознавание образов», был заимствован, как и многие другие термины информатики, из англоязычных работ) этот класс задач оказался связан с понятием «образа» («pattern»). К сожалению, в свое время не обратили внимания на его многозначность — термин «pattern», кроме значения «образ», имеет еще и значение «модель», «стиль», «режим», «закономерность», «образ действий». В современном распознавании и особенно искусственном интеллекте его употребляют в самом широком смысле, имея в виду некоторое структурированное приближенное (частичное) описание (эскиз) изучаемого объекта или явления, причем частичная определенность описания является принципиальным свойством образа. Образ допускает рекурсивное определение: символ является образом, список символов является образом, образами являются только те выражения, которые построены в соответствии с двумя указанными условиями. Списочная запись позволяет использовать одно и то же представление для описания образа произвольного типа независимо от его «содержания». Дополнительным достоинством записи этого типа служит возможность пользоваться одними и теми же алгоритмами для работы с образами с различными денотатами. Естественно также допускать, что образ состоит из двух групп символов, представляющих соответственно переменные и постоянные характеристики объекта описания.

Основное назначение описаний (образов) — это их использование в процессе установления соответствия объектов, т. е. при доказательстве их идентичности, аналогичности, подобия, сходства и т. п., осуществляемом путем сравнения (сопоставления). Два образа считаются подобными, если удается установить их соответствие. Можно, в частности, считать, что имеет место соответствие двух образов, если можно достичь их идентичности, подставляя вместо переменных какие-либо выражения.

Сопоставление образов представляет собой основную задачу распознавания и играет существенную роль в информатике в целом. Эта задача возникает, в частности, в различных разделах искусственного интеллекта, например в понимании естественного языка, символьной обработке алгебраических выражений, эксперт-

[1] В обзоре используются материалы из работ авторов, в частности [11, 20, 21].

ных системах, преобразовании и синтезе программ ЭВМ. Процедура сопоставления оказалась столь существенной для искусственного интеллекта, что во многие языки программирования, используемые в искусственном интеллекте, она входит в качестве примитива.

Отметим, что в различных задачах понятию образа придается различный смысл. Так, скажем, в распознавании (в классических моделях) образ обычно описывается вектором признаков, каждый элемент которого представляет числовое значение одного из признаков, характеризующих соответствующий объект. В структурной модели распознавания в качестве образа выступает некоторое высказывание, порождаемое той грамматикой, которая характеризует класс, которому данный образ принадлежит. В задачах обработки текста роль образа исполняет некоторая цепочка — в результате процедура установления соответствий сводится к поиску вхождений этой цепочки (образа) в тексте.

Термин «распознавание» в равной мере относится как к процессам восприятия и познания, свойственным человеку и живым организмам в целом, так и к попыткам реализовать и использовать «механические» анологи (по функции и результату) этих процессов, исследование и синтез которых составляют предмет распознавания как раздела информатики. В данной главе будут рассматриваться лишь последние.

Итак, целью создания автоматизированных вычислительных систем распознавания является автоматизация группы процессов восприятия и познания, связанных с поиском, выделением, идентификацией, классификацией и описанием образов на основе анализа реальных данных, полученных тем или иным способом. Обычно поиск и выделение образов осуществляются на начальном этапе анализа в процессе обработки исходных данных и выполняются для того, чтобы получить некоторые промежуточные результаты (т. е. преобразовать исходные данные в некоторую другую форму), «лучше» представляющие образы с точки зрения решения соответствующей задачи. Следующий этап — разработка «классификатора» — обычно включает анализ выборочных (преобразованных) данных, синтез модели, учитывающей изменчивость образов, принадлежащих некоторому классу, выбору из заданного набора характеристик некоторого их подмножества, адекватно характеризующего отдельные классы объектов, определение методов выделения указанного подмножества и разработку собственно алгоритма распознавания (классификации).

1.2. Гносеологические аспекты распознавания

Параллельно развитию все более тонких методов распознавания росла потребность в создании некой регулярной основы и стандартизированных механизмов для сопоставления отдельных эвристических алгоритмов распознавания по вычислительной сложности, эффективности, точности и быстродействию, для выбора оптимального алгоритма в модели и в конечном счете, автоматизации выбора и синтеза алгоритмических процедур для решения конкретной задачи распознавания. В результате эти попытки, не будучи подкреплены надежным фундаментом математической теории распознавания, приводили лишь к спонтанным достижениям частного характера и способствовали осознанию принципиальной ограниченности возможностей самых изощренных эвристических моделей распознающих алгоритмов при их «независимом» применении и необходимости установления связей между отдельными моделями, т. е. разработки общей теории распознавания. Это развитие, естественно, усилило интерес к гносеологическому статусу таких фундаментальных понятий распознавания, как образ, класс, распознающий алгоритм. К этим проблемам обращались и философы, интересовавшиеся гносеологическими аспектами кибернетики и информатики (см., например, работы В. С. Тюхтина [31]), и математики, занимавшиеся разработкой теории распознавания. Поскольку в Гл. 2 достаточно детально изложены важнейшие принципы современной математической теории распознавания, основанной на так называемом алгебраическом подходе к задачам распознавания и классификации [20], здесь мы остановимся лишь на идее комбинаторной регулярности У. Гренандера [7] и парадигматическом символе С. Ватанабэ [36], оказавших существенное влияние на становление и развитие теории распознавания.

При построении своей теории У. Гренандер исходил из того, что поиск регулярности — это доминирующая тема в попытках человека понять окружающий мир, а любая попытка такого рода базируется на неявном или явном допущении о подчиненности явлений природы и событий искусственного мира, созданного человеком, определенным законам, определяющим упорядоченность и структуру. Он опирался также на следующий тезис Д. Юма, лежащий в основе рассуждения методом неполной индукции: «Если бы мы руководствовались разумом, то следовали бы принципу, что те события, опыт столкновения с которыми у нас отсутствует, должны иметь сходство с такими событиями, опытом столкновения с которыми мы располагаем, и что закономерности в природе всегда остаются неизменными» [33]. В основе теории Гренандера лежит представление о структурированности мира, т. е. существовании достаточно определенной регулярности, проявляющейся в виде постоянных связей, закономерностей. Отправной точкой при построении теории служит объект — образ как таковой — и возникающие в этой связи проблемы природы образа, прототипа, класса. Образы рассматриваются в рамках точного формализма, который используется в качестве основы для синтеза и анализа образов, что способствует пониманию того, каким способом образы строятся и обрабатываются. В результате процедуры, обеспечивающие описание, аппроксимацию, восстановление и распознавание образов, принимают вид естественных следствий процедур формирования и преобразования объектов.

Как известно, обнаружение или «изобретение» принципов описания регулярностей, а также их логический анализ и получение соответствующих следствий реализуются посредством так называемых формальных систем, под которыми понимается ряд базисных утверждений, процедур и правил, указывающих, каким образом их применять для объяснения соответствующих явлений. Формальная система, описывающая регулярность, должна обладать определенным постоянством относительно времени и пространства. Если подобная система применима лишь к какому-то определенному времени и какому-то определенному месту, то мы имеем дело не с законом природы, а с данными, результатами изолированных наблюдений, поскольку, как правило, предполагается, что законы, порядок, образы — это нечто большее, чем просто изолированные факты. Законы имеют дело с несколькими альтернативами, интересные законы — со значительным числом альтернатив. Поэтому, по У. Гренандеру, образ следует соотносить с некоторым ансамблем возможных случаев, а порядок в таком ансамбле рассматривается как единообразное выполнение определенных свойств.

Основным объектом теории образов У. Гренандера служат комбинаторные регулярные структуры — регулярные конфигурации — логические конструкции, позволяющие определять различные типы регулярности. С формальной точки зрения речь идет о построении новых объектов посредством комбинирования заданных объектов в соответствии с определенными правилами построения объектов.

Постулируется, что образы формируются из простых стандартных элементов — образующих. Это неделимые элементы (атомы), которые выбираются в соответствии с «физической» природой изучаемых объектов или явлений. В качестве таких элементов могут выступать абстрактные символы, множества, отношения или функции. Эти образующие могут представляться абсолютно разными, но их роль в порождении регулярных структур идентична.

С. Ватанабэ [36] рассматривает основные понятия распознавания, используя концепцию «парадигматического символа» (от греческого «paradeigma» — пример, образец) и опираясь на замечание Л. Витгенштейна «видеть нечто$_1$ как нечто$_2$». Образ для него — противоположность хаосу, некая сущность, нечетко определенная, но допускающая приписывание ей имени, т. е. нечто. В этом смысле образ, по Ватанабэ, соответствует «нечто$_2$», причем последнее не существует как таковое в языке на том же уровне, как «нечто$_1$» присутствует на изображении. В некоторых случаях «нечто$_2$» имеет имя, в некоторых — не имеет его, в некоторых — структура безымянной конструкции должна описываться перечислением составляющих ее элементов и указанием хорошо определенного способа их соединения между собой. Заметим, кстати, что, как свидетельствуют многочисленные данные психологии восприятия и психиатрии (зрительные иллюзии, двухзначные изображения типа «утка/кролик» Л. Витгенштейна, гештальтистские данные по переключениям «фигура/фон», пятна Роршаха), существенным оказывается то обстоятельство, что «нечто$_2$» порождается посредством взаимодействия внешнего раздражителя и процессов мышления, обеспечивающих установление связи рассматриваемого объекта с каким-то другим подобным ему или, по крайней мере, родственным.

Замена в формуле Л. Витгенштейна глагола «видеть» на «распознавать» и «нечто$_2$» на «образ» приводит к идее класса, объединяющего ряд индивидуальных объектов, и, естественно, представлению о распознавании как об идентификации некоторого объекта в качестве элемента некоторого известного нам множества, т. е. как о процессе отображения, ставящего в соответствие различным элементам одного множества один и тот же элемент другого.

Распознавание обычно связывают с двумя функциями: отнесением некоторого объекта к классу объектов, неизвестному классификатору объектов, и идентификации некоторого объекта в качестве элемента известного классификатору класса. Первая функция представляет собой процесс выделения новых классов — так называемой кластеризации, а вторая — это собственно распознавание (распознавание в узком смысле или распознавание на парадигму, по С. Ватанабэ). Последнее толкование распознавания опирается на представление об образе как о некотором объекте, выступающем в качестве выборочного образца объектов, составляющих некоторый класс. При этом предполагается, что некоторая группа объектов описывается («представляется», характеризуется) типичным примером, образчиком, прототипом. С другой стороны, этимология понятия «образ» допускает и толкование, не требующее существования какого-то «главного объекта», являющегося родоначальником всех остальных разновидностей «образа». Эта двойственность отражается в эпицентрическом и агрегатном подходе к определению образа и оказывает очень существенное влияние на методологию распознавания в целом.

1.3. Цели распознавания. *Характеризация и типы задач распознавания*

Центральная задача распознавания образов — построение на основе систематических теоретических и экспериментальных исследований эффективных вычислительных средств для отнесения формализованных описаний ситуаций и объектов к соответствующим классам. В основе такого отнесения (распознавания, классификации) лежит получение некоторой агрегированной оценки ситуации, исходя из ее описания. При условии установления соответствия между классами эквивалентности, заданными на множестве решений и множестве объектов распознавания (ситуаций), автоматизация процедур распознавания становится элементом автоматизации процессов принятия решений.

Задачи распознавания представляют собой, по существу, дискретные аналоги задач поиска оптимальных решений. К ним относится широкий класс задач, в которых по некоторой, обычно весьма разнородной, быть может, неполной, нечеткой, искаженной и косвенной информации требуется установить, обладают ли изучаемые (весьма сложные, в некотором смысле «комплексные») ситуации (объекты, явления) фиксированным конечным набором свойств, позволяющим отнести их к определенному классу,— задачи распознавания и классификации, или по аналогичного рода информации о конечном множестве достаточно однотипных процессов следует выяснить, в какой области из конечного числа областей будут находиться эти процессы через определенный период времени — задачи прогнозирования. К задачам этого вида сводятся задачи технической и медицинской диагностики, геологического прогнозирования (в частности, восстановление геофизических полей), прогнозирования свойств химических соединений, сплавов и новых материалов, распознавания и характеризации свойств динамических и статических объектов в сложной фоновой обстановке и при наличии активных и пассивных помех по изображениям, получаемым с помощью разнообразных технических средств, прогнозирования хода строительства крупных объектов, обработки данных дистанционного исследования природных ресурсов, прогнозирования урожая, обнаружения лесных пожаров, управления производственными процессами (прогнозирование возможностей входа значений параметров быстротекущих процессов в критические области) и др.

1.3.1. Общая характеристика задач распознавания. Практические задачи, для решения которых целесообразно применять методы распознавания, отличаются рядом специфических особенностей.

1. Информационные задачи, решение которых осуществляется посредством применения к доступным исходным данным системы преобразований, включающей в общем случае два основных этапа: а) приведение исходных данных к некоторому стандартному виду, удобному для распознавания — синтез формализованного описания ситуации (объекта) на основе имеющейся разнородной информации (эмпирических данных, результатов измерений, знаний о логических аспектах изучаемых явлений (процессов), сведений о конструкции, назначении и эксплуатационных характеристиках (возможно, предполагаемых) объекта, экспертных данных, имеющейся априорной семантической и синтаксической информации); б) собственно распознавание — преобразование формализованного описания в стандартизированную матрицу ответов, соответствующую выбору в качестве ответа (классификационного решения) одной из некоторого конечного фиксированного набора возможностей (указание принадлежности ситуации (объекта) определенному классу).

2. В задачах имеется возможность вводить понятие некоторого подобия между объектами (ситуациями), точнее, между их

описаниями — формулировать обобщенное понятие близости в качестве основания для зачисления ситуаций (объектов) в один и тот же класс или разные классы.

3. Имеется возможность оперировать определенным набором прецедентов — примеров, классификация которых (в смысле решаемой задачи) известна и которые (в виде стандартных формализованных описаний) могут быть предъявлены алгоритму распознавания для настройки на задачу в процессе обучения.

4. Задачи, для которых трудно строить формальные теории и применять классические математические методы, поскольку в ситуациях, в которых они возникают, имеет место один из двух следующих случаев: а) уровень формализации соответствующей предметной области и/или доступная информация таковы, что не могут составить основу для синтеза математической модели, отвечающей классическим математическим или математическо-физическим канонам и допускающей изучение классическими аналитическими или численными методами; б) математическая модель в принципе может быть построена, однако ее синтез или изучение связаны с такими затратами (сбор необходимой информации, вычислительные ресурсы, время), что они существенно превышают выигрыш, приносимый искомым решением, либо выходят за пределы существующих технических возможностей, либо делают решение задачи просто бессмысленным.

5. В задачах «по определению» существует «плохая» исходная информация, характеризующая сложную в семантическом и структурном отношении ситуацию (объект в некоторой среде) — это ограниченная, неполная (с пропусками), разнородная, косвенная (характеристики внешних проявлений функционирования процесса, причем не всегда относящиеся к принципиальным особенностям лежащего в его основе механизма), нечеткая, неоднозначная, вероятностная. В целом это задачи, в которых известно слишком мало, чтобы можно было пользоваться классическими методами решения (моделями), но все-таки известно достаточно, чтобы решение было возможно.

1.3.2. Типы задач распознавания. 1. Отнесение предъявленного объекта (ситуации) по его формализованному описанию к одному из заданных классов — задача распознавания (обучение с учителем).

2. Разбиение множества ситуаций (объектов) по их формализованным описаниям на систему непересекающихся подмножеств (классов) — задача автоматической классификации (таксономия, кластер-анализ, обучение без учителя).

3. Определение информативного набора признаков для построения формализованного описания объекта распознавания; оценка информативности отдельных признаков и их сочетаний — задача выбора информативного набора признаков при распознавании.

4. Построение формализованного описания объекта распознавания — задача приведения исходных данных к виду, удобному для распознавания.

5. Задача *1* с учетом динамичности объекта (ситуации).

6. Задача *2* с учетом динамичности объектов (ситуаций).

7. Задачи *5*, *6*, в которых решение должно относиться к некоторому моменту времени в будущем — задача прогнозирования.

1.3.3. Виды исходной информации в задачах распознавания и прогнозирования. Все перечисленные в разд. 1.3.2 задачи могут решаться при задании исходных данных в одной из следующих форм либо в их сочетаниях:

1) изображения, полученные в различных частях спектра излучений (оптические, инфракрасные, ультразвуковые и т. д.) различными способами (телевизионные, фотографические, лазерные, радиолокационные, радиационные и т. д.) и преобразованные в цифровую форму;

2) сигналы (длинные числовые последовательности);

3) экспертные данные, числовые и другие виды символьной информации общего вида;

4) серии изображений («фильмы») любого вида из перечисленных в п. 1.

Глава 2

РАСПОЗНАВАНИЕ ОБРАЗОВ. МАТЕМАТИЧЕСКАЯ ТЕОРИЯ

2.1. Эволюция задачи распознавания образов и подходов к ее решению

Проблема распознавания в течение достаточно продолжительного времени привлекает внимание специалистов в области прикладной математики, а затем и информатики. Отметим, в частности, работы Р. Фишера, выполненные в 1920-х годах и приведшие к формированию дискриминантного анализа, постановку в начале 1940-х годов А. Н. Колмогоровым и А. Я. Хинчиным задачи о разделении смеси двух распределений, теорию статистических решений и массу работ 1950—1960-х годов, посвященных поиску и применению алгоритмов, обеспечивающих отнесение нового объекта к одному из заданных классов или разделение некоторого множества объектов на несколько непересекающихся классов. К середине 1970-х годов облик распознавания как самостоятельного научного направления стал несколько меняться, появилась возможность создания нормальной математической теории распознавания.

Одной из предпосылок этой возможности явилось выделение и отработка в процессе решения прикладных задач обработки информации ряда моделей алгоритмов распознавания — семейств алгоритмов для решения классификационных задач. К этому времени были изучены и получили практическое распространение следующие модели.

1. Модели, основанные на использовании принципа разделения (*R*-модели). Эти модели различаются главным образом заданием класса поверхностей, среди которых выбирается поверхность (или набор поверхностей), в некотором смысле наилучшим образом разделяющая элементы разных классов (см., например, [30 (гл. 2), 26 (гл. 4)]).

2. Статистические модели. Этот тип моделей алгоритмов распознавания основан на использовании аппарата математической статистики. Они применяются в тех случаях, когда известны или могут быть просто определены вероятностные характеристики классов, например соответствующие функции распределения (см., например, [5, 6 (гл. 2, 3), 30 (гл. 4, 6, 7)]).

3. Модели, построенные на основе так называемого «метода потенциальных функций» (П-модели). В основе этой модели лежит заимствованная из физики идея потенциала, определенного для любой точки пространства и зависящего от того, где расположен источник потенциала. В качестве функции принадлежности объекта классу используется потенциальная функция — всюду положительная и монотонно убывающая функция расстояния [1].

4. Модели вычисления оценок (голосования) (Г-модели). Эти модели основаны на принципе частичной прецедентности. Анализируется «близость» между частями описаний ранее классифицированных объектов и объекта, который надо распознать. Наличие близости служит частичным прецедентом и оценивается по некоторому заданному правилу (посредством числовой оценки). По набору оценок близости вырабатывается общая оценка распознаваемого объекта для класса, которая и является значением функции принадлежности объекта классу [6 (гл. 3), 12, 16, 23].

5. Модели, основанные на исчислении высказываний, в частности на аппарате алгебры логики (Л-модели). В этих моделях классы и признаки объектов рассматриваются как логические переменные, а описание классов на языке признаков представляется в форме булевых соотношений (см., например, [6 (гл. 4), 24]).

Внешне достоинства, достижения и перспективы распознавания воспринимаются в «классификационной» по существу перспективе. Не менее существенной является, однако, и другая сторона дела. Она заключается в том, что становление распознавания служит отличной моделью развития математической теории обработки и преобразования информации, развития, в процессе которого эвристические (по крайней мере по существу) методы получают строгое обоснование и начинают применяться в рамках вполне формализованных регулярных процедур. Интересно отметить, что само распознавание сегодня является достаточно разработанным вариантом такой теории, поскольку позволяет разрешать ее основную задачу — синтезировать и выбирать алгоритмические средства для извлечения полезной информации из данных того рода, который был охарактеризован выше.

Известно, что к постановке задачи распознавания прибегают в тех случаях, когда трудно строить формальные теории и применять классические математические методы, что происходит обычно по двум причинам: а) уровень формализации соответствующей предметной области и/или доступная информация таковы, что не могут составить основу для синтеза математической модели, отвечающей классическим математическим или математико-физическим канонам и допускающей изучение классическими аналитическими или численными методами; б) математическая модель в принципе может быть построена, однако ее синтез или изучение связаны с такими затратами, что они существенно превышают выигрыш, приносимый искомым решением, либо выходят за пределы существующих технических возможностей, либо делают решение задачи просто бессмысленным.

Таким образом, «двойственность» распознавания проявилась в том, что решение таких задач ввело в обиход большое число некорректных (эвристических) алгоритмов. Довольно долго подавляющее большинство приложений теории распознавания было связано с плохо формализованными областями — медициной, геологией, социологией, химией и т. д. Сегодня в этих областях еще трудно строить формальные теории и применять стандартные математические методы. В лучшем случае удается дать математическое оформление некоторым интуитивным принципам и затем применять полученные «эмпирические формализмы» для решения частных задач. Это обстоятельство определило тот факт, что на первом этапе развития теории и практики распознавания возникло большое число различных методов и алгоритмов, применявшихся без какого-либо серьезного обоснования для решения практических задач. При исследовании задачи или класса задач на базе так называемых «правдоподобных» рассуждений предлагался нестрогий, но содержательно разумный метод решения и основанный на нем алгоритм; обоснование же производилось непосредственно в эксперименте с задачами. Алгоритмы, выдержавшие подобную экспериментальную проверку, т. е. приносившие успех при решении определенных практических задач, применяются, несмотря на отсутствие математических обоснований.

Стало очевидным, что появление каждого эвристического алгоритма такого рода можно рассматривать как некоторый эксперимент, а со всем множеством экспериментов и их результатов работать как с новым для математики множеством объектов, т. е. изучать с помощью строгих математических методов множество некорректных процедур решения плохо формализованных задач. Поэтому второй этап развития теории распознавания отличался, с одной стороны, попытками ставить и решать задачу выбора в конкретной ситуации наилучшего в некотором смысле алгоритма, с другой — попытками переходить от описания отдельных некорректных алгоритмов к описанию принципов их формирования, т. е. попытками строить единообразные описания для множеств эвристических, но успешно решающих реальные задачи процедур. Подобное множество задается указанием переменных, объектов, функций, параметров и точным определением областей их вариации. Фиксация этих переменных, объектов, функций, параметров позволяет выделить из соответствующего множества, т. е. модели, некоторый конкретный алгоритм. Впервые в виде модели был представлен класс алгоритмов вычисления оценок, позднее появились описания и других моделей.

Потребность в синтезе моделей алгоритмов распознавания в первую очередь определялась необходимостью фиксировать каким-то образом класс алгоритмов при выборе оптимальной или хотя бы приемлемой процедуры решения конкретной задачи. В свою очередь, попытки построения таких моделей породили интерес к собственно «математическим» свойствам алгоритмов распознавания, в особенности к проблемам их строгого обоснования. Оказалось, что получение описания класса алгоритмов распознавания представляет собой задачу, сходную с построением классического определения алгоритма. Следовательно, необходимым условием построения теории распознавания являлось проведение классических алгоритмических исследований для понятия «алгоритм распознавания».

Анализ совокупности некорректных алгоритмов распознавания позволяет по мере их накопления выделять и описывать не только отдельные частные алгоритмы, но и принципы их формирования. Эти принципы, действующие уже над подмножествами алгоритмов и формулируемые сначала также в плохо формализованном виде, затем могут реализовываться в виде точных математических описаний. На этом этапе эвристический характер имеет собственно выбор принципа, а алгоритмы, порождаемые на основе соответствующего принципа, могут строиться стандартным образом. Именно в таком смысле формализация различных принципов построения распознающих алгоритмов приводит к появлению моделей распознающих алгоритмов.

Переход к моделям распознающих алгоритмов сам по себе не привел ни к созданию некой универсальной модели, ни к формализации процесса выбора определенной модели для решения конкретной задачи распознавания. Но появление моделей позволило ставить и решать в рамках определенной модели задачу выбора алгоритма, экстремального по функционалу качества классификации или прогноза. Построение таких оптимальных алгоритмов обычно приводится к исследованию, реализации и разработке вычислительных схем для нестандартных экстремальных задач.

Параметризация ряда алгоритмов (моделей) распознавания и возможность на основе имеющейся информации о классах определять значения параметров действительно позволяют выбирать корректные алгоритмы для некоторых подклассов задач. В большинстве практических случаев, однако, оказывается, что подкласс этот довольно узок, так как в противном случае при синтезе модели алгоритмов распознавания, описании классов и выборе признаков объектов распознавания необходимо было бы использовать весьма значительный объем априорной информации, которую можно получить, лишь располагая достаточно точной моделью изучаемых объектов и явлений. Кроме того, построение оптимального алгоритма в многопараметрической модели связано с решением трудных экстремальных задач (часто NP-полных). Достаточно часто не удается отыскать глобальный экстремум, использование же алгоритмов, соответствующих локальному экстремуму, значительно ухудшает качество распознавания и не позволяет реализовать истинные возможности модели. Иногда оказывается, что использование малопараметрических моделей, допускающих отыскание глобального экстремума, даст больший эффект, чем применение локально-экстремального алгоритма из многопараметрической модели; нет к тому же и гарантии, что оптимальный в модели алгоритм останется таким же и при работе с объектами, не участвовавшими в обучении [4, 18, 20].

Обоснование на втором этапе проводится одним из следующих трех способов:

1) экспериментально — возможность получить некое приемлемое с точки зрения пользователя «решение» поставленной задачи с помощью соответствующего алгоритма распознавания рассматривается в качестве обоснования допустимости его использования при решении данной задачи;

2) при помощи решения оптимизационной задачи и использования оптимального в рамках выбранной модели алгоритма распознавания — обоснование состоит в том, что применяется наилучший из возможных модели алгоритмов распознавания;

3) обоснование проводится так же, как и в п. 2, но к тому же доказывается, что при выполнении «ряда» «естественных» гипотез (условий), справедливых для изучаемого класса задач, алгоритмы, оптимальные в используемой модели, действительно обеспечивают высокую точность распознавания, т. е. обосновывается как выбор алгоритмов, так и выбор модели.

Очередной этап развития распознавания был связан с изучением строения совокупности некорректных алгоритмов в целом. Поскольку оказалось, что обогащение модели часто не удается сопроводить эквивалентным улучшением результатов и к тому же существует естественная граница сложности любой модели, возникла идея выбирать алгоритмы из имеющихся семейств и, используя соответствующие операции над алгоритмами (корректирующие операции), непосредственно строить из исходных алгоритмов оптимальный.

Одним из первых вариантов этой идеи явился так называемый «корректор по результатам», предусматривавший формирование решения задачи распознавания на основе результатов обработки исходной информации отдельными алгоритмами [29]. Оказалось, однако, что не существует в некотором естественном смысле «хороших» простых операций, которые позволяли бы проводить необходимую коррекцию даже в таком случае, когда в качестве допустимых алгоритмов рассматриваются ответы «да», «нет», «не знаю». Дело в том, что пространство исходных информаций и множество возможных ответов определяются содержательной постановкой задачи. Поэтому первое состоит из достаточно сложно организованных элементов (обычно векторов очень большой размерности), а второе — весьма бедно ($\{0,1\}$).

В качестве выхода из этой ситуации были предложены способ определения алгоритма распознавания, в рамки которого укладываются все существующие типы алгоритмов, и так называемый алгебраический подход к задачам распознавания и классификации, обеспечивающий эффективное исследование и конструктивное описание классов алгоритмов распознавания. Как мы уже отмечали, получение описания классов алгоритмов распознавания представляет собой задачу, сходную с построением общего определения алгоритма (аналогично тому, как это делалось в классических работах Черча, Тьюринга, А. А. Маркова и других). В этих работах интуитивное представление об алгоритме как закономерной массовой процедуре переведено в строгие определения, задающие математическую модель понятия алгоритма. Доказательство соответствия этой модели интуитивным представлениям невозможно при помощи математического аппарата, так как требует сопоставления формально заданного объекта с объектами, для которых не существует строгих формальных описаний. Ценность математической модели алгоритма состоит в том, что она хорошо согласуется с интуитивными представлениями об алгоритме, причем такое согласие устанавливается посредством опроса специалистов. В сущности, состоятельность определения алгоритма подтверждается тем, что подавляющее большинство специалистов (экспертов) считают это определение правильным. Таким образом, даже в рамках наиболее строго формализованной математической дисциплины — математической логики и теории алгоритмов — одно из основных понятий обосновывается лишь с помощью своего рода обработки результатов наблюдений.

Для построения строгой теории распознавания было необходимо провести аналогичные разработки для понятия «распознающий алгоритм», т. е. требовалось, как и в случае понятия «алгоритм», перевести интуитивные (эвристические, некорректные) представления на язык определений. Последнее эквивалентно необходимости построить математическую модель алгоритма распознавания, столь же убедительную (по экспертному критерию), как и математическая модель понятия «алгоритм».

Итак, необходимым условием создания строгой теории распознавания является проведение классических алгоритмических исследований для понятия «алгоритм распознавания». Соответствующее определение должно удовлетворять и некоторым другим условиям, главное из которых требует, чтобы с этим определением можно было работать на последующих этапах исследования алгоритма распознавания. В реальных задачах это требование оказывается достаточно сильным [20].

Изучение моделей распознающих алгоритмов позволило получить интересные теоретические результаты и решать разнообразные прикладные задачи. Вместе с тем данному методу решения задач распознавания присущи и некоторые серьезные недостатки, которые, по-видимому, не могут быть устранены при рассмотрении лишь отдельных моделей. Для преодоления этих трудностей была предложена общая теория распознающих алгоритмов, построенная на основе алгебраического подхода к решению задач распознавания и классификации, обеспечивающего эффективное исследование и конструктивное описание класса алгоритмов распознавания и предусматривающего введение такого определения алгоритма распознавания, в рамки которого укладываются все существующие модели алгоритмов [18—20].

Алгебраический подход предусматривает обогащение исходных эвристических семейств алгоритмов при помощи алгебраических операций и построение семейства, гарантирующего получение корректного алгоритма, обеспечивающего решение изучаемого класса задач. В его основе лежит идея индуктивного порождения математических объектов посредством обобщенного индуктивного определения. Выделяются базисные алгоритмы и модели распознавания и вводятся операции над ними, позволяющие последовательно порождать новые алгоритмы и модели. Выясняются условия, при которых данное семейство алгоритмов является базисным относительно введенных операций, а также свойства, которыми должна обладать модель для того, чтобы в ней нашелся алгоритм, правильно классифицирующий все объекты произвольной конечной выборки. Формируются методы построения таких алгоритмов. Смысл этого подхода состоит в том, что семейство таких алгоритмов рассматривается как некоторая алгебра, операции которой позволяют на основе базиса семейства алгоритмов строить такое расширение этого семейства, которое содержит корректный алгоритм, правильно классифицирующий конечную выборку по всем классам.

В алгебраическом подходе существенно используются особенности структуры, свойственные любой процедуре распознавания. Он предусматривает введение так называемого пространства оценок, промежуточного по отношению к исходным описаниям и допустимым ответам. Алгоритм распознавания при этом рассматривается как суперпозиция двух операторов. Первый из этих операторов — распознающий — в качестве ответов формирует элементы, называемые оценками, а второй (решающее правило) — по оценкам определяет окончательные ответы. Таким образом, необходимость иметь дело с «неудобными» пространствами исходных описаний и допустимых ответов уступает место возможности вести коррекцию в пространстве оценок (чаще всего оно представляет собой множество действительных чисел).

Важным в алгебраическом подходе является понятие полноты, связывающее отдельные задачи и модели алгоритмов: полнота некоторой задачи относительно модели означает, что при произвольном наборе априорных классификаций для рассматриваемых объектов в рамках модели может быть построен алгоритм, дающий всегда правильный ответ. Из полноты некоторой задачи относительно модели непосредственно следует существование в этой модели алгоритма, обеспечивающего абсолютную точность на материале обучения. Существенно, что построение экстремального алгоритма оказывается в большинстве случаев задачей, сравнительно легко разрешаемой стандартными математическими методами.

В рамках алгебраического подхода проведен ряд исследований, посвященных изучению и обоснованию развитых методов (некоторые из этих исследований нашли отражение в выпусках 1, 2 ежегодника). Оказалось, что проблема границы множества корректирующих операций, переход за пределы которой в процессе *расширения не дает реального эффекта, связана с фиксацией* допустимого способа использования информации применяемыми алгоритмами. Формализация и последующее исследование содержательного представления о допустимом способе использования информации алгоритмами распознавания позволили получить ряд окончательных оценок для моделей алгоритмов и множеств корректирующих операций. Так, в частности, получена универсальная верхняя граница степени для множеств операций полиномиального типа, установлены нижние границы сложности для моделей распознающих операторов вычисления оценок и для моделей распознающих операторов, основанных на принципе разделения.

Показано, что возникающие при использовании алгебраического подхода семейства алгоритмов имеют ограниченную емкость, и это обеспечивает корректность применения таких семейств в случае, когда выполнены некоторые достаточно общие гипотезы статистического характера [27, 28]. Оказалось, что во многих случаях экстремальные алгоритмы, формируемые в рамках алгебраического подхода, имеют ненулевой радиус устойчивости. Это означает, что при малом в некотором смысле изменении исходной информации классификация, порождаемая экстремальным алгоритмом, сохраняется, т. е. при выполнении достаточно общих допущений о компактности почти всюду имеет место сходимость классификаций, порождаемых экстремальными алгоритмами, к истинной. Проведены также исследования, связанные с изучением возможностей наиболее простого представления экстремальных алгоритмов.

Параллельно процессу перехода в распознавании от отдельных алгоритмов к моделям развивалась и другая ветвь исследований, связанных с использованием алгебраических методов для расширения типов исходных информаций, допустимых в задачах распознавания. В этой связи отметим теорию образов У. Гренандера и развиваемую в рамках алгебраического подхода дескриптивную теорию анализа изображений (см. Гл. 3, а также [7—11, 13, 18]).

Резюмируя изложенное, нам хотелось бы подчеркнуть, что методология распознавания используется в информатике в двух качествах:

— во-первых, по прямому назначению для решения задач распознавания в классическом смысле;

— во-вторых, как средство точного исследования плохо определенных задач.

В последнем случае эта методология реализовывается приблизительно следующим образом. Пусть, например, имеются некоторые данные, полученные в результате физического или имитационного эксперимента. Эти данные в некотором весьма ограниченном смысле характеризуют изучаемые объект или явление; необходимо попытаться свести их воедино с тем, чтобы установить, какие закономерности отражаются в имеющемся материале. Для этого выдвигается некоторая простая гипотеза, которой придается математический облик, и делается попытка «объяснить» имеющийся материал с ее помощью. Последовательное использование ряда эвристик (реализаций гипотезы) может позволить угадать модель. В противном случае происходит переход к поиску в рамках модели, порождаемой эвристикой, а затем к поиску оптимального (адекватного) эвристического принципа — модели. Если оказывается, что соответствующего принципа не существует или им нельзя практически воспользоваться, то следует сформировать некоторый конгломерат принципов, обеспечивающий выделение «федеративного» принципа; именно этот верхний уровень и соответствует возможностям и назначению алгебраического подхода.

2.2. Математическая постановка задачи распознавания

Пусть дано множество M объектов ω; на этом множестве существует разбиение на конечное число подмножеств (классов) Ω_i, $i = 1, \ldots, m$, $M = \bigcup_{i=1}^{m} \Omega_i$. Разбиение M определено не полностью. Задана лишь некоторая информация I_0 о классах Ω_i. Объекты ω задаются значениями некоторых признаков x_j, $j = 1, \ldots, N$ (этот набор всегда один и тот же для всех объектов, рассматриваемых при решении определенной задачи). Совокупность значений признаков x_j определяет описание $I(\omega)$ объекта ω. Каждый из признаков может принимать значения из различных множеств допустимых значений признаков, например, из следующих: $\{0,1\}$ — признак не выполнен или выполнен соответственно; $\{0,1, \Delta\}$, Δ — информация о признаке отсутствует; $\{0, 1, \ldots, d-1\}$ — степень выраженности признака имеет различные градации, $d > 2$; $\{a_1, \ldots, a_d\}$ — признак имеет конечное число значений, $d > 2$; $[a, b]$, $(a, b]$, $[a, b)$, a, b — произвольные числа или символы $-\infty$, $+\infty$; значениями признака x_j являются функции некоторого класса; значениями признака x_j являются функции распределения некоторой случайной величины. Описание объекта $I(\omega) = (x_1(\omega), \ldots, x_N(\omega))$ называют стандартным, если $x_j(\omega)$ принимает значение из множества допустимых значений.

Задача распознавания со стандартной информацией состоит в том, чтобы для данного объекта ω и набора классов $\Omega_1, \ldots, \Omega_m$ по обучающей информации $I_0(\Omega_1, \ldots, \Omega_m)$ о классах и описанию $I(\omega)$ вычислить значения предикатов $P_i(\omega)$ — "$\omega \in \Omega_i$", $i = 1, \ldots, m$. Информация о вхождении объекта ω в класс Ω_i кодируется символами «1» ($\omega \in \Omega_i$), «0» ($\omega \notin \Omega_i$), Δ — неизвестно, принадлежит ли ω классу Ω_i или нет, и записывается в виде так называемого информационного вектора

$$\tilde{\alpha}(\omega) = (\alpha_1(\omega), \ldots, \alpha_m(\omega)) \quad \alpha_i \in \{0, 1, \Delta\}. \tag{2.1}$$

Стандартной информацией

$$I_0(\Omega_1, \ldots, \Omega_m) \tag{2.2}$$

называют совокупность множеств $(I(\omega_1), \ldots, I(\omega_{r_m}))$ и $(\tilde{\alpha}(\omega_1), \ldots, \tilde{\alpha}(\omega_{r_m}))$ (предполагается, что среди информационных векторов нет вектора вида $(\Delta, \ldots, \Delta)$). Априорная информация в задаче распознавания с непересекающимися классами часто задается в виде так называемой таблицы обучения $T_{N,m}$ (см. табл. 2.1). Очевидно, что объекты $\omega_1, \ldots, \omega_{r_1}$ принадлежат классу Ω_1, объекты $\omega_{r_{i-1}}, \ldots, \omega_{r_i}$ принадлежат классу Ω_1 и объекты $\omega_{r_{m-1}+1}, \ldots, \omega_{r_m}$ — классу Ω_m.

2.3. Синтез модели эвристического алгоритма распознавания

Как уже отмечалось, анализ совокупности некорректных алгоритмов позволяет выявлять принципы их формирования, а формализация последних позволяет получать математические модели эвристических алгоритмов распознавания. Проиллюстрируем процесс построения такой модели на примере формализации принципа разделения, состоящего в том, что во многих задачах, где описания объектов задаются наборами значений числовых признаков (объекты суть точки n-мерного пространства), такие описания, принадлежащие разным классам, могут быть разделены поверхностями достаточно простого вида.

Рассмотрим одну из возможных формализаций. Воспользуемся простейшим классом разделяющих поверхностей — гиперплоскостями

$$\sum_{i=1}^{n} a_i x_i + a_{n+1} = 0. \tag{2.3}$$

Пусть множество допустимых объектов разделено на два класса: K_1, K_2, $K_1 \cap K_2 = \varnothing$. Пусть также известно, что объекты $S_1, \ldots, S_m$ принадлежат K_1, объекты $S_{m+1}, \ldots, S_q$ — K_2. Эти объекты, вообще говоря, неравнозначны. Поэтому введем их числовые характеристики $\gamma(S_i) = \gamma_i$ — вес объекта S_i, $i = 1, 2, \ldots, m, m+1, \ldots, q$. Таким образом, множество алгоритмов характеризуется заданием параметров $a_1, \ldots, a_{n+1}$ — коэффициентов в уравнении гиперплоскости и $\gamma_1, \ldots, \gamma_q$ — весов объектов, классификация которых проведена ранее. Процесс распознавания для $I(S) = (\alpha_1, \ldots, \alpha_n)$ производится следующим образом.

Пусть

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i + a_{n+1}. \tag{2.4}$$

Таблица 2.1. Стандартная форма задания таблицы обучения $T_{N,m}$

Объекты	Признаки и их значения						Классы
	x_1	x_2		x_j		x_N	
ω_1	$a_{1,1}$	$a_{1,2}$	$\ldots$	$a_{1,j}$	$\ldots$	$a_{1,N}$	
ω_2	$a_{2,1}$	$a_{2,2}$	$\ldots$	$a_{2,j}$	$\ldots$	$a_{2,N}$	Ω_1
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	
ω_{r_1}	$a_{r_1,1}$	$a_{r_1,2}$	$\ldots$	$a_{r_1,j}$	$\ldots$	$a_{r_1,N}$	
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
$\omega_{r_{i-1}+1}$	$a_{r_{i-1}+1,1}$	$a_{r_{i-1}+1,2}$	$\ldots$	$a_{r_{i-1}+1,j}$	$\ldots$	$a_{r_{i-1}+1,N}$	
$\omega_{r_{i-1}+2}$	$a_{r_{i-1}+2,1}$	$a_{r_{i-1}+2,2}$	$\ldots$	$a_{r_{i-1}+2,j}$	$\ldots$	$a_{r_{i-1}+2,N}$	Ω_i
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	
ω_{r_i}	$a_{r_i,1}$	$a_{r_i,2}$	$\ldots$	$a_{r_i,j}$	$\ldots$	$a_{r_i,N}$	
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
$\omega_{r_{m-1}+1}$	$a_{r_{m-1}+1,1}$	$a_{r_{m-1}+1,2}$	$\ldots$	$a_{r_{m-1}+1,j}$	$\ldots$	$a_{r_{m-1}+1,N}$	
$\omega_{r_{m-1}+2}$	$a_{r_{m-1}+2,1}$	$a_{r_{m-1}+2,2}$	$\ldots$	$a_{r_{m-1}+2,j}$	$\ldots$	$a_{r_{m-1}+2,N}$	Ω_m
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	
ω_{r_m}	$a_{r_m,1}$	$a_{r_m,2}$	$\ldots$	$a_{r_m,j}$	$\ldots$	$a_{r_m,N}$	
ω'	b_1	b_2	$\ldots$	b_j	$\ldots$	b_N	$\Omega_?$

Разделим объекты $S_1, \ldots, S_m$ на множества K_1^+, K_1^-: $S_i \in K_1^+$, если $f(I(S_i)) \geqslant 0$; $S_i \in K_1^-$, если $f(I(S_i)) < 0$. Аналогично объекты $S_{m+1}, \ldots, S_q$ разделим на множества K_2^+, K_2^-. Рассмотрим величины

$$\gamma(K_1^+) = \sum_{S_i \in K_1^-} \gamma(S_i), \qquad \gamma(K_1^-) = \sum_{S_j \in K_1^-} \gamma(S_j) \tag{2.5}$$

и аналогичные им величины $\gamma(K_2^+)$, $\gamma(K_2^-)$.

Вычисляем $f(I(S))$. Сопоставим S два числа: $\Gamma_1(S)$, $\Gamma_2(S)$ — соответственно значение функции принадлежности S классам K_1, K_2.

Если $f(I(S)) \geqslant 0$, то

$$\Gamma_1(S) = \frac{\gamma(K_1^+) + \gamma(K_2^-)}{\gamma(K_1^-) + \gamma(K_2^+)}, \qquad \Gamma_2(S) = \frac{\gamma(K_2^+) + \gamma(K_1^-)}{\gamma(K_1^+) + \gamma(K_2^-)}. \tag{2.6}$$

При $f(I(S)) < 0$:

$$\Gamma_1(S) = \frac{\gamma(K_1^-) + \gamma(K_2^+)}{\gamma(K_1^+) + \gamma(K_2^-)}, \tag{2.7}$$

и аналогично вычисляется $\Gamma_2(S)$.

По значениям $\Gamma_1(S)$ и $\Gamma_2(S)$ принимается решение о зачислении S в K_1 или K_2. Эта процедура задается решающим правилом. Рассмотрим класс решающих правил, определяемых параметром:

если $\Gamma_1(S) - \Gamma_2(S) > \delta$, то $S \in K_1$,

если $\Gamma_2(S) - \Gamma_1(S) > \delta$, то $S \in K_2$, (2.8)

если $|\Gamma_1(S) - \Gamma_2(S)| \leqslant \delta$, то решение не принимается, алгоритм отказывается от классификации S.

Таким образом, построена одна из возможных моделей, основанных на принципе разделения. Эта модель основывается на гипотезах: а) элементы классов K_1 и K_2 разделяются гиперплоскостью (по крайней мере значительная часть элементов, классификация которых представляет интерес); б) элементы классов не равнозначны по важности; меру этой важности можно выразить числом.

Реализация гипотез проведена при построении моделей

$$M(a_1, \ldots, a_{n+1}, \gamma_1, \ldots, \gamma_q, \delta), -\infty < \gamma_i, \quad a_i < +\infty, \quad \delta \geqslant 0. \tag{2.9}$$

Задание значений всех параметров модели определяет ее элемент — конкретный алгоритм распознавания.

2.4. Синтез экстремального в модели алгоритма распознавания

Если на уровне создания отдельных фиксированных алгоритмов основная проблематика связана с построением эффективных вычислительных схем и проведением экспериментов — решением прикладных задач, то на уровне моделей возникает множество новых математических задач. Среди них особо следует выделить проблемы синтеза алгоритмов, экстремальных по качеству распознавания в рамках данной модели. Функционал качества алгоритма может определяться различными способами. Обычно его определение базируется на следующем принципе. Задается способ построения объектов каждого из классов. Оценивается для фиксированного алгоритма из данной модели, какую долю объектов он классифицирует правильно, т. е. отнесет к данному классу. Полученная величина усредняется по классам и называется функционалом качества алгоритма. Задача состоит в том, чтобы в рамках модели найти алгоритм с максимальным значением функционала качества. Например, может быть задан следующий закон порождения классов K_1, K_2. Пусть описания $I(S)$ объектов S являются наборами $(a_1(S), \ldots, a_n(S))$ числовых признаков $-\infty < a_i(S) < +\infty$, $i = 1, 2, \ldots, n$.

В n-мерном пространстве заданы два нормальных распределения соответственно с математическими ожиданиями m_1, m_2 и дисперсиями σ_1, σ_2. Производится случайный выбор точек (описаний объектов) и разыгрывается по заданным законам класс, в который они зачисляются. После этого объект S, занесенный, например, в K_1, с вероятностью p причисляется к обучающей выборке, с вероятностью $1-p$ — к контрольной. То же самое делается с объектами из K_2. Пусть таким образом сформированы обучающая и контрольная выборки. В первую из них зачислены из K_1 объекты $S_{11}, \ldots, S_{1m}$, из K_2 — $S_{21}, \ldots, S_{2t}$; во вторую из K_1 — объекты $S_{31}, \ldots, S_{3v}$, из K_2 — $S_{41}, \ldots, S_{4u}$. В модели строится алгоритм A, который по описаниям $I(S_{11}), \ldots, I(S_{1m})$, $I(S_{21}), \ldots, I(S_{2t})$ дает максимальное значение функционала качества $\varphi(A) = q'/q''$, где q' — число объектов из контрольной выборки правильно классифицированных алгоритмов A; $q'' = v + u$ — число объектов в контрольной выборке.

Значение $\varphi(A)$ есть случайная величина, и ее характеристики (моменты) дают представление о точности модели на определенном типе задач распознавания. Вычисление этих характеристик является обычно далеко не тривиальным. Результаты в таких задачах удается получить лишь для сравнительно простых моделей и законов образования классов.

Более стандартным является подход, когда при фиксированной начальной информации I_0 и модели требуется найти в модели алгоритм, позволяющий максимально точно классифицировать данную совокупность S_i, $i = 1, 2, \ldots, m$, контрольных объектов, принадлежность или непринадлежность которых классам $K_1, \ldots, K_l$ известна.

Естественно, что информация типа $S_i \overline{\in} K_j$, $S_i \in K_j$ не вводится в алгоритм. Построение экстремальных в модели на заданной контрольной выборке алгоритмов приводит к решению и исследованию новых типов экстремальных задач. Таким исследованиям посвящено большое число работ, особенно в R-моделях и Γ-моделях. Пусть, например, даны описания объектов $I(S_1), \ldots, I(S_m)$ из K_1, $I(S_{m+1}), \ldots, I(S_q)$ из K_2, $I(S_i) = (\alpha_{i1}, \ldots, \alpha_{il})$. Начальной информации нет. Строится R-модель, разбиение проводится гиперплоскостью $f(\vec{x}) = \sum_{i=1}^{n} a_i x_i + a_{n+1}$. Параметрами модели являются коэффициенты гиперплоскости $a_1, \ldots, a_{n+1}$.

Решающее правило. Если $f(I(S_i)) \geqslant 0$, то $S_i \in K_1$ при $f(I(S_i)) < 0$ объект S_i заносится в K_2, $i = 1, 2, \ldots, q$.

Для каждого S_i нетрудно написать условие правильной классификации. Написав эти условия последовательно для $S_1, \ldots, S_m$, $S_{m+1}, \ldots, S_q$, получим систему линейных неравенств с неизвестными $a_1, \ldots, a_{n+1}$:

$$
\begin{aligned}
&a_1\alpha_{11} + \ldots + a_n\alpha_{1n} + a_{n+1} \geqslant 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_1\alpha_{m,1} + \ldots + a_n\alpha_{m,n} + a_{n+1} \geqslant 0,
\end{aligned}
\qquad (2.10)
$$

$$a_1\alpha_{m+1,1} + \ldots + a_n\alpha_{m+1,n} + a_{n+1} < 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_1\alpha_{q1} + \ldots + a_n\alpha_{qn} + a_{n+1} < 0.$$

Система (2.10) вообще говоря, несовместна. Для построения искомого алгоритма требуется найти максимальную совместную подсистему в (2.10). Решив ее, получим значения $a_1, \ldots, a_{n+1}$, а следовательно, экстремальный на выборке $S_1, \ldots, S_q$ алгоритм.

Выделение максимальной совместной подсистемы даже для линейных систем является трудной задачей, и ее решение требует создания специальных методов. В более сложных моделях построение экстремальных на заданной выборке алгоритмов также приводится к отысканию максимальных совместных подсистем, но неравенства в аналогах (2.10) не являются линейными.

2.5. Модель АВО: Алгоритмы распознавания, основанные на вычислении оценок

Первые алгоритмы распознавания, относящиеся к классу вычисления оценок (сокращенное название АВО), были предложены в 60-е годы [14], а на протяжении 70-х годов подробно исследованы и развиты [6, 12, 16, 18, 20, 23]; подробную библиографию по АВО можно найти в работах [6, 18, 20]. В основе алгоритмов этого класса лежит очень естественный эвристический принцип, которым пользуются часто и охотно — принцип прецедентности или частичной прецедентности, т. е. *принятие решений по аналогии, а именно:* в аналогичных (или хотя бы сходных) ситуациях следует действовать аналогично.

«Принцип действия» АВО состоит в вычислении приоритетов (оценок сходства), характеризующих «близость» распознаваемого и эталонного объекта (объектов) по системе ансамблей признаков, представляющей собой систему подмножеств заданного множества признаков.

Значение класса АВО определяется целым рядом обстоятельств.

1. Этот класс алгоритмов послужил полигоном для отработки математической теории распознавания, известной теперь как «алгебраический подход к решению задач распознавания и классификации» и посвященной построению обоснования для алгоритмов решения плохо формализованных задач, связанных с обработкой и анализом данных [20].

2. Применительно к этому классу впервые было сформулировано понятие «модели распознающего алгоритма» и были реально построены и применены для решения практических задач соответствующие модели.

3. Построение модели распознающего алгоритма позволяет перейти к систематическому выбору алгоритма для решения конкретной задачи, формализовав его. Подобный выбор соответствует выделению из некоторого известного множества алгоритмов распознавания $\{A\}$ такого алгоритма A, на котором некоторый функ-

ционал φ_A, характеризующий качество распознающего алгоритма A, обращается в максимум. Как известно, классическая задача оптимизации состоит в том, чтобы найти экстремум некоторого функционала при наличии ограничений на варьируемые переменные, причем как сам функционал, так и переменные могут быть статическими и динамическими. В результате возникает экстремальная задача для специального функционала в некотором специальным образом выбранном пространстве. Именно в этом смысле иногда говорят о синтезе алгоритма распознавания как о решении дискретной экстремальной задачи.

В качестве соответствующего функционала, заданного на пространстве распознающих алгоритмов, в классе АВО используется ошибка распознавания, возникающая при использовании конкретного алгоритма. В связи с этим возникает необходимость уточнить понятие ошибки распознавания. Довольно часто задачи распознавания и классификации обладают весьма важной особенностью: число объектов, принадлежность которых к определенным классам известна, составляет лишь незначительную часть от числа гипотетически возможных («допустимых») объектов. Очевидно, в таких ситуациях остается лишь вычислять ошибку или некоторую «обобщенную» ошибку по всем объектам, принадлежность которых определенным классам уже известна, и минимизировать именно этот функционал. С другой стороны, прекрасно известно, что при использовании ограниченного «экспериментального материала» переход к другому экспериментальному материалу может провести к кардинальному изменению функционала ошибки. Для практических приложений распознавания альтернатива такова: либо в принципе недопустимо решать задачу распознавания на ограниченном материале, либо необходимо пользоваться функционалом ошибки по этому материалу. Снятие этой проблемы связано с отысканием критериев устойчивости статистической выборки по функционалу ошибки (некоторые аспекты этой проблемы рассматриваются в монографии В. Н. Вапника [4]).

Известны, правда, и такие задачи, при решении которых априори формируется гипотеза о принадлежности различных объектов соответствующим классам, после чего на известном статистическом материале проверяется сформированная гипотеза. К сожалению, в типичных для распознавания приложениях, таких, как, скажем, медицинская диагностика, геологическое прогнозирование или социологические исследования, в связи с невысокой степенью формализации соответствующей предметной области исследователям обычно не удается сформулировать искомые разумные гипотезы (иногда же их выдвижение требует привлечения такого количества исходных сведений и таких сведений, что их наличие делает решение собственно задачи распознавания уже излишним).

4. Поскольку оптимизационные процедуры реализуются на специальным образом выбираемом пространстве, то для решения задачи синтеза оптимального алгоритма необходимо определить класс алгоритмов, на котором производится оптимизация, таким

образом, чтобы можно было ее реально осуществить. Это значит, в сущности, что нужна, во-первых, модель, единообразно представляющая достаточно широкий класс алгоритмов распознавания, и, во-вторых, эта модель должна задаваться рядом параметрических объектов, характеризующих данный класс алгоритмов распознавания. Подобная модель позволяет установить взаимно-однозначное соответствие между алгоритмами и набором числовых параметров (класс алгоритмов — область многомерного пространства). В таком случае задание конкретного алгоритма A, принадлежащего рассматриваемому классу алгоритмов распознавания $\{A\}$, позволяет сопоставить ему значение функционала обобщенной ошибки φ_A и, следовательно, определить функционал, оценивающий качество алгоритма распознавания, на точках соответствующего параметрического пространства.

В-третьих, значения этого функционала должны вычисляться эффективно. Задание параметрической модели алгоритма обеспечивает возможность подбирать и изменять его параметры таким образом, чтобы значение функционала качества изменялось в нужном направлении. Если существует эффективный способ вычисления функционала φ_A, то принципиально возможно построение алгоритма A^*, на котором достигается экстремум φ_A. Хотя известно, что последняя задача не всегда может быть решена до конца (в ряде ситуаций приходится ограничиваться локальными экстремумами), в тех случаях, когда абсолютно экстремальный алгоритм может быть найден, имеется гарантия, что при данном исходном материале в данном классе алгоритмов для заданного контрольного материала не существует лучшего алгоритма распознавания, чем A^* [16, 20]. Опыт решения задач распознавания свидетельствует о том, что часто основная «различающая» информация заключена не в отдельных признаках, а в различных их сочетаниях. Уже в середине 60-х годов предпринимаются попытки строить алгоритмы распознавания, позволяющие учитывать информацию, заключенную в комбинациях признаков. Наибольшую известность среди алгоритмов этого типа приобрели так называемые тестовые алгоритмы (впервые описаны в [14]), программа «Геометрия» (блок «Кора») [2] и алгоритм «Кора» [3]. Вычислительная сложность (объем перебора) заставила в двух последних алгоритмах ограничиваться конъюнкциями сложности три. При использовании тестовых алгоритмов возникали проблемы, связанные с синтезом совокупности всех тупиковых тестов (см. ниже) для данной задачи. Класс АВО доводит идею использования совокупностей признаков до логического конца: поскольку не всегда известно, какие именно сочетания признаков информативны в наибольшей степени, то в АВО степень похожести объектов вычисляется в процессе сопоставления всех возможных (или определенных — для случая, когда известны сочетания признаков, обладающие наибольшей разделяющей силой) сочетаний признаков, входящих в описание объектов. Как уже отмечалось, для вычисления соответствующих оценок близости объектов в АВО имеются несложные аналитические формулы, снимающие перебор при реализации процедуры распознавания (точнее — при настройке параметров алгоритма на задачу в процессе обучения). Кроме того, АВО позволяет учитывать различия в информативности (разделяющей силе) отдельных признаков и их сочетаний, а также различия в репрезентативности отдельных объектов, включенных в таблицу обучения.

5. Важным отличием АВО от других классов алгоритмов распознавания являются значительно более слабые требования к исходной информации, так как они не предусматривают наличия сведений о моментах и других статистических характеристиках. Исходный материал может представляться не только в числовой форме, но и задаваться описаниями на естественном языке.

6. Класс АВО позволяет решать не только статические, но и динамические задачи — к последним часто сводятся задачи прогнозирования.

Алгоритм распознавания, основанный на принципе прецедентности или частичной прецедентности, сравнивает описания распознаваемого объекта $I(\omega')$ с $T_{N,m}$ и принимает решение о том, к какому классу следует отнести этот объект. Решение выносится на основе вычисления степени сходства распознаваемого объекта (строки) со строками, принадлежность которых к заданным классам известна.

Пусть заданы стандартные описания объектов $\{\omega^{\Omega_i}\}$, $\omega^{\Omega_i} \in \Omega_i$ и $\{\omega^{\overline{\Omega}_i}\}$, $\omega^{\overline{\Omega}_i} \overline{\in} \Omega_i$. Необходимо определить принадлежность классу Ω_i, $i = 1, \ldots, m$, предъявленного для распознавания объекта ω'. Если введен способ определения величины близости для некоторых частей описания $J(\omega')$ и соответствующих частей описаний $\{J(\omega^{\Omega_i})\}$ и $\{J(\omega^{\overline{\Omega}_i})\}$, то можно сформировать «обобщенную близость» между объектом ω' и множествами объектов $\{\omega^{\Omega_i}\}$ и $\{\omega^{\overline{\Omega}_i}\}$ соответственно. В простейшем случае обобщенная близость приравнивается сумме близостей между частями описаний. В результате характеристику вида $\Gamma_i(\omega') = \Gamma_i^{\Omega_i} - \Gamma_i^{\overline{\Omega}_i}$, где $\Gamma_i^{\Omega_i}$ и $\Gamma^{\overline{\Omega}_i}$ — значения соответствующих обобщенных близостей, естественно считать значением функции принадлежности объекта ω' классу Ω_i. Величина $\Gamma_i(\omega')$ называется оценкой объекта ω' по классу Ω_i (иногда мы будем обозначать ее как $\Gamma(\omega', \Omega_i)$).

Описания объектов $\{\omega'\}$, предъявленные для распознавания, переводятся алгоритмом распознавания в числовую матрицу $\{\Gamma_i\}_{\{\omega'\}\times m}$ — матрицу оценок. Эта процедура включает два этапа: сначала подсчитывается оценка ω' по каждой строке из $T_{N,m}$, а затем полученные оценки используются для получения суммарных оценок по каждому из классов Ω_i. Применение решающего правила к матрице оценок позволяет получить матрицу $\{\alpha_i\}_{\{\omega'\}\times m}$ информационных векторов объектов $\{\omega'\}$.

Рассмотрим процедуру построения оценок $\Gamma_i(\omega')$, используемую в тестовых алгоритмах и АВО.

В основе тестовых алгоритмов лежит понятие теста [32]. Тестом таблицы $T_{N,m}$ называется совокупность столбцов $x_{t_1}, \ldots, x_{t_q}$, таких, что после удаления из $T_{N,m}$ всех столбцов, за исключением имеющих номера $t_1, \ldots, t_q$, в полученной таблице $T_{N-q,m}$ все пары строк, принадлежащих разным классам, различны. Тест $\{x_{t_1}, \ldots, x_{t_q}\}$ называется тупиковым, если никакая его часть не является тестом.

Пусть $\{T\}$ — множество всех тупиковых тестов $T_{N,m}$ и $T = \{x_{t_1}, \ldots, x_{t_q}\} \in \{T\}$. Выделим в описании распознаваемого объекта $I(\omega')$ часть $(b_{t_1}, \ldots, b_{t_q})$, соответствующую признакам $x_{t_1}, \ldots, x_{t_q}$, и сопоставим ее со всеми частичными описаниями $(a_{r_{t_1}}, \ldots, a_{r_{t_q}})$ объектов $I(\omega_r)$ таблицы $T_{N,m}$, $r = (r_{i-1} + 1), \ldots, r_i$, $i = 1, \ldots, m$. Подсчитаем число совпадений $\Gamma_T(\omega', \Omega_i)$ частичных описаний $(b_{t_1}, \ldots, b_{t_q})$ со всеми частичными описаниями $(a_{r_{t_1}}, \ldots, a_{r_{t_q}})$ объектов i-го класса. Величина $\Gamma_T(\omega', \Omega_i)$ представляет собой число строк этого класса, близких распознаваемой строке ω' по тесту T, т. е. оценку строки ω' для класса Ω_i по тесту T. Аналогичным образом вычисляется оценка для ω' по остальным тестам (для всех классов). Величина

$$\Gamma(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{T \in \{T\}} \Gamma_T(\omega', \Omega_i) \qquad (2.11)$$

представляет собой оценку объекта ω' по классу Ω_i.

Известны разновидности тестовых алгоритмов, в которых при формировании оценок $\Gamma_T(\omega', \Omega_i)$ учитываются различия в представительности («важности») отдельных строк таблицы $T_{N,m}$ и признаков, включенных в стандартные описания объектов. Для этого используются числовые коэффициенты — веса признаков и веса объектов. Существует довольно много способов введения таких весов, чаще всего они задаются из эвристических соображений, исходя из специфики решаемой задачи, с помощью экспертной оценки и т. п. Для тестовых алгоритмов в свое время была предложена весьма естественная мера важности признака — информационный вес [14]

$$p(x_j) = \frac{r_{x_j}(N, m)}{r(N, m)}, \qquad (2.12)$$

где $r(N, m)$ — число тупиковых тестов таблицы $T_{N,m}$, $r_{x_j}(N, m)$ — число тупиковых тестов $T_{N,m}$, содержащих признак x_j. Чем в большее число тупиковых тестов входит признак x_j, тем больше его информационный вес $p(x_j)$, тем значительнее его роль в описании объектов $T_{n,m}$.

Если учитываются веса признаков $p(x_1), \ldots, p(x_N)$ и объектов таблицы $T_{N,m}$ $\gamma(\omega_1), \ldots, \gamma(\omega_{n_m})$, то каждое совпадение частичного описания распознаваемого объекта $(b_{t_1}, \ldots, b_{t_q})$ с частичным описанием объекта из $T_{N,m}$ $(a_{r_{t_1}}, \ldots, a_{r_{t_q}})$, соответствующих некоторому тесту T, оценивается величиной вида

$$\Gamma_T(\omega', \omega_r) = \gamma(\omega_r)(p(x_{t_1}) + \ldots + p(x_{t_q})),$$
$$r = (r_{i-1}+1), \ldots, r_i; \qquad \omega_r \in \Omega_i. \tag{2.13}$$

В результате оценка объекта ω' по классу Ω_i принимает следующий вид:

$$\Gamma(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{T \in \{T\}} \sum_{r=r_{i-1}+1}^{r_i} \Gamma(\omega', \omega_r). \tag{2.14}$$

Переход от тестовых алгоритмов к АВО связан с расширением видов подмножеств множества признаков, по которым проводится сопоставление распознаваемого объекта с объектами из $T_{N,m}$, и построением эффективных формул вычисления оценок $\Gamma(\omega', \Omega_i)$ для различных случаев задания подмножеств множества признаков (в АВО они называются опорными множествами алгоритма распознавания). В тестовых алгоритмах в качестве системы опорных множеств алгоритма используются множества тупиковых тестов (при отсутствии пересечений классов в $T_{N,m}$ и бинарных признаках). В АВО рассматривается два случая: наличие [16, 23] и отсутствие ограничений на систему опорных множеств алгоритма [12]. В первом случае наиболее распространенными являются системы опорных множеств, составленные изо всех подмножеств множества признаков фиксированной длины q, $q = 2, \ldots, N-1$, либо из всех непустых подмножеств множества признаков.

Рассмотрим полный набор признаков $\langle x_1, \ldots, x_N \rangle$ и выделим систему подмножеств множества признаков (систему опорных множеств алгоритма) $S_1, \ldots, S_l$. Удалим произвольный подпабор признаков из строк $\omega_1, \omega_2, \ldots, \omega_{r_m}$, ω' и обозначим полученные строки через $S\tilde{\omega}_1, S\tilde{\omega}_2, \ldots, S\tilde{\omega}_{r_m}$, $S\tilde{\omega}'$. Правило близости, позволяющее оценить похожесть строк $S\tilde{\omega}'$ и $S\tilde{\omega}_r$, состоит в следующем. Пусть «усеченные» строки содержат q первых признаков, т. е. $S\tilde{\omega}_r = (a_1, \ldots, a_q)$ и $S\omega' = (b_1, \ldots, b_q)$, и заданы пороги $\varepsilon_i, \ldots, \varepsilon_q$, δ. Строки $S\tilde{\omega}_r$ и $S\omega'$ считаются похожими, если выполняется не менее чем δ неравенств вида $|a_j - b_j| \leqslant \varepsilon_j$, $j = 1, \ldots, q$. Величины $\varepsilon_1, \ldots, \varepsilon_q$, δ входят в качестве параметров в модель класса алгоритмов типа АВО.

Рассмотрим процедуру вычисления оценок по подмножеству S_1. Для остальных подмножеств она полностью аналогична. В табл. 2.1 $T_{N,m}$ выделяются столбцы, соответствующие признакам, входящим в S_1; остальные столбцы вычеркиваются. Проверяется близость строки $S_1\tilde{\omega}'$ со строками $S_1\tilde{\omega}_1, \ldots, S_1\tilde{\omega}_{r_1}$, принадлежащими классу Ω_1. Число строк этого класса, близких по выбранному критерию классифицируемой строке $S_1\tilde{\omega}'$, обозначается через $\Gamma_{S_1}(\omega', \Omega_1)$; эта величина представляет собой оценку строки ω' для класса Ω_1 по опорному множеству S_1. Аналогичным образом вычисляются оценки для остальных классов: $\Gamma_{S_1}(\omega', \Omega_2), \ldots, \Gamma_{S_1}(\omega', \Omega_l)$. Применение подобной процедуры ко всем остальным опорным множествам алгоритма позволяет получить систему оценок $\Gamma_{S_1}(\omega', \Omega_1)$, $\Gamma_{S_2}(\omega', \Omega_m)$, ..., $\Gamma_{S_l}(\omega', \Omega_1)$, ..., $\Gamma_{S_l}(\omega', \Omega_m)$.

Величины

$$\Gamma(\omega', \Omega_1) = \Gamma_{S_1}(\omega', \Omega_1) + \Gamma_{S_2}(\omega', \Omega_1) + \ldots + \Gamma_{S_l}(\omega', \Omega_1) = \\ = \sum_{S_A} \Gamma(\omega', \Omega_1); \tag{2.15}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\Gamma(\omega', \Omega_m) = \Gamma_{S_1}(\omega', \Omega_m) + \Gamma_{S_2}(\omega', \Omega_m) + \ldots + \Gamma_{S_l}(\omega', \Omega_m) = \\ = \sum_{S_A} \Gamma(\omega', \Omega_m)$$

представляют собой оценки строки ω' для соответствующих классов по системе опорных множеств алгоритма S_A. На основании анализа этих величин принимается решение либо об отнесении объекта ω' к одному из классов Ω_i, $i = 1, \ldots, m$, либо об отказе от его распознавания. Решающее правило может принимать различные формы, в частности, распознаваемая строка может быть отнесена к классу, которому соответствует максимальная оценка, либо к тому, чья оценка превышает оценки всех остальных классов не меньше чем на определенную пороговую величину η_1, либо величина отношения соответствующей оценки к сумме оценок для всех остальных классов должна быть не менее величины некоторого порога η_2 и т. д. Параметры типа η_1 и η_2 также включаются в модель АВО.

Пример. Заданы таблица обучения и распознаваемый объект ω' (табл. 2.2).

Пусть $S_1 = \langle x_1, x_2 \rangle$, $S_2 = \langle x_3, x_4 \rangle$, $S_3 = \langle x_5, x_6 \rangle$; строки будем считать близкими, если они полностью совпадают. Применение описанной выше процедуры вычисления оценок позволяет получать следующее:

$$S_1\colon \Gamma_{S_1}(\omega', \Omega_1) = 1; \quad \Gamma_{S_1}(\omega', \Omega_2) = 2;$$

$$S_2\colon \Gamma_{S_2}(\omega', \Omega_1) = 2; \quad \Gamma_{S_2}(\omega', \Omega_2) = 1;$$

$$S_3\colon \Gamma_{S_3}(\omega', \Omega_1) = 1; \quad \Gamma_{S_3}(\omega', \Omega_2) = 0.$$

$$\Gamma_{S_A}(\omega', \Omega_1) = \Gamma_{S_1}(\omega', \Omega_1) + \Gamma_{S_2}(\omega', \Omega_1) + \Gamma_{S_3}(\omega', \Omega_1) = \\ = 1 + 2 + 1 = 4;$$

$$\Gamma_{S_A}(\omega', \Omega_2) = \Gamma_{S_1}(\omega', \Omega_2) + \Gamma_{S_2}(\omega', \Omega_2) + \Gamma_{S_3}(\omega', \Omega_2) = \\ = 2 + 1 + 0 = 3.$$

Согласно решающему правилу, реализующему принцип простого большинства голосов, строка ω' зачисляется в класс Ω_1, так как $\Gamma_{S_A}(\omega', \Omega_1) > \Gamma_{S_A}(\omega', \Omega_2)$.

Определение распознающей части класса АВО сводится к формализации следующих этапов, соответствующих последователь-

Таблица 2.2

Объекты	Признаки и их значения						Классы
	x_1	x_2	x_3	x_4	x_5	x_6	
ω_1	0	0	0	0	0	0	Ω_1
ω_2	0	1	0	0	1	1	
ω_3	1	1	0	1	1	1	
ω_4	0	1	0	1	0	1	Ω_2
ω_5	1	1	1	1	1	1	
ω_6	1	1	0	0	0	1	
ω'	1	1	0	0	0	0	$\Omega_?$

ности реализации процедуры распознавания: 1) выделяется система опорных множеств алгоритма, по которым производится анализ распознаваемых объектов; 2) вводится понятие близости на множестве частичных описаний объектов; 3) задаются правила: а) вычисления оценки для пар объектов по значению степени подобия эталонного и распознаваемого объекта; б) формирования величин оценок для каждого из эталонных классов по фиксированному опорному множеству на основе оценок для пар объектов; в) формирования суммарной оценки для каждого из эталонных классов по всем опорным множествам; г) принятия решения, обеспечивающего на основе оценок для классов отнесение распознаваемого объекта к одному из классов или отказывающего этому объекту в классификации.

Фиксация структурных параметров, т. е. способа выбора системы опорных множеств, типа функции близости, правил вычисления оценок и решающего правила, определяет выбор подкласса алгоритмов типа АВО, а задание соответствующих параметров — конкретный алгоритм типа АВО. Модель класса АВО — параметрическая, т. е. имеет место взаимно-однозначное соответствие между конкретными алгоритмами и наборами числовых параметров. В таком случае задание конкретного алгоритма, принадлежащего рассматриваемому классу, позволяет сопоставить ему значение некоторого функционала качества распознавания (например, число ошибок и отказов от распознавания на таблице обучения) и, следовательно, определить последний на точках параметрического пространства алгоритма.

Если строить вычислительную процедуру по приведенному описанию алгоритма, то при большой мощности системы опорных множеств вычислительная сложность оказывается высокой. Так, при выборе в качестве системы опорных множеств алгоритма всех подмножеств множества признаков мощности q число опорных множеств равно C_N^q, а число слагаемых в формуле, определяющей величину $\Gamma_{S_A}(\omega', \Omega_i)$, равно $(r_i - r_{i-1})\, C_N^q$.

Как уже отмечалось, существенным достоинством, отличающим АВО, является то обстоятельство, что для вычисления оценок, определяющих принадлежность распознаваемого объекта одному из заданных классов, существуют простые аналитические формулы, заменяющие сложные переборные процедуры (возникающие при вычислении оценок близости по системе опорных множеств). Поскольку эффективность (в вычислительном смысле) вычисления функционала качества в АВО полностью определяется эффективностью процедуры вычисления оценок, то принципиально возможно построение оптимального алгоритма. В случаях, когда может быть найден абсолютный экстремальный алгоритм, имеется гарантия, что при заданном исходном материале $T_{N,m}$ в данном классе не существует лучшего алгоритма распознавания.

Известны два метода, комбинация которых позволяет находить достаточно простые формулы для практически интересных моделей АВО при условии, что используются пороговые функции близости (принимают значения 0 или 1) и $p(S) = p_{x_{t_1}} + \ldots + p_{x_{t_q}}$ (вес опорного множества равен сумме весов входящих в него признаков).

1. Первый метод [20] использует свойство оценок для класса $\Gamma(\omega', \Omega_i)$, когда в качестве оценки для класса по опорному множеству S_u, $u = 1, \ldots, l$, используется оценка вида (2.13):

$$\Gamma_{S_u}(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{\omega_r \in \Omega_i} \gamma(\omega_r)(p_{t_1} + \ldots + p_{t_q})\, B_{S_u}(\omega', \omega_r), \tag{2.16}$$

где $t_1, \ldots, t_q$ — совокупность единичных координат характеристического вектора, определяющего опорное множество S_u, а B_{S_u} — функция близости частичных описаний объектов $S\omega'$ и $S\omega_r$, принимающая значения «1» или «0» в зависимости от числа выполненных неравенств вида $|a_j - b_j| \leqslant \varepsilon_j$, $j = t_1, \ldots, t_q$ (см. выше).

Пусть $v_j(\omega', \omega_r)$ — число различных значений, которые может принимать количество опорных множеств $S_u \in S_A$, содержащих некоторый фиксированный признак x_j и таких, что $B(S\omega', S\omega_2) = 1$. Можно показать, что оценка для класса Ω_i принимает вид

$$\Gamma(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{\omega_r \in \Omega_i} \gamma(\omega_r) \sum_{j=1}^{N} p_j v_j(\omega', \omega_r). \tag{2.17}$$

Если число различных значений $v_j(\omega', \omega_r)$ невелико, то внутренняя сумма свертывается в небольшое число слагаемых (практически исключается суммирование по системе опорных множеств) и сложность вычисления $\Gamma(\omega', \Omega_i)$ становится пропорциональной длине обучающей выборки. Доказано, в частности, что при использовании пороговой функции близости, значение которой определяется лишь числом выполненных и невыполненных неравенств вида $|a_j - b_j| \leqslant \varepsilon_j$, величины $v_j(\omega', \omega_r)$, $j = 1, \ldots, N$, принимают не более двух значений [20].

Приведем аналитические формулы, обеспечивающие эффективное вычисление оценок $\Gamma(\omega', \Omega_i)$ для двух способов введения ограничений на систему опорных множеств алгоритма [23]:

а) S_A совпадает с системой всех подмножеств мощности q множества признаков $\{1, \ldots, N\}$:

1° функция близости принимает значение «1», если выполняется не менее δ (порог по числу выполненных неравенств вида $|a_{t_u} - b_{t_u}| \leqslant \varepsilon_{t_u}$):

$$\Gamma(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{\omega_r \in \Omega_i} \gamma(\omega_r) \left(\sum_{j=0}^{\delta} C^{q-j}_{z(\omega', \omega_r)} C^{j}_{N-z(\omega', \omega_r)} \right), \qquad (2.18)$$

где $z(\omega', \omega_r)$ — число выполненных неравенств вида $|a_j - b_j| \leqslant \leqslant \varepsilon_j$ для пары (ω', ω_r) $j = 1, \ldots, N$;

2° функция близости принимает значение «1», если выполнены все неравенства вида $|a_{t_u} - b_{t_u}| \leqslant \varepsilon_{t_u}$ $(\delta = 0)$, т. е. в опорное множество входят только совпадающие по порогу ε_{t_u} признаки:

$$\Gamma(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{\omega_r \in \Omega_i} \gamma(\omega_r)\, C^{q}_{z(\omega', \omega_r)}; \qquad (2.19)$$

б) S_A совпадает с системой всех непустых подмножеств множества признаков $\{1, \ldots, N\}$:

1° функция близости та же, что и в случае а1°:

$$\Gamma(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{\omega_r \in \Omega_i} \gamma(\omega_r) \left[(2^{z(\omega', \omega_r)} - 1) \sum_{j=0}^{\delta} C^{j}_{N-z(\omega', \omega_r)} \right]; \qquad (2.20)$$

2° функция близости та же, что и в случае а2°:

$$\Gamma(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{\omega' \in \Omega_i} \gamma(\omega_r)\, (2^{z(\omega', \omega_r)} - 1). \qquad (2.21)$$

2. Второй способ получения эффективных формул вычисления оценок [12] основан на следующих двух утверждениях:

а) если система опорных множеств S_A состоит из непересекающихся подмножеств, то оценка $\Gamma(\omega', \Omega_i)$ для такого алгоритма равна сумме оценок для алгоритмов, опорными множествами которых являются подмножества, образующие S_A (каждому подмножеству соответствует отдельный алгоритм):

$$S_A = \bigcup_u S_u; \qquad \Gamma_A(\omega', \Omega_i) = \sum_u \Gamma^{A_u}(\omega', \Omega_i); \qquad (2.22)$$

б) если характеристическая функция алгоритма класса АВО является элементарной конъюнкцией, то задача построения эффективной формулы вычисления оценок обеспечивается описанным выше способом *1*.

Практика распознавания показывает, что в некоторых случаях априори известны поднаборы признаков, которые следует учитывать при сопоставлении распознаваемого объекта с объектами таблицы обучения. Эти подмножества признаков не всегда совпадают с частными случаями; подмножества могут иметь различную длину, могут задаваться запрещенные или «избыточные» комбинации и т. п. В [12] аналитические формулы вычисления оценок получены для случая произвольных опорных множеств.

Вывод этих формул основан на введении характеристической булевой функции системы опорных множеств алгоритма f_{S_A} и установлении взаимно-однозначного соответствия между подмножествами множества признаков и булевыми векторами длины N (вершинами единичного N-мерного куба).

Пример. Рассмотрим те же, что и в предыдущем примере, таблицу обучения и распознаваемый объект, но ограничимся лишь первыми четырьмя признаками. Закодировав вхождение признака в опорное множество через «1», а невхождение через «0», каждому подмножеству множества признаков $\langle x_1, x_2, x_3, x_4\rangle$ можно сопоставить бинарный вектор, или, что то же самое, вершину единичного четырехмерного куба. На множестве этих векторов можно определить характеристическую булеву функцию, единицы которой будут определять подмножества множества признаков, включенные в систему опорных множеств алгоритма S_A.

Пусть $S_A = \{S_1, S_2\}$, $S_1 = \langle x_2, x_3\rangle$ (вершина), $S_2 = \langle x_1, x_2, x_3\rangle$. В таком случае $f_{S_A} = \bar{x}_1x_2x_3\bar{x}_4 \vee x_1x_2x_3\bar{x}_4 = x_2x_3\bar{x}_4$.

В [12] показано, что в тех случаях, когда множество единиц f_{S_A} образует в единичном N-мерном кубе интервал или сумму непересекающихся интервалов, можно построить эффективные формулы вычисления оценок. Напомним, что подмножество вершин единичного N-мерного куба называется интервалом, если оно соответствует некоторой элементарной конъюнкции. Очевидно, что все грани, ребра и вершины единичного N-мерного куба являются интервалами.

Система опорных множеств организована следующим образом (соответствующий интервал представлен ребром, соединяющим вершины): в нее включены все признаки, входящие в дизъюнктивную нормальную форму (ДНФ) характеристической функции без отрицания (x_2 и x_3), не включены признаки, входящие в ДНФ с отрицанием (x_4), а по остальным признакам (x_1) происходит полная вариация, т. е. рассматриваются подмножества, как включающие, так и не включающие эти признаки.

Эффективная формула для вычисления оценок в тех случаях, когда характеристической функции системы опорных множеств соответствует интервал, а функция близости предполагает, что $\delta = 0$, имеет вид

$$\Gamma(\omega', \Omega_i) = \frac{1}{r_i - r_{i-1}} \sum_{\omega_r^* \in \Omega_i} \gamma(\omega_r^*) \, 2^{z^*(\omega', \omega_r^*)}, \qquad (2.23)$$

где ω_r^* — «эффективный» для ω' объект таблицы обучения. В (2.23) учитывается вклад только тех объектов $T_{N,m}$ («эффективных»), постоянная часть которых близка (в смысле используемой функции близости) постоянной части ω'; $z^*(\omega', \omega_r^*)$ — число выполненных неравенств вида $|a_j - b_j| \leqslant \varepsilon_j$ на варьируемой части.

Таким образом, при условии $\varepsilon_1, \ldots, \varepsilon_4 = 0$ и $\gamma_1, \ldots, \gamma_6 = 1$ и учитывая, что в Ω_1 для ω' эффективны объекты ω_2 и ω_3, а в Ω_2 — объекты ω_4 и ω_6, $z^*(\omega', \omega_2) = 0$, $z^*(\omega', \omega_3) = 1$, $z^*(\omega', \omega_4) = 0$, $z^*(\omega', \omega_6) = 1$, получаем $\Gamma(\omega', \Omega_1) = (1/3)(1 \cdot 2^0 + 1 \cdot 2^1) = 1$ и $\Gamma(\omega', \Omega_2) = (1/3)(1 \cdot 2^0 + 1 \cdot 2^1) = 1$. Полученный результат означает, что при указанном выборе системы опорных множеств объект ω' не классифицируется.

Если характеристической функции соответствует сумма непересекающихся интервалов (представляется ортогональной ДНФ), как например, в случаях $S_A = \{S_1, S_2, S_3, S_4, S_5\}$, $S_1 = \langle x_2, x_3 \rangle$, $S_2 = \langle x_1, x_2, x_3 \rangle$, $S_3 = \langle x_1, x_3, x_4 \rangle$, $S_4 = \langle x_1, x_3 \rangle$, $S_5 = \langle x_1, x_2, x_4 \rangle$, $f_{S_A} = x_2 x_3 \bar{x}_4 \bigvee x_1 \bar{x}_2 x_3 \bigvee x_1 \bar{x}_3 x_4$, то при вычислении оценок по формуле (2.23) она применяется к каждому интервалу отдельно, после чего результаты суммируются.

В [12] показано, что сложность формулы вычисления оценок АВО при произвольной S_A пропорциональна сложности ДНФ, представляющей характеристическую функцию системы опорных множеств алгоритма. Это означает, что построение простой формулы вычисления оценок $\Gamma(\omega', \Omega_i)$ связано с задачей минимизации булевых функций в классе ДНФ, а точнее — с задачей построения кратчайшей ортогональной ДНФ или ДНФ, в которой каждый интервал имеет небольшое число пересечений с соседними. В общем случае задача такого синтеза неразрешима, и потому следует пользоваться приближенными алгоритмами, обеспечивающими получение «достаточно простых» ортогональных ДНФ или ДНФ с небольшим числом взаимных пересечений интервалов [12, 17].

Таким образом, если для вычисления расстояний $\rho_j(a_j, b_j)$, $j = 1, \ldots, N$, существует эффективный алгоритм и число операций при одном таком вычислении не превосходит некоторой величины Q, то число операций при вычислении всех величин $\Gamma(\omega', \Omega_i)$, $i = 1, \ldots, m$, не превосходит $2QNm$. Число операций при распознавании одного объекта в фиксированном алгоритме класса АВО пропорционально «площади» таблицы T_{Nm} с коэффициентом пропорциональности, не превосходящим $2Q$ (см. табл. 2.1). Обоснование сведения задачи построения экстремальных алгоритмов класса АВО к отысканию экстремумов функции многих переменных выполнено в [16]. Для проведения оптимизации могут быть применены методы переборного типа (при небольшом числе параметров), градиентного типа или случайного поиска.

Отметим, что существуют и другие способы задания системы опорных множеств. Так, в последнее время были изучены модели АВО с опорными множествами, задаваемыми на $T_{N,m}$ локальными окрестностями невысоких порядков, так называемыми двухиндексными опорными множествами, а также предложена модель ДАВО (АВО с двухмерными опорными множествами [10, 11].

Класс АВО успешно используется для решения задач медицинской диагностики, геологического прогноза, обработки социологической информации, идентификации технологических процессов и управления ими, прогноза свойств синтезируемых химических соединений, оптимального выбора алгоритма и оптимизации процессов принятия решений, профессионального отбора, автоматизации обработки экспериментальных данных и т. д. Алгоритмы этого класса позволяют решать задачи распознавания всех основных типов: отнесение объекта к одному и заданных классов, автоматическую классификацию, выбор системы признаков для описания объектов распознавания и оценку их информативности. Кроме того, модель АВО послужила основой для определения первого параметрического класса алгоритмов распознавания изображений [10, 11].

2.6. Распознавание при представлении исходных данных в виде длинных последовательностей [2]

Одна из наиболее универсальных форм представления информации в задачах распознавания — длинные последовательности. Поскольку вектор признаков у любого объекта можно описать бинарной последовательностью, эта форма является в некотором смысле наиболее общей. Специфика длинных последовательностей заключается в том, что большой объем информации делает работу с ними как целыми практически неосуществимой. Возникает задача поиска более компактных описаний, достаточных для распознавания, причем с учетом дополнительных трудностей, возникающих в практических задачах из-за неполноты данных и влияния шумовых помех. Вопросы построения подобных наборов признаков получили в последние годы достаточно широкое освещение в литературе.

Наиболее естественный способ построения простых и эффективных наборов признаков для описания и распознавания длинных последовательностей заключается в вычислении некоторых частотных характеристик и сравнении их с эталонами. В рамках такого подхода необходимо зафиксировать некоторый набор векторов малой размерности, вычислить частоты вхождения векторов из этого набора в последовательность в качестве подпоследова-

---

[2] Разд. 2.6 написан Ю. Г. Сметаниным.

тельности и считать полученные частоты признаками, описывающими объект. Большим достоинством таких признаков является легкость их вычисления и небольшой объем памяти, необходимой для хранения и обработки. Кроме того, подобные признаки позволяют устранить влияние трудностей, как правило, возникающих в практических задачах: наличие небольших шумовых помех, искажающих вид последовательности, слабо сказывается на величинах частот; смещение точки, с которой начинается дискретизация и отсчет признаков, также не играет существенной роли; выбор порядка сканирования многомерных объектов, существенно изменяющий вид последовательности, до известной степени может быть нейтрализован надлежащим подбором характерных конфигураций.

Для иллюстрации последнего утверждения можно привести следующий пример. Пусть имеется бинарное изображение размера $n \times n$, сканируемое слева направо и сверху вниз. Для точки с координатами (i, j), получающей в последовательности номер $k = ni + j$, соседи по горизонтали получают номера $k - 1$ и $k + 1$, а соседи по вертикали — номера $k - n$ и $k + n$. Если пренебречь краевыми эффектами, то число подслов вида «11» — это число появлений двух единиц подряд по горизонтали; для определения аналогичной характеристики по вертикали необходимо подсчитать число векторов «11» в совокупности пар $(a_1\ a_{n+1})$, (a_2, a_{n+2}), $\ldots$, $(a_{n\cdot n-n}, a_{n\cdot n})$, т. е. в совокупности подпоследовательностей длины 2, у которых номер второго элемента отличается от номера первого элемента на n. Подсчет числа векторов «11» в совокупности пар (a_1, a_2), (a_2, a_3), $\ldots$, $(a_{n\cdot n-1}, a_{n\cdot n})$, (a_1, a_{n+1}), $(a_2 a_{n+2})$, $\ldots$, $(a_{n\cdot n-n}, a_{n\cdot n})$ дает число соседств единиц по горизонтали и по вертикали, т. е. характеристику, инвариантную относительно поворота изображения на $90°$.

К сложностям, возникающим при использовании подобных частотных признаков, следует отнести прежде всего отсутствие четких критериев релевантности и методов выбора признаков.

В практических задачах иногда удается найти эвристические процедуры выбора небольшого числа признаков для распознавания длинных последовательностей. Эти признаки обычно имеют следующий вид. Пусть $a = (a_1, \ldots, a_N)$ — последовательность, а $r_1, r_2, \ldots, r_{k-1}$ — целые неотрицательные числа. Подпоследовательностью $a\ (i;\ r_1, \ldots, r_{k-1})$ с задержками $r_1, \ldots, r_{k-1}$ называется подпоследовательность вида $(a_i, a_{i+r_1}, \ldots, a_{i+r_1+\ldots+r_{k-1}})$, причем если некоторое $r_d = 0$, то элемент $a_{i+r_1+\ldots+r_{d-1}+r_d}$ входит в подпоследовательность только один раз, а не два. Очевидно, $a\ (i;\ r_1, \ldots, r_{k-1}) \in E_{k-\nu}$, где ν — число нулей в наборе $(r_1, \ldots, r_{k-1})$, а E_l — двоичный l-мерный куб. Для каждого вектора $\gamma \in E_{k-\nu}$ символом $t_\gamma\ (a;\ r_1, r_2, \ldots, r_{k-1})$ будет обозначаться *число вхождений вектора* γ *в совокупность* $\{a\ (i;\ r_1, \ldots, r_{k-1})\}$, $i = 1, 2, \ldots, N - (r_1 + \ldots + r_{k-1})$. В качестве признаков для распознавания выбираются величины $(r_1, \ldots$

. . ., r_{k-1}), а в качестве значений этих признаков — величины $t_\gamma(a;\ r_1,\ \ldots,\ r_{k-1})$, $0 < k < N$, $\gamma \in E_{k-\nu}$.

Распознавание может проводиться на основании единственного признака, определяемого следующим образом. Методом последовательного перебора (от меньших задержек к большим) находится такое значение r_1, для которого величина $t_{\alpha\beta}(a^i;\ r_1) - t_{\alpha\beta}(b^j;\ r_1)$, где a^i — элементы класса A, а b^j — элементы класса B, при некоторых α, $\beta \in \{0, 1\}$ принимает значение, превышающее заданный порог. Аналогично определяется величина r_2, для которой достаточно велико значение $t_{\bar\beta\gamma}(b;\ r_2) - t_{\bar\beta\gamma}(a^r;\ r_2)$, $\beta, \gamma \in \{0, 1\}$. В качестве единственного признака для распознавания используется величина $t_{\alpha\beta\gamma}(a;\ r_1,\ r_2)$. Поскольку значение $t_{\beta\gamma}(x;\ r_2)$ тем меньше, чем больше значение $t_{\bar\beta\gamma}(x;\ r_2)$, такое слияние обеспечивает большее отличие величины $t_{\alpha\beta\gamma}(a^i;\ r_1,\ r_2)$ от величины $t_{\alpha\beta\gamma}(b;\ r_1,\ r_2)$, чем отличие каждого из «отдельных» признаков, и «объединенный» признак работает лучше.

Задача восстановления последовательностей по частотам подпоследовательностей с фиксированными задержками оказывается тесно связанной с задачей восстановления слов по фрагментам, возникающей в теории кодирования [25]. В частности, достаточность единственного признака для дихотомии вытекает из того факта, что для любых векторов a, $b \in E_n$, существуют задержка r, $0 \leqslant r \leqslant [N/2]$, и вектор $\gamma \in E_2 \cup E_1$, такие, что $t_\gamma(a;\ r) \neq t_\gamma(b;\ r)$. Этот факт дает возможность в дальнейшем ограничиться случаем подпоследовательностей с одной задержкой. Кроме того, выбор этой задержки можно осуществить таким образом, что для любой тройки векторов a, b, $c \in E_n$ найдется вектор $\gamma \in E_2 \cup E_1$ такой, что

$$|\,t_\gamma(a;\ r) - t_\gamma(b;\ r)\,| \neq |\,t_\gamma(a;\ r) - t_\gamma(c;\ r)\,|. \tag{2.24}$$

Различимость произвольных последовательностей по частотам подпоследовательностей с одной задержкой позволяет применить для распознавания конструкцию алгебраического расширения совокупности алгоритмов вычисления оценок, основанных на использовании этих признаков.

Алгоритм вычисления оценок в этом случае задается следующим образом. Пусть $a_1, \ldots, a_m$ — обучающая выборка двоичных последовательностей длины N: для каждой из этих последовательностей задан информационный вектор, определяющий ее принадлежность к классам $K_1, \ldots, K_l$, причем $\tilde K_u \neq \tilde K_v$ при $u \neq v$, где $\tilde K_u = \{a_i \mid a_i \in K_u\}$, $\tilde K_v = \{a_i \mid \tilde a_i \in K_v\}$. Пусть $a^1, \ldots, a^q$ — распознаваемые последовательности. Для каждой пары a^i, a^k выбирается признак r^{ik} и вектор γ такие, что

$$t_\gamma(a^i,\ r^{ik}) \neq t_\gamma(a^k;\ r^{ik}). \tag{2.25}$$

Такая процедура выполняется для всех пар a^i, a^k. Отобранные признаки r^{ik} нумеруются произвольным образом. Для работы используются описания обучающей и распознаваемой выборок наборами этих признаков:

$$a_i \to S_i = (t_{i1}, \ldots, t_{is}), \qquad a^p \to S^p = (t_1^p, \ldots, t_S^p). \qquad (2.26)$$

Условие (2.24) гарантирует неизоморфность каждой пары описанных таким образом объектов из распознаваемой выборки в смысле [19]. Вместе с условием $\tilde{K}_u \neq \tilde{K}_v$ это обеспечивает регулярность задачи распознавания в смысле [19]. Отсюда следует, что алгебраическое замыкание класса алгоритмов вычисления оценок для задачи распознавания

$$\langle S_1, \ldots, S_m, \tilde{K}_1, \ldots, \tilde{K}_l, S^1, \ldots, S^q \rangle,$$

где векторы S заданы признаками вида $t_\delta(\alpha; r)$, все элементы a^u различны, $a^u \neq a^v$ при $u \neq v$, а для классов $K_1, \ldots, K_l$ выполняются условия $\tilde{K}_u \neq \tilde{K}_v$ при $u \neq v$, является корректным. Предлагаемый метод можно использовать и в случае целочисленных последовательностей, $a_i \in \{0, 1, \ldots, p\}$ без перекодировки их в двоичную форму, поскольку условие (2.24) верно и для p-значных последовательностей. Предлагаемый метод непосредственно обобщается на случай двухмерных массивов. Доказательство утверждения, аналогичного (2.24), практически ничем не отличается. Можно показать, что число признаков в сформулированном выше алгоритме вычисления оценок в худшем случае $q - 1$.

2.7. Основные положения алгебраического подхода

В алгебраическом подходе существенно используются особенности строения, свойственные любой процедуре распознавания. При решении каждой конкретной задачи распознавания рассматривается информация, на базе которой происходит распознавание, перечень классов и объекты $\omega', \ldots, \omega^q$, для которых надо решить, к каким из заданных классов они принадлежат. Любой алгоритм распознавания должен по стандартной информации вида (2.2) и описанию $I(\omega)$ определить информационный вектор этого объекта (2.1), т. е. вычислить значения предикатов $P_i(\omega)$ — «$\omega \in \Omega_i$», $i = 1, \ldots, m$.

Таким образом, любой алгоритм распознавания переводит задачу распознавания z с q распознаваемыми объектами и m классами в матрицу ответов — информационную матрицу, строками которой являются информационные векторы для каждого из распознаваемых объектов ω^j, $j = 1, \ldots, q$.

$$\begin{Vmatrix} \alpha_{11} \ldots \alpha_{1i} \ldots \alpha_{1m} \\ \ldots\ldots\ldots\ldots \\ \alpha_{j1} \ldots \alpha_{ji} \ldots \alpha_{jm} \\ \ldots\ldots\ldots\ldots \\ \alpha_{q1} \ldots \alpha_{qi} \ldots \alpha_{qm} \end{Vmatrix}, \qquad (2.27)$$

где элемент матрицы α_{ji} принимает значения 1, 0 или Δ и указывает, какое именно значение вычислил распознающий алгоритм для свойства $P_i(\omega^j)$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, q$.

Очевидно, что применение разных распознающих алгоритмов к одной и той же задаче z может приводить к получению различных информационных матриц. Поэтому, естественно, возникает проблема построения корректирующих методов, которые позволяли бы, сопоставляя информационные матрицы различных алгоритмов, вырабатывать единую матрицу ответов, причем по возможности с минимальным числом ошибок. Это непростая проблема, поскольку не существует удобных операций над элементами «да», «нет», «не знаю» («1», «0», «Δ»), обладающих естественными свойствами типа ассоциативности, коммутативности и т. д. (доказательство этого факта можно найти в [20]). Поэтому выполнение корректирующих операций над информационными матрицами в некотором смысле просто невозможно. Более тщательное исследование различных моделей распознающих алгоритмов позволяет, однако, уловить специфику организации процесса перевода исходной обучающей информации в информационную матрицу ответов.

Рассмотрим несколько примеров.

В статистических алгоритмах распознавания процесс начинается с формирования матрицы вероятностей $\| P_{ji} \|_{q \times m}$, где P_{ji} — вероятность принадлежности j-го объекта i-му классу; затем по этой матрице вероятностей строится матрица окончательных ответов.

В АВО процесс начинается с построения матрицы оценок (голосов) $\| \Gamma_i(\omega^j) \|_{q \times m}$, где $\Gamma_i(\omega^j)$ — числовая оценка принадлежности j-го объекта i-му классу; затем по этой матрице принимается окончательное решение о принадлежности распознаваемого объекта определенному классу.

Аналогичный анализ можно было бы провести и для других семейств алгоритмов распознавания. Важно в данном случае то обстоятельство, что процесс переработки распознающим алгоритмом исходной информации в информационную матрицу ответов распадается на две последовательные стадии. На первой из них осуществляется перевод исходной информации в числовую матрицу стандартного размера с числом строк, равным числу распознаваемых в задаче z объектов, и числом столбцов, равным числу классов, рассматриваемых при решении z. Вторая стадия работы распознающего алгоритма сводится к переработке этой числовой матрицы в матрицу окончательных ответов с тем же числом строк и столбцов.

Существование разделения процесса переработки информации в распознающем алгоритме на две стадии, продемонстрированное на приведенных примерах, содержательно, строго обосновывается теоремой, утверждающей, что всякий алгоритм распознавания можно представить в виде двух последовательно выполняемых алгоритмов [20]. Первый из них — алгоритм «B» — переводит обучающую информацию и описания распознаваемых объектов в числовую матрицу размерности $q \times m$, а второй — алгоритм «С» — переводит последнюю в матрицу ответов, составленную из символов 1, 0 и Δ и имеющую ту же размерность.

Алгоритм «С» можно в принципе сделать одинаковым для всех распознающих алгоритмов. Это так называемое пороговое решающее правило с положительными порогами:

$$1)\ C\left(\| a_{ji} \|_{q\times m}\right) = \| C(a_{ji}) \|_{q\times m} \tag{2.28}$$

— это означает, что решающее правило применяется к числовой матрице $\| a_{ji} \|_{q\times m}$ поэлементно;

$$2)\ C(a_{ji}) = \begin{cases} 1, & a_{ji} > d_2, \\ 0, & a_{ji} < d_1 \quad 0 < d_1 < d_2, \\ \Delta, & d_1 \leqslant a_{ji} \leqslant d_2; \end{cases} \tag{2.29}$$

здесь d_1 и d_2 — произвольные фиксированные положительные числа.

Итак, вторая часть распознающего алгоритма очень проста и к тому же стандартна для распознающих алгоритмов. Из этого следует, что основная часть процесса переработки информации приходится на первую часть распознающего алгоритма, т. е. на алгоритм «В», который принято называть распознающим оператором, переводит исходную информацию для данной задачи z в числовую матрицу стандартных размеров, т. е.

$$B(I, \Omega^m, \omega^q) = \| a_{ji} \|_{q\times m}. \tag{2.30}$$

Если имеется несколько или целое семейство распознающих алгоритмов, то, исключив из рассмотрения составляющую «С» распознающего алгоритма, мы получаем набор или семейство соответствующих распознающих операторов. Над ними как над отображениями на матрицы стандартных размеров можно определить операции сложения, умножения и умножения на число.

Пусть B_1 и B_2 — некоторые распознающие операторы:

$$\begin{aligned} B_1(I, \Omega^m, \omega^q) &= \| a^1_{ji} \|_{q\times m}, \\ B_2(I, \Omega^m, \omega^q) &= \| a^2_{ji} \|_{q\times m}. \end{aligned} \tag{2.31}$$

Суммой распознающих операторов B_1 и B_2 называется распознающий оператор $B^+ = B_1 + B_2$, определяемый следующим соотношением:

$$B^+(I, \Omega^m, \omega^q) = \| a^1_{ji} + a^2_{ji} \|_{q\times m}. \tag{2.32}$$

Применение распознающего оператора B^+ к задаче z состоит в применении к этой задаче распознающих операторов B_1 и B_2 и поэлементном сложении полученных матриц.

Аналогичным образом определяется произведение $B(\cdot) = B_1 B_2$ распознающих операторов B_1 и B_2:

$$B(\cdot)(I, \Omega^m, \omega^q) = \| a^1_{ji} \cdot a^2_{ji} \|_{q\times m}. \tag{2.33}$$

Отметим, что при определении произведения распознающих операторов используется не стандартное, а поэлементное перемножение соответствующих матриц.

Умножение распознающего оператора B на число определяется как

$$(c \cdot B)(I,\ \Omega^m,\ \omega^q) = \| c \cdot a_{jl} \|_{q \times m}. \tag{2.34}$$

Введенные операции сложения, умножения и умножения на число распознающих операторов обладают всеми свойствами сложения и умножения чисел, а именно коммутативностью, ассоциативностью, дистрибутивностью и т. д. Если задано семейство алгоритмов распознавания $\{A\}$, то, согласно теореме о представлении алгоритма распознавания, можно задать семейства распознающих операторов $\{B\}$ и решающих правил $\{C\}$. Замыкание $L\ \{B\}$ операторов семейства $\{B\}$ операциями (2.32) и (2.34) называют линейным замыканием $\{B\}$, а множество

$$L\ \{A\} = L\ \{B\} \cdot \{C\} \tag{2.35}$$

линейным замыканием семейства алгоритмов $\{A\}$. Операторы линейного замыкания представим через операторы B_g, принадлежащие $\{B\}$, т. е.

$$a_1 B_1 + \ldots + a_g B_g + \ldots + a_k B_k, \tag{2.36}$$

где a_g — действительные числа, $g = 1, \ldots, k$.

Замыкание $\mathfrak{N}\ \{B\}$ операторов семейства $\{B\}$ относительно операций (2.32)—(2.34) называют алгебраическим замыканием $\{B\}$, а множество

$$\mathfrak{N}\ \{A\} = \mathfrak{N}\ \{B\} \cdot \{C\} \tag{2.37}$$

алгебраическим замыканием семейства алгоритмов $\{A\}$. В таком случае можно перейти к записи операторных многочленов, т. е. выражений вида

$$\sum C_{g_1} \ldots C_{g_k} \cdot B_{g_1} \ldots B_{g_k}, \tag{2.38}$$

где $C_{g_1}, \ldots, C_{g_k}$ — константы, $B_{g_1}, \ldots, B_{g_k}$ — распознающие операторы исходных распознающих алгоритмов.

Совокупность операторов из $\mathfrak{N}\ \{B\}$, представимых такими многочленами степени не выше k, называют алгебраическим замыканием $\mathfrak{N}_k\ \{B\}$ множества $\{B\}$, $k = 2, 3, \ldots, v$, а множество

$$\mathfrak{N}_k\ \{A\} = \mathfrak{N}_k\ \{B\} \cdot \{C\} \tag{2.39}$$

алгебраическим замыканием степени k семейства $\{A\}$ алгоритмов распознавания.

Введение операций над распознающими операторами обеспечивает возможность расширять исходную совокупность распознающих операторов, а следовательно, и распознающих алгоритмов. Действительно, если $L\ (B_1, \ldots, B_k)$ — некоторый операторный многочлен, то выражение $L(B_1, \ldots, B_k)\, C(d_1, d_2)$ определяет уже некоторый распознающий алгоритм. В данном случае символ произведения означает последовательное выполнение соответствующих частей алгоритма, $C(d_1, d_2)$ — пороговое решающее правило с порогами d_1 и d_2.

Расширенные таким образом распознающие алгоритмы обладают очень сильными корректирующими свойствами. Доказано [20], что даже в тех случаях, когда в исходном семействе распознающих алгоритмов отсутствует алгоритм, правильно решающий данную задачу распознавания z, то при выполнении просто проверяемых допущений относительно исходной обучающей информации и описания распознаваемых объектов искомый правильно решающий задачу z алгоритм существует в расширении, причем он может быть выписан в явном виде.

Для иллюстрации рассмотрим в качестве исходного множества некорректных распознающих процедур (эвристик) совокупность АВО. Как уже отмечалось, алгоритм этого класса задается параметрами $\varepsilon_1, \ldots, \varepsilon_N$ (порог точности для признаков), $p_1, \ldots, p_n$ (веса признаков), $\gamma_1, \ldots, \gamma_{r_m}$ (веса объектов в таблице обучающей информации), l (число признаков в опорных множествах алгоритма). Последний параметр используется при задании лишь некоторых разновидностей АВО.

Рассмотрим задачу распознавания z с обучающими объектами $\omega_1, \ldots, \omega_{r_m}$, $\omega_r = (a_{r_1}, \ldots, a_{r_n})$, $r = 1, \ldots, r_m$, классами $\Omega_1, \ldots, \Omega_m$, которые могут быть пересекающимися; решение z предполагает распознавание объектов $\omega^1, \ldots, \omega^q$, $\omega^j = (b^{1N}, \ldots, b^{qN})$, $j = 1, \ldots, q$. Пусть также в множестве возможных значений признака x_t, $t = 1, \ldots, N$, введен способ измерения расстояния $\rho_t(y, z)$.

Установлено, что при выполнении достаточно естественных условий в алгебраическом расширении может быть явно построен распознающий алгоритм, правильно решающий заданную z. Распознающий оператор этого распознающего алгоритма имеет следующий вид:

$$R_A = (d_1 + d_2) \sum_{j=1}^{q} \sum_{i=1}^{m} C_{ji} B_{ji}^{k_{ji}}. \tag{2.40}$$

В этом выражении для распознающего алгоритма все обозначения сохраняют свой смысл, а k_{ji} — степень операторного многочлена (порядок алгебраического замыкания). В качестве распознающих операторов B_{ji} выступают фиксированные распознающие операторы из класса АВО, следовательно, каждый из распознающих операторов полностью описан, т. е. заданы все параметры $\varepsilon(j, i)$, $p(j, i)$, $\gamma(j, i)$ и $k(j, i)$. Значения параметров вычисляются с помощью специальных процедур по обучающей информации I и описаниям распознаваемых объектов. Каждый распознающий оператор B_{ji} задается, таким образом, набором из $2N + r_m + 1$ чисел. Величина k_{ji} также вычисляется по обучающей информации и описаниям распознаваемых объектов, но во всех случаях [22]

$$kj_i \leqslant \{(\ln q + \ln m + | \ln (d_1 + d_2) | - | \ln d_1 |) [| \ln (1 - (q + m - 2 - \eta)^{-1}) |]^{-1}\} + 1, \qquad (2.41)$$

где η — сколь угодно малое число.

Показатель степени k_{ji} можно понизить до $q - 1$, но в этом случае исчезает возможность простой записи распознающих операторов.

Константы C_{ji} не всегда удается определять только на основе обучающей информации. В таком случае вся исходная информация делится на две части: объекты, включенные в первую часть, рассматриваются как материал обучения, остальные объекты образуют распознаваемый массив $\omega^1, \ldots, \omega^q$. На самом же деле для последних объектов известна принадлежность к заданным классам. В этом случае считается, что $C_{ji} = 1$, если $\omega^j \in \Omega_i$, и $C_{ji} = 0$, если $\omega^j \in \Omega_i$.

В алгебраической теории распознающих алгоритмов доказывается, что распознающий оператор типа R_A сохраняет классификацию, если объекты $\omega', \ldots, \omega^q$ заменяются объектами из некоторой их окрестности. Это означает, что алгоритм с распознающим оператором вида R_A имеет ненулевой радиус устойчивости классификации. Отсюда несложно выводится теорема о том, что если границы классов достаточно гладки и число q достаточно велико, то алгоритм с распознающим оператором вида R_A дает правильный ответ почти всегда, т. е. с вероятностью, близкой к 1. Этот факт имеет место в принципе и при более общих предположениях о строении классов. Он верен, если выполнена так называемая гипотеза компактности. В том случае, когда никакой информации о строении классов нет и содержательные соображения не дают оснований полагать, что требования гипотезы компактности выполнены, с помощью аппарата, предложенного В. Н. Вапником и А. Я. Червоненкисом [5], можно доказать, что распознающий алгоритм с распознающим оператором типа R_A дает правильный ответ на всей совокупности допустимых объектов распознавания с некоторой гарантированной вероятностью. Эта вероятность тем ближе к 1, чем больше величина q — число объектов, зачисляемых во вторую часть разбиения обучающего массива. Эту часть обучающего массива часто называют контрольной.

При разбиении совокупности объектов с известной классификацией на две части контрольную группу целесообразно делать по возможности большей. При этом необходимо лишь обеспечить выполнение двух условий: а) на обучающей информации классы должны различаться; б) для всякой пары объектов ω^u, ω^v из контрольного массива должен существовать хотя бы один объект ω_r из обучающего массива и хотя бы один признак x_t, $r = r(u, v)$, $t = t(u, v)$, такие, что

$$\rho_t(a_{rt}, b_{ut}) \neq \rho_t(a_{rt}, b_{vt}). \qquad (2.42)$$

Другими словами, это означает, что для любых объектов из контрольного массива в обучающем массиве найдутся объект и признак, такие, что некий способ измерения расстояния ρ_t в множестве значений x_i-го признака дает различные значения для расстояний от найденного обучающего объекта до первого и второго контрольных.

Резюмируем основные особенности и результаты алгебраического подхода к построению общей теории распознающих алгоритмов. Итак, на основе изучения различных моделей алгоритмов оказалось возможным сформулировать общее определение распознающего или классифицирующего алгоритма и изучить свойства множества таких алгоритмов. Оказалось, что множество распознающих алгоритмов является алгеброй, причем операции этой алгебры обладают набором свойств, позволяющих детально изучить множество распознающих алгоритмов. С помощью соответствующих алгебраических методов решены задачи нахождения базисов в совокупности алгоритмов распознавания, выяснены условия, при которых модель содержит алгоритм, который абсолютно точно классифицирует любую заданную конечную выборку объектов, и т. д. Алгебраические методы позволяют также эффективно решать задачу выбора экстремального алгоритма уже не в рамках отдельной модели, а во всем множестве распознающих алгоритмов.

В рамках алгебраического подхода показано, что если набор нестрогих алгоритмов, предназначенных для решения задач определенного вида, имеет систему сравнительно просто проверяемых свойств, то он может быть пополнен при помощи формальных процедур до системы алгоритмов, обладающих следующим свойством: в пополненной системе алгоритмов, построенных на базе «интуитивных» алгоритмов, существует алгоритм, дающий точное решение каждой задачи данного вида.

Таким образом, «интуитивные алгоритмы», семейства которых обычно строятся при решении плохо формализованных проблем пополняются до множества строгих алгоритмов.

Введение дополнительных ограничений на классы решаемых задач и множества эвристических алгоритмов позволяет не только доказать указанные теоремы существования, но и указать методы поиска в пополнениях таких корректных алгоритмов.

Следовательно, вместо построения формальных моделей в различных областях, плохо поддающихся формализации, возможен принципиально иной путь. Достаточно построить семейство «интуитивно разумных» алгоритмов для решения соответствующих задач, затем ввести алгебру на множестве таких задач и построить алгебраическое замыкание «интуитивного» семейства алгоритмов. В этом замыкании оказывается принципиально разрешимой любая задача из множества задач, связанных с исследованием плохо формализованных ситуаций. Естественно, что и сама совокупность задач должна быть формально описана: необходимо задание множества исходной информации и множества вопросов (предикатов), на которые следует ответить в процессе решения задачи.

Основной результат алгебраической теории распознающих алгоритмов состоит в том, что для описания класса алгоритмов, правильно классифицирующих конечную выборку по всем классам, достаточно взять любую полную модель, рассмотреть линейное замыкание совокупности ее распознающих операторов и присоединить к ней любое корректное решающее правило. Такой алгоритм окажется оптимальным по любому функционалу качества распознающего алгоритма.

Модель распознающих алгоритмов является полной, если в ее линейном замыкании существуют такие распознающие операторы, что матрицы, полученные из начальной информации в результате их применения, образуют базис в пространстве числовых матриц размерности $q \times m$. Решающее правило является корректным, если для любой конечной совокупности допустимых объектов существует матрица числовых оценок (получаемая в результате применения распознающего оператора) такая, что применение к ней этого решающего правила переводит ее в истинную информацию матрицу окончательных ответов.

Известен ряд результатов, связанных с проверкой полноты широко применяемых моделей алгоритмов.

Показано, что класс алгоритмов над стандартной начальной информацией, определяемых кусочно-линейными разделяющими поверхностями и весами объектов в обучающей выборке, является полным [20]. Там же показано, что класс АВО с пороговыми функциями близости, определяемый заданием порогов точности для признаков, весов признаков и весов объектов в обучающей выборке, является корректным над начальной информацией, если на последнюю наложены некоторые ограничения и класс распознающих операторов пополнен оператором, выделяющим максимальные элементы числовых матриц.

Эти доказательства основаны на применении общих критериев, позволяющих устанавливать корректность соответствующих замыканий. Сами замыкания, однако, содержат бесконечные множества алгоритмов. Теоремы существования, естественно, сами по себе не дают возможности реально построить алгоритм, корректный для данной задачи, если ее информационная матрица неизвестна. В связи с этим была показана возможность эффективного выделения в замыкании конечного множества алгоритмов, среди которых находится алгоритм, корректный для данной задачи (информационная матрица задачи при этом неизвестна). Аналогичное построение выполняется и при наличии сведений об информационной матрице [19].

Кроме того, доказано, что для каждой задачи z, задаваемой исходной информацией I и набором распознаваемых объектов ω^q, существует окрестность в пространстве задач с центром в (I, ω^q), для которой корректный алгоритм не меняется при сохранении информационной матрицы. Продемонстрирован способ вычисления радиуса этой окрестности и установлено, следовательно, что корректные алгоритмы в замыканиях устойчивы [19].

Если на заданных классах гипотеза компактности оказывается справедливой, то может быть указан способ построения корректных алгоритмов. Если же в любой окрестности задачи существуют задачи с различными информационными матрицами, то построение одного корректного алгоритма для задач из окрестности принципиально невозможно. В этом случае можно выделять минимальные по мощности конечные множества алгоритмов, корректные для задач из окрестности. Способ построения таких множеств описан в работе [19].

Большая часть исследований в рамках алгебраического подхода выполнена в расширениях модели АВО, однако совершенно аналогично они могут быть проведены для любых моделей, расширения которых являются полными и, следовательно, корректными.

Принципиальную схему класса алгоритмов распознавания можно описать следующим образом [18]:

1) формирование моделей распознающих алгоритмов;

2) выбор экстремального алгоритма и модели;

3) разбиение экстремальных алгоритмов на классы; построение в каждом классе линейного замыкания; выбор алгоритма, оптимального по коэффициентам линейной комбинации и параметрам решающего правила;

4) выработка оптимальной корректирующей операции и применение ее для корректировки оптимальных алгоритмов, полученных на этапе 3.

Подробное изложение и доказательство результатов, полученных в рамках алгебраического подхода, можно найти в работах [19, 20].

Глава 3
РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ

3.1. Общая характеристика проблемы

В последние десятилетия, особенно начиная с 70-х годов, среди прикладных задач обработки и преобразования информации все более существенное место начинают занимать задачи, в которых исходная информация представляется в виде изображений. Этот процесс отражает появление новых технических средств сбора и воспроизведения информации, обеспечивающих эффективное и наглядное представление зарегистрированных и накопленных данных в виде изображений и рост известности и популярности распознавания как новой информационной технологии — мощной, практичной и в некотором смысле универсальной методологии математической обработки и оценивания информации и выявления скрытых закономерностей [11].

Проблема распознавания изображений все еще имеет довольно изменчивый облик: она может возникать в виде собственно задач распознавания, задач анализа сцен, задач понимания изображений, а также в виде так называемой проблемы машинного зрения. Объектами распознавания /анализа/ понимания могут служить изображения, полученные в различных частях полного спектра излучений (оптические, инфракрасные, ультразвуковые и т. д.) различными способами (телевизионные, фотографические, лазерные, радиолокационные, радиационные и т. д.), преобразованные в цифровую форму и представленные в виде некоторой целочисленной матрицы.

Роль изображения как объекта информационной технологии определяется тем, что изображение является специфическим видом информации, в котором объединены и смешаны исходная (отображаемая) информация и форма ее представления, информационная и физическая модели отображаемых объектов, явлений, процессов. К сожалению, эта специфика изображения еще не понята и не изучена ни в информатике, ни в исследованиях, посвященных изучению зрительного восприятия человека и животных (психология, психофизика, нейрофизиология).

Изображения обладают информационной емкостью, компактностью и наглядностью, а зрение является наиболее естественным для человека механизмом восприятия информации о внешнем мире и из внешнего мира. Это к тому же наиболее древний в эволюционном смысле способ представления и восприятия информации. В процессе восприятия изображения человек, судя по всему, не строит словесное описание изображения, но оперирует с ним как с неким целостным образом или системой таких образов, прибегая к неязыковому внутреннему представлению. При разработке методов и систем автоматизированного распознавания изображений приходится отыскивать способы эффективной формализации изображений для того, чтобы иметь возможность работать с представлениями (описаниями), отражающими семантику изображения, информацию, заключенную в его внутренней структуре и структуре внешних связей части реального мира (сцены), воспроизводимой с помощью изображения.

Специфика, сложность и следующая из них трудность задач распознавания изображений определяется необходимостью достижения компромисса между весьма противоречивыми факторами, отражающими требования к анализу, природу зрительного восприятия, способы получения, формирования и воспроизведения изображения и существующие математические и технические возможности работы с ними. Основным, очевидно, является противоречие между природой изображения и анализом, основанным на использовании формального аппарата (в сущности, модели) объекта. Оно выражается в том, что для использования преимуществ представления информации в виде изображения необходимо придать этой информации «неизобразительный» вид, ибо соответствующие алгоритмы приспособлены для переработки лишь неких символических описаний. Изображения по природе своей — прекрасный объект для применения параллельных методов обработки, однако большинство известных методов распознавания имеет по, следовательный характер, в частности, потому, что организация процедуры распознавания предусматривает учет результатов, получаемых на промежуточных этапах. Подавляющее большинство методов работы с изображениями являются чисто эвристическими, и достоинства их, в сущности, определяются тем, сколь успешно им удается преодолевать «изобразительность» характера изображения «неизобразительными» средствами, т. е. опираться на процедуры, не зависящие от организации обрабатываемой информации в виде изображения.

Анализ современного состояния проблемы распознавания изображений приводит к следующим выводам:

1) математическая теория распознавания изображений до сих пор не создана, и, к сожалению, специфика изображений не позволяет непосредственно воспользоваться методами и средствами классической («одномерной») теории распознавания образов и цифровой обработки сигналов;

2) отсутствие математической теории распознавания изображений препятствует обоснованным и систематизированным разработке, выбору, сопоставлению и применению алгоритмов распознавания изображений, а также получению осбоснованной и надежной оценки их эффективности и адекватности;

3) задачу распознавания изображений следует ставить, исследовать и решать как математическую задачу;

4) разработка информационной теории распознавания изображений предполагает:

а) математическую постановку, характеризацию и систематизацию задач распознавания изображений;

б) разработку, исследование, характеризацию и систематизацию методов и средств построения моделей изображений, ориентированных на задачу распознавания;

в) разработку, исследование, характеризацию и систематизацию преобразований, обеспечивающих приведение изображения к виду, удобному для распознавания;

г) разработку формальных конструкций для описания моделей алгоритмов распознавания изображений и использование последних в качестве базы знаний для определения и характеризации классов алгоритмов распознавания изображений;

д) стандартизацию моделей для отдельных классов изображений применительно к задаче распознавания;

е) разработку систематических методов автоматического выбора оптимальной модели алгоритма распознавания для конкретного изображения;

ж) реализацию перечисленных выше методов и моделей в программном обеспечении вычислительных систем, предназначенных для разработки, исследования и оценки алгоритмов распознавания изображений.

3.2. Типы задач распознавания изображений

При работе с исходной информацией, задаваемой в виде изображений, возникают следующие задачи распознавания:

1) сопоставление двух изображений в целом для установления их принадлежности к одному классу (определяется, представляют ли изображения один и тот же объект или сцену);

2) сопоставление изображения в целом с набором или серией последовательных (по времени) изображений, представляющих некоторый класс изображений (т. е. объектов или сцен) (цель — та же, что и в задаче 1);

3) задачи 1 и 2 для случая нескольких классов;

4) поиск на предъявленном на распознавание изображении некоторой регулярности/нерегулярности (объекта, ситуации), на которую следует обратить внимание, хотя она и не задавалась в априорном перечне эталонов (ассоциативный поиск; ограниченно детерминированный набор классов — задачи логической и семантической фильтрации в сочетании с самообучением);

5) поиск на предъявленном на распознавание изображении регулярности/нерегулярности/фрагмента заданного вида;

6) разбиение множества изображений на непересекающиеся подмножества (задача автоматической классификации);

7) решение задачи автоматической классификации на одном изображении (разбиение изображения на однородные области, группы объектов, сегментация области, выделение признаков объектов);

8) совместное решение задач *6* и *7*;

9) автоматическое выделение непроизводных элементов, характерных объектов изображения, признаков-объектов, пространственных и логических отношений для синтеза формализованных представлений и описаний изображения;

10) приведение изображения к виду, удобному для распознавания; автоматический синтез формализованных представлений и описаний изображения;

11) решение задач *9* и *10* в диалоговом режиме;

12) задачи восстановления

— пропущенных кадров в последовательности изображений;
— изображений в целом по их фрагментам;
— фрагментов изображения (и объектов) на основе непроизводных элементов, признаков и порождающих процедур с учетом контекста изображения в целом;
— траектории задачи по ее фрагментам и неизвестных фрагментов траекторий;

13) выбор и формирование траектории задачи распознавания изображений (в смысле задачи распознавания со стандартной обучающей информацией);

14) решение задач *1—13* в случаях наличия на изображении динамических объектов, сложной фоновой обстановки (в том числе динамических и статических помех) и с учетом способов получения, формирования и представления изображений.

3.3. Математическая постановка задачи распознавания изображений

При анализе и распознавании изображений обрабатываемая информация представляется числовой матрицей, воспроизводящей свойства изображаемого объекта (сцены) и деформации, связанные со способом и процессом получения изображений. Для формализации процесса распознавания изображений определим три множества (моделей) изображений, на которых постулируется существование классов эквивалентности, и множества допустимых преобразований, заданных на классах эквивалентности. Введение классов эквивалентности на множестве моделей изображений отражает гипотезу, согласно которой всякое изображение обладает определенной регулярностью или смесью регулярностей различных типов. Задачу распознавания при этом предположении можно свести к разделению изображений, сохраняющих собственную регулярность, и изображений с нарушениями собственной регулярности (могут, естественно, ставиться задачи обнаружения на изображении регулярности или нарушения регулярности определенных типов).

Рассмотрим следующую модель. Пусть I — некоторое истинное изображение изучаемого объекта. Процесс получения, формирования, дискретизации и т. д. (все процедуры, необходимые для того, чтобы с изображением можно было работать) можно рассматривать как передачу истинного изображения по каналу с помехами. В результате предметом анализа служит не истинное, а некоторое реальное — наблюдаемое — изображение I^*. В процессе анализа последнее должно быть классифицировано, т. е. должен быть определен его прототип в истинном классе эквивалентности K_l, либо на наблюдаемом изображении I^* следует обнаружить регулярность (регулярности) заданного вида J^R.

Таким образом, можно определить множества $\{I\}$, $\{I^*\}$ и $\{I^R\}$, $\{I\} = \bigcup\limits_l K_l$ и преобразования формирования $\{T^F\}$ и распознавания $\{T^R\}$ изображений:

$$T^F\colon I \Rightarrow J^*, \tag{3.1}$$

$$T^R\colon I^* \Rightarrow J^R. \tag{3.2}$$

Итак, распознавание изображений сводится к определению на классах эквивалентности множества $\{I\}$ алгебраических систем преобразований $\{T^F\}$ и $\{T^R\}$, применению их к наблюдаемым изображениям I^* для: а) анализа «назад» — разделения изображений в соответствии с характером регулярности (восстановление истинных изображений, т. е. указание классов эквивалентности, к которым они относятся); б) анализа «вперед» — поиска на изображении I^* регулярностей определенного вида I^R и их локализации.

Такая постановка задачи распознавания позволяет определить класс процедур обработки изображений, характеризующийся фикспрованной структурой процесса, интерпретация (конкретная реализация) которой зависит от целей и типа анализа. В процессе распознавания выделяются следующие основные этапы:

1°. Синтез модели наблюдаемого изображения $\mathfrak{R}\{I^*\}$ — это так называемый этап приведения к виду, удобному для распознавания, т. е. получение некоторого формализованного описания анализируемого изображения, пригодного для обработки его соответствующими преобразованиями — алгоритмическими процедурами распознавания.

2°. Логическая фильтрация изображения. На этом этапе производится предобработка наблюдаемого изображения, обеспечивающая его предварительную классификацию, которая необходима для выбора множества преобразований $\{T^F\}^{-1}$. Предполагается существование соответствия между типом и/или характером модели $\mathfrak{R}(I^*)$ и классом эквивалентности, определенным на множестве $\{T^F\}$. Предполагается, кроме того, и наличие слабой эквивалентности на множестве $\{T^F\}$, что дает возможность сопоставлять в смысле такой слабой эквивалентности подмножества $\{T^F\}$ классам эквивалентности моделей истинных изображений $K^i[\mathfrak{R}(I)]$.

3°. Установление класса эквивалентности истинного изображения $K_i[\mathfrak{R}(I^*)]$, порождающего данное наблюдаемое изображение I^*. Для этого к модели $\mathfrak{R}(I^*)$ применяются обратные преобразования формирования $\{T^F\}^{-1}$. Кроме того, на основе анализа этапа 2° выдвигается гипотеза о классе эквивалентности, истинном для $\mathfrak{R}(I^*)$. Это дает возможность применять к модели истинного изображения-прототипа $\mathfrak{R}(I^*)$ преобразования T^F для проверки допустимости порождения рассматриваемого I^* в соответствующем классе эквивалентности и сопоставлять результаты применения $(T^F)^{-1}$ к $\mathfrak{R}(I^*)$ и T^F к $K_i[\mathfrak{R}(I^*)]$. И те и другие преобразования могут в соответствии с методологией алгебраического подхода применяться в форме линейных и алгебраических замыканий соответствующих преобразований:

$$L\{T^F\}^{-1}:\mathfrak{R}(I^*)\Rightarrow K_i[\mathfrak{R}(I^*)]', \tag{3.3}$$

$$L\{T^F\}:K_i[\mathfrak{R}(I^*)]\Rightarrow \mathfrak{R}(I^*)'. \tag{3.4}$$

При достижении эквивалентности $K_i[\mathfrak{R}(I^*)]'$ и $K_i[\mathfrak{R}(I^*)]$, $\mathfrak{R}(I^*)$ и $\mathfrak{R}(I^*)'$ или эквивалентности промежуточных результатов преобразований процессов анализа «вперед» и «назад» процесс прекращается. Этот механизм восстановления класса эквивалентности, истинного для $\mathfrak{R}(I^*)$, называется процедурой реверсивного алгебраического замыкания [8].

4°. По $K_i[\mathfrak{R}(I^*)]$ выбираются преобразования распознавания T^R, поскольку предполагается также и соответствие подмножеств преобразований $\{T^R\}$ и $K_i[\mathfrak{R}(I)]$:

$$K_i:\ \{T^R\}\Rightarrow T^R(K_i). \tag{3.5}$$

5°. «Распознавание»: выявление на I^* искомых регулярностей

I^R посредством применения к $\mathfrak{R}(I^*)$ анализа «вперед» и одновременного применения к $\mathfrak{R}(I^R)$ преобразований $\{T^R\}^{-1}$, т. е. анализа назад. В данном случае, как на этапе 3°, применяется процедура реверсивного алгебраического замыкания, но на этот раз для установления того, может ли искомая регулярность $\mathfrak{R}(I^R)$ порождаться моделью наблюдаемого изображения $\mathfrak{R}(I^*)$:

$$L[T^R(K_i)]:\mathfrak{R}(I^*)\Rightarrow\mathfrak{R}(I^R), \tag{3.6}$$

$$L[T^R(K_i)]^{-1}:\mathfrak{R}(I^R)\Rightarrow\mathfrak{R}(I^*)'. \tag{3.7}$$

По достижении эквивалентности $\mathfrak{R}(I^R)$ и $\mathfrak{R}(I^R)'$, $\mathfrak{R}(I^*)$ и $\mathfrak{R}(I^*)'$ или эквивалентности промежуточных результатов преобразований процессов анализа «вперед» и «назад» процесс прекращается. При отсутствии эквивалентности выполняются новые итерации для этапов 1°—5° с другими $\mathfrak{R}(I^*)$ и гипотезами о $K_i[\mathfrak{R}(I^*)]$.

В рамках данной постановки распознавание сводится к определению на классах эквивалентности множества $\{I\}$ алгебраических систем преобразований $\{T^F\}$ и $\{T^R\}$ и применению их к наблюдаемым изображениям $\{I^*\}$ в соответствии с методом реверсивного алгебраического замыкания для: а) анализа «назад» — разделения изображений в соответствии с характером их регулярности («восстановление» истинных изображений — указание классов эквивалентности, к которым они относятся); б) анализа «вперед» — поиска на изображениях I^* регулярностей определенного вида I^R и их локализации.

Итак, математическая постановка задачи распознавания изображений имеет следующий вид:

а) дано

1° $\{I\}$ — множество идеальных изображений;
2° $\{I^*\}$ — множество наблюдаемых изображений;
3° $\{I^R\}$ — множество изображений — результатов реализации процесса распознавания (множество решений);
4° $\{T^F\}$ — множество допустимых преобразований формирования изображения;
5° $\{T^R\}$ — множество допустимых преобразований распознавания изображения;

б) пусть

1° $\{I\}=\bigcup\limits_i K_i,\quad K_i,\quad i=1,\ldots,l;$

2° $\exists\mathfrak{R}(I)\in K_i(\mathfrak{R}(I))\,|\,\forall I\in K_i;$

3° $K_i(\mathfrak{R}(I))=K_i(\mathfrak{R}(I^*))\,|\,t^F:I\Rightarrow I^*,\quad t^F\in\{T_i^F\};$

4° $\langle K_i(\mathfrak{R}(I))\rangle\underset{R}{\Leftrightarrow}(\{T_i^F\}\in\{T^F\},\quad \{T_i^F\}^{-1}\in\{T^F\}^{-1});$

5° $\langle K_i(\mathfrak{R}(I))\rangle\underset{R}{\Leftrightarrow}(\{T_i^R\}\in\{T^R\},\quad \{T_i^R\}^{-1}\in\{T^R\}^{-1};$

6° $K_i(\mathfrak{R}(I))\underset{R_M}{\Leftrightarrow}K_i(\mathfrak{R}(I^R)),\quad t^R:I^*\Rightarrow I^R,\ t^R\in\{T_i^R\},$

где K_l — класс эквивалентности изображений; $\mathfrak{R}(I)$ — формальное описание изображения; R_M — морфизм; R — отношение соответствия;

в) требуется: вычислить значение предиката

$$P(\mathfrak{R}'(I^R) \underset{l}{\in} K_l(\mathfrak{R}(I^R)) \mid \{t^R\} : \mathfrak{R}(I^*)$$

$$\{t^R\} \in \{T_i^R\}, \quad i = 1, \ldots, l; \tag{3.7}$$

г) структура решения задачи распознавания:

1° $\mathfrak{R}(I^*)$;

2° $P((\mathfrak{R}(I^*)) \in K_l(\mathfrak{R}(I^*))) = ?$

3° $L(t_i^{F-1}) : \mathfrak{R}(I^*) \underset{K_l}{\Rightarrow} \mathfrak{R}'(I^*)$; $\quad L(t_i^F) : \mathfrak{R}(I) \Rightarrow \mathfrak{R}'(I)$;

4° $P_{K_l}(\mathfrak{R}'(I^*) \underset{R_{K_l}}{\Leftrightarrow} \mathfrak{R}'(I)) = ?$ $\quad (P_{K_l} = 1) \Rightarrow 5°$,

$(P_{K_l} = 0) \Rightarrow (2° \div 4°)$,

5° $L(t_i^{R-1}) : \mathfrak{R}(I^R) \underset{K_l}{\Rightarrow} \mathfrak{R}'(I^R)$, $\quad L(T_i^R) : \mathfrak{R}(I^*) \underset{K_l}{\Rightarrow} \mathfrak{R}''(I^*)$;

6° $P_R(\mathfrak{R}'(I^R) \underset{R_{K_l}}{\Leftrightarrow} \mathfrak{R}''(I^*)) = ?$ $\quad (P_R = 1) \Rightarrow \mathrm{stop}$;

$(P_R = 0) \Rightarrow (2° \div 5°)$,

где L — линейное (алгебраическое) замыкание; R_{K_l} — отношение эквивалентности, заданное на классе K_l.

3.4. Дескриптивная теория распознавания изображений

Принципиальные особенности этой теории определяются ее целями — созданием регулярных методов для выбора и синтеза алгоритмических процедур обработки информации в задачах распознавания изображений, специфическими особенностями постановки и решения задачи распознавания при задании исходной информации в виде изображения и структурой модели алгоритма распознавания изображений. Остановимся на основных положениях дескриптивной теории.

1. Принцип порождения. В процесс распознавания включается информация, отражающая процесс формирования зрительного образа. Выяснение строения (структуры) образа сводится к выяснению того, какие подобразы можно выделять в целостном образе, насколько они могут или должны быть элементарны и в каких отношениях пребывают эти элементы. Основная задача при этом — изучение и использование структур взаимоотношений элементов, составляющих образ. В результате описание сложного объекта на изображении строится как иерархическая структура, образованная более простыми объектами, т. е. возникает возможность в явном виде использовать и представлять иерархическую (древовидную) структурную информацию, которую несет изображение. Применительно к распознаванию такой метод задания изображений может реализовываться как обнаружение и обработка регулярных структур. Представление изображения иерархической структурой естественно приводит к комбинаторным регулярным структурам. Обращение к ним позволяет, оперируя весьма ограниченным количеством атомарных («непроизводных») элементов и ограниченным количеством правил комбинирования, при помощи неограниченного (например, рекуррентного) применения последних к исходным элементам и результатам реализации соответствующих рекуррентных процедур получать практически неограниченное разнообразие описаний. Использование комбинаторной регулярности в качестве механизма описания структуры изображений обеспечивает чрезвычайно экономное расходование средств описания.

Подобный способ задания («индуктивного порождения») некоторого класса объектов носит в математике название обобщенного индуктивного определения. Оно осуществляется по следующей схеме: 1) задаются некоторые (исходные) объекты определяемого класса; задаются некоторые правила, позволяющие из уже определенных объектов получать другие объекты определяемого класса; 3) объектами определяемого класса являются только те, которые построены в соответствии с пп. *1* и *2* этого определения.

Смысл принципа порождения состоит, таким образом, в том, что формализованное описание изображения при распознавании задается как некоторая связанная между собой структурными отношениями система объектов, выделяемых на изображении при помощи системы преобразований и задаваемых преобразованиями, указывающими допустимый способ их построения. Отметим, что эти преобразования («порождающие процедуры») обладают функциональной полнотой относительно соответствующего класса эквивалентности, индуцированного на множестве идеальных изображений.

2. Формализованное описание. В процессе распознавания изображений в качестве исходной информации используется формализованное описание изображения — модель изображения. Все преобразования, применяемые к формализованному описанию изображения, вводятся ради достижения одной из следующих трех целей: а) получение нового формализованного описания изображения; б) приведение изображения к виду, удобному для распознавания; в) получение агрегированной оценки формализованного описания, т. е. переход от пространства исходных информаций к пространству оценок; на последнем обычно реализуются процессы принятия классификационных решений при распознавании.

3. Специфика формализованного представления изображения. Под формализованным представлением понимается некоторая формальная схема, предназначенная для получения в явном виде объектов изображения и порождающих процедур, т. е. стандартизированное формальное описание форм и поверхностей, образующих изображение. Под формализованным описанием понимается конкретная реализация формализованного представления. В качестве исходной информации при синтезе формализованных представлений изображения можно использовать значения яркостей, определяемых геометрическими свойствами — пространственной организацией и отражательной способностью видимых поверхностей, освещением сцены и позицией наблюдателя. Система представлений, используемых при распознавании изображения, должна включать характерные объекты изображения, которым можно ставить в соответствие признаки, указывающие значения таких переменных, как яркость, ориентация, размеры, местоположение; характерные объекты должны соответствовать реальным физическим «особенностям» поверхностей. Это означает, что структура и свойства реального мира — объекта изображения, воспроизводимые при помощи физических ограничений, играют центральную роль в процессах получения информации о поверхности.

Следующие четыре аспекта определяют специфику формализованного представления изображения.

а. Масштабная и морфологическая многоуровневость. При распознавании изображений используется система представлений, включающая несколько слоев формализованных описаний изображения, каждый из которых соответствует своему масштабному или морфологическому уровню. Разделение на морфологические уровни определяется типами и сложностью характерных объектов и непроизводных элементов, задаваемых и порождаемых на каждом уровне. Разделение на масштабные уровни определяется масштабом характерных объектов и непроизводных элементов, задаваемых и порождаемых на каждом уровне, а также соответствующими масштабными преобразованиями. Многоуровневость представления обеспечивает регуляризацию выбора системы порождающих преобразований и создает информационную избыточность, компенсирующую «частичность» и информационную неполноту формализованных описаний, соответствующих отдельным уровням. Отметим, что одним из вариантов задания морфологических уровней служат классы Павлидиса [11].

б. Синтаксическая и реляционная информация. В силу принципа порождения информация, характеризующая синтаксическую структуру изображения, существенно используется при синтезе представлений. Реляционная информация играет заметную роль при задании представления в виде некоторой формальной конструкции, поскольку обеспечивает возможность определения связей между объектами представления. Заметим, что доминирующая роль синтаксической и структурной информации в представлениях изображений в неявном виде отражалась в широком использовании структурных методов при анализе изображений, особенно на ранних этапах эволюции анализа изображений.

в. П р и з н а к и. Они используются при распознавании изображений в двух качествах: как непроизводные элементы, сборка которых посредством порождающих преобразований обеспечивает получение характерных объектов представления; как характеристики, которые ставятся в соответствие характерным объектам для того, чтобы отразить и зафиксировать типы изменений переменных, снимаемых с изображения и полезных, например, для определения изменений ориентации видимой поверхности относительно наблюдателя и расстояний до него. Признаки обеспечивают возможность использования локальной и глобальной информации, содержащейся в изображении, и определяют соотношение используемой при формировании описания глобальной и локальной информации.

При работе с изображениями возникают новые типы признаков (относительно классических моделей распознавания со стандартной информацией), позволяющие воспроизвести двухмерный характер объекта распознавания. Введение понятия «признак изображения» основано на концепции локальных алгоритмов вычисления информации [15] и, в частности, предполагает существенное использование понятия «локальная окрестность». Признак рассматривается как проявление на изображении некоторого интересного в контексте задачи свойства, которое может либо иметь некоторую семантическую нагрузку, либо воспроизводить физические или геометрические свойства отображаемой сцены, либо иметь некоторую количественную меру (числовое значение функций или характеристик, сопоставляемых фрагменту изображения). Задается признак изображения через предикат, определяющий выраженность или проявление соответствующего свойства на некоторой локальной окрестности изображения — опорном множестве предиката признака (опорном множестве признака). Это опорное множество определяется как минимальная локальная окрестность, на которой сохраняется устойчивая вычислимость (в ослабленном случае — вычислимость) предиката признака, причем устойчивая вычислимость характеризуется локальностью вычисления (сложностью и потребными ресурсами алгоритма) и допустимым диапазоном изменений значений пикселов в локальной окрестности. (Отметим, что один из вариантов приведения изображения к виду, удобному для распознавания,— это определение разбиения изображения на опорные множества признаков или построение оптимальной упаковки изображения опорными множествами признаков.)

В качестве оценки «свойства» может использоваться, в частности, некоторая числовая величина, отражающая свойства локального участка изображения (распределение значений пикселов на этом участке; наличие или отсутствие некоторого характерного объекта на этом участке; тип формы объекта, выделяемого на участке; мощность локальной окрестности). Кроме того, к числу полезных локальных признаков, т. е. признаков, вычисляемых по локальной окрестности изображения, следует отнести: характеристики, основанные на шенноновской мере, вычисляемой для распределения типов окрестностей отдельных элементов изображения; характеристики, основанные на распределении типов булевых функций, задаваемых на окрестностях элементов изображения; характеристики, основанные на распределении типов частично определенных булевых функций, задаваемых на окрестностях элементов изображения; характеристики, основанные на числовых оценках свойств графов связности и частично определенных графов связности однородных частей окрестностей элементов изображения.

Признаки, используемые при синтезе модели изображения, классифицируются по функции (порождающие (дескриптивные) — непроизводные элементы, признаки - характерные объекты, признакопорождающие процедуры и параметрические), по характеру отображаемой информации (глобальные и локальные), по способу получения (измеряемые или выделяемые на изображении и вычисляемые), по математическому аппарату, используемому для формирования или вычисления (статистические, алгебраические, топологические, спектральные, геометрические, матричные), по типу изображений, к которым они относятся (яркостные, бинарные, текстурные), по типам объектов, представляемых с их помощью (остовные, контурные, сегментационные).

Заметим, в частности, что в представлении типа «первоначальный эскиз» для задания информации, характеризующей изменения яркости, распределения и геометрические характеристики двухмерных изображений, используются непроизводные элементы следующих типов: пересечения нулевого уровня, пятна, обрывы, нарушения непрерывности яркостных переходов, границы, линии. В представлении же типа «двух с половиной мерный эскиз» для представления задания информации, характеризующей геометрические свойства видимых поверхностей (ориентация, глубина, контуры нарушений непрерывности), используются непроизводные элементы следующих типов: локальная ориентация, расстояние от наблюдателя, нарушения непрерывности [11].

г. Модели. Как следует из принципа порождения и видов признаков, используемых для синтеза формализованных представлений изображений, в процессе распознавания изображения представляются моделями двух принципиально различных типов — дескриптивными (порождающими) и признаковыми (параметрическими). Первые отражают структурную организацию изображения и являются той информацией, которая подвергается обработке в процессе преобразования представлений. Вторые отражают реальные физические свойства изображаемых объектов, а также свойства числовой матрицы, посредством которой задается наблюдаемое изображение. Признаковые модели, как правило, удобнее использовать для получения агрегированных оценок формализованного описания. В целом дескриптивные модели скорее характеризуют процессы перехода между представлениями различных уровней, в то время как признаковые модели целесообразнее ставить в соответствие результатам преобразований представлений, относящимся к конкретным уровням.

Заметим, что установление соответствий этих двух типов моделей применительно к некоторому уровню или шагу процесса распознавания может служить критерием управления ходом процесса и правилом остановки.

4. **Двойственность.** Принципиальным свойством процесса распознавания изображений является наличие в нем двойственностей двух типов: двойственности между формализованным описанием изображения и процедурой распознавания и двойственности между формализованным описанием и формализованным представлением изображения. Под двойственностью в данном случае понимается свойство внутренней симметрии, присущее ряду аксиоматических теорий и выражающееся во взаимозаменяемости некоторых основных понятий. Как уже отмечалось выше, при распознавании изображений изменяются понятия начальной и финальной информации. Многоуровневость процесса распознавания приводит к тому, что он сводится к некоторой последовательности преобразований формализованных описаний изображений, относящихся к различным морфологическим или масштабным уровням. При этом формализованное описание изображения выступает в роли как начальной, так и финальной информации, поскольку при анализе изображений в качестве искомого результата может требоваться и классификационное решение, и некоторое формализованное описание, соответствующее определенным условиям, определяемым контекстом задачи.

Двойственность второго типа связана с тем, что тип выбираемого представления в значительной мере определяет средства, используемые при формировании соответствующего формализованного описания (характерные объекты, непроизводные элементы), последнее же, в свою очередь, влияет на типы формализованных представлений, включаемых в траекторию преобразований, реализующую процесс распознавания (это связано, в частности, с проблемой допустимости порождающих преобразований).

5. **Характеризация процесса распознавания.** Организация процесса распознавания изображений подчиняется нескольким основным принципам, определяющим методы формирования распознающих преобразований, структуру этих преобразований, механизм организации процесса в целом и модель алгоритма распознавания изображений.

а. При выборе и формировании распознающих преобразований основой является методология алгебраического подхода к задачам распознавания и классификации, включающая, как известно, три основных этапа: выбор и применение базовых эвристических процедур, синтез и параметризацию модели эвристических процедур и выполнение оптимизации в модели, синтез процедуры решения задачи посредством коррекции на множестве базисных моделей эвристических процедур (с использованием механизмов линейного и алгебраического замыканий).

б. Модель алгоритма распознавания изображений состоит в отличие от классической модели алгоритма распознавания не из двух, а из трех элементов:

оператора приведения изображения к виду, удобному для распознавания:

$$R_f^n(I_n) = P_n(I, \Omega^m, \omega'), \quad n = 1, \ldots, t, \tag{3.9}$$

где I_n — изображение, соответствующее некоторому морфологическому (масштабному) уровню формализованного описания, $P_n(I, \Omega^m, \omega')$ — модель изображения на n-м уровне формализованного описания, полученная в результате действия оператора приведения изображения к виду, удобному для распознавания;

распознающего оператора, осуществляющего преобразование формализованного описания в некоторую числовую матрицу $\| a_{ji} \|$ стандартного размера с числом строк, равным числу распознаваемых в задаче объектов, и числом столбцов, равным числу классов, рассматриваемых при решении задачи:

$$B(I, \Omega^m, \omega') = \| a_{ji} \|_{q\times m}; \tag{3.10}$$

решающего правила, обеспечивающего переработку полученной числовой матрицы в матрицу окончательных ответов $\| a^{ji} \|_{q\times m}$ с тем же числом строк и столбцов:

$$c(\| a_{ji} \|_{q\times m}) = \| a_{0ji} \|_{q\times m}. \tag{3.11}$$

Очевидно, что главной функцией оператора приведения является получение такого формализованного описания изображения, которое удобно для применения к нему распознающего оператора (3.10), т. е. получение формализованного описания, включающего числовые оценки информации, содержащейся в изображении, которые учитывали бы двухмерный характер этой информации. Заметим, что в классе ДАВО (алгоритмов вычисления оценок по двухмерной информации) [10] в качестве оператора приведения используется сетка дискретизации с ячейками произвольной формы (правильной), каждая из которых покрывает некоторую совокупность смежных пикселов исходного изображения. Конкретный способ подсчета оценки для локальной окрестности определяется набором структурных параметров, указывающих размер ячейки, ее форму, мощность, ориентацию, способ подсчета (правило — усреднение, правило большинства, то же с учетом весов пикселов по вертикали, диагонали, горизонтали и т. д.), веса пикселов, значения порогов дискретизации. Это означает, что оператор приведения задается параметрической моделью, с которой можно работать точно так же, как и с другими моделями эвристических процедур распознавания.

Указанная тройка (3.9—3.11) является стандартным элементом процесса распознавания изображений.

в. В целом структура процесса распознавания имеет иерархический характер в силу двойственности, существующей между применением распознающих процедур и формализованными описаниями изображений. Интересно, что преобразования представлений осуществляются не только «по вертикали» — между различными морфологическими и масштабными уровнями, но и «по горизонтали» — при оптимизации выбора формализованного представления и описания в пределах одного уровня. Таким образом, траектория процесса распознавания включает вертикальные участки (при этом контроль за ходом процесса осуществляется посредством преобразований по горизонтали) и горизонтальные участки (при этом контроль за ходом процесса осуществляется посредством преобразований по вертикали). Основным механизмом контроля служит установление соответствия между порождающими и параметрическими моделями на отдельных уровнях иерархии описания изображения.

г. Достижение соответствия формализованных описаний, синтезируемых в процессе распознавания, а также проверка выполнения условий правила остановки процесса осуществляются с помощью механизма реверсивного алгебраического замыкания (см. раздел «Математическая постановка задачи распознавания изображений»). Смысл преобразований (3.3)—(3.7) состоит в том, что процесс распознавания имеет итерационный и реверсивный характер; в результате достижение соответствия обеспечивается на некотором промежуточном этапе (в смысле формализованных представлений и описаний) и преобразования применяются параллельно к исходным моделям и к финальным моделям, если их вид задан, либо к соответствующим моделям-гипотезам. Естественно, этот механизм используется и при вертикальных, и при горизонтальных преобразованиях иерархии формализмов изображений. При этом основным структурным элементом служит тройка операторов (3.9)—(3.11), а увеличение разрешающей мощности (и сложности) преобразований осуществляется по схеме «базовая эвристика — модель эвристической процедуры — коррекция на множестве базисных моделей».

6. **Логическая фильтрация.** Достаточно очевидно, что в процессе распознавания изображения интенсивно используются знания о предметной области, характере задачи, физических реальностях объекта изображения (сцены), универсальных физических, логических и математических законах, которым, естественно, подчиняется и объект изображения (например, геометрические закономерности), способах получения и формирования изображения и обстоятельствах, этим процессам сопутствовавших. Эти знания используются при разбиении на классы эквивалентности, выдвижении гипотез о типе и характере финальной информации в процессе распознавания, выборе и сопоставлении базовых эвристик и моделей, формировании правил остановки и контроля процесса распознавания, выборе и назначении непроизводных элементов и характерных объектов, типов и уровней формализованных представлений изображения. Основные формы, в которых представляются эти априорные значения,— семантическая информация и системы физических ограничений. Основным механиз-

мом реализации этих знаний служит логический анализ, и в частности логический вывод.

7. **Комплекс преобразований,** обеспечивающих решение конкретной задачи распознавания изображений, образует траекторию решения задачи. Формирование траектории в некотором смысле аналогично этапу обучения или настройки алгоритма на задачу в классическом варианте распознавания. Очевидно, что для задач распознавания изображений одного класса (анализ идеальных изображений, относящихся к одному и тому же классу эквивалентности) траектории должны обладать близостью в пространстве преобразований; кроме того, должна иметь место и близость соответственных формализованных представлений и описаний. В таком случае задача синтеза траектории решения новой задачи сводится к задаче выбора оптимальной траектории из множества траекторий, поставленных в соответствие некоторому классу эквивалентности (по крайней мере после того, как класс эквивалентности для наблюдаемого изображения восстановлен). Другим вариантом этой задачи служит задача восстановления траектории по ее фрагментам, поставленным в соответствие некоторым типам описаний, представлений и их преобразований. Эта задача может возникать в связи с проблемой эффективного использования знаний о задаче и, в частности, на этапе логической фильтрации. Очевидно, что пучки траекторий задачи могут служить основой для синтеза соответствующих параметрических моделей, на которых может ставиться и при определенных условиях разрешаться задача оптимизации выбора и т. д., что естественно приводит решение задачи о синтезе траектории к схеме алгебраического подхода «эвристика — модель — коррекция».

Важнейшей открытой проблемой дескриптивной теории распознавания изображений в настоящее время является систематизация средств описания изображения в задачах распознавания (главным образом признаков, непроизводных элементов и характерных объектов) и методов синтеза формализованных представлений и описаний (главным образом порождающих процедур).

Заключение

РАСПОЗНАВАНИЕ — ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ РАЗРАБОТКИ АЛГОРИТМИЧЕСКИХ БАЗ ЗНАНИЙ

Данный обзор посвящен итогам развития и современному состоянию распознавания — информационной технологии (одной из первых реально и широко используемых практически), обладающей развитым математическим аппаратом и исключительными прикладными возможностями. Проанализировав основные достижения распознавания образов и рассмотрев проблему распознавания изображений, мы пришли к следующим главным выводам:

а) распознавание образов располагает развитой и в определенном смысле законченной математической теорией, сформированной на базе так называемого алгебраического подхода;

б) задача распознавания со стандартной информацией и задача распознавания изображений различны в столь значительной мере, что, хотя последние ставятся и решаются в рамках методологии распознавания, они не допускают непосредственного использования методов и средств классической теории распознавания и требуют разработки специального направления теории распознавания, предназначенного для работы с изображениями;

в) теория распознавания изображений может быть развита в рамках дескриптивного подхода в виде дескриптивной теории распознавания изображений.

Целью алгебраического подхода к распознаванию является получение алгоритма, обеспечивающего выделение из представленных данных всей полезной информации и получение решения, полностью и точно соответствующего «информативности» этой информации. Такое решение характеризуется минимальной (относительной) вычислительной сложностью, устойчивостью к шуму и искажениям в исходной информации и статистической надежностью. В процессе решения существенно используются принцип прецедентности, формализация понятия обобщенной близости, автоматизация настройки алгоритма на задачу, в том числе автоматизация выбора класса алгоритмов, оптимального для рассматриваемого класса задач, и принцип коррекции окончательного решения посредством расширения базового множества моделей алгоритмов, используемых для его получения.

Процесс решения имеет многоуровневый характер. На первом этапе строится эвристическая модель алгоритма, отражающая специфику задачи. На следующем этапе работа ведется с моделями семейств алгоритмов, порождаемых на основе принципа, выбираемого эвристически, стандартным образом. На этом этапе оптимизация алгоритма распознавания осуществляется в рамках отдельных моделей. На третьем этапе искомый алгоритм синтезируется из алгоритмов, принадлежащих разным моделям.

Таким образом, алгебраический подход к обработке информации в задачах распознавания, прогнозирования и искусственного интеллекта обеспечивает реализацию идеологии, позволяющей синтезировать алгоритм, который при выполнении определенных нежестких и просто проверяемых условий точно решает конкретную задачу. Это своего рода методология автоматизации синтеза алгоритмов распознавания — САПР для алгоритмов распознавания и прогнозирования, обеспечивающая возможность предварительно анализировать предъявленную задачу, учитывать ее особенности, после чего выбирать метод решения и на его основе предлагать соответствующий алгоритм. Главное отличие алгебраического подхода от других методов распознавания состоит в том, что в них отсутствует этап три, а следовательно, нет и реальной возможности получать точное решение. Смысл этого этапа состоит в снятии трудностей, возникающих на этапе 2 (неточность моделей эвристических алгоритмов распознавания, сложности, возникающие при реальном проведении оптимизации в многопараметрическом пространстве), и возможности обеспечивать получение абсолютно точного в указанном выше смысле решения в отличие от, как правило, лишь локально-экстремальных решений этапа 2.

В развитии идеологии алгебраического подхода в задачах распознавания для работы с изображениями сформулирована и разрабатывается дескриптивная теория распознавания изображений. Она предусматривает решение задач, связанных с получением формализованных представлений и описаний изображений как объектов распознавания и синтезом процедур их распознавания, при помощи изучения внутреннего строения структуры и содержания изображения как результата тех порождающих операций, при помощи которых изображение может быть построено из непроизводных элементов и объектов, выделяемых на изображении на различных этапах его анализа. Поскольку этот способ характеризации изображения является операциональным, весь процесс обработки и распознавания изображений, включая построение формализованного описания — модели изображения (траектория задачи), рассматривается как реализация над изображением некоторой системы преобразований, определенных на классах эквивалентности, которые представляют ансамбли допустимых изображений (ансамбль также задается дескриптивно — системой прототипов и функционально полным относительно класса эквивалентности набором порождающих преобразований). В процессе распознавания участвует иерархия формализованных описаний и представлений изображений, в частности модели, относящиеся к различным морфологическим и масштабным уровням представления, многоуровневые модели, позволяющие в процессе распознавания выбирать и изменять необходимую степень подробности описания объекта распознавания.

Возникновение дескриптивного подхода объясняется рядом неотъемлемых свойств изображения как способа представления информации и особенностями организации и реализации процесса распознавания изображений. К их числу, очевидно, в первую очередь необходимо отнести следующие:

а) синтез формализованного описания изображения как объекта распознавания становится самостоятельной задачей, которая ставится и решается в рамках процесса распознавания, а не вне его, как это обычно происходит в классических методах распознавания со стандартной информацией;

б) формализованное описание изображения должно сохранять преимущества, возникающие, собственно, в силу использования изображения как средства представления исходной информации; поэтому при его построении само изображение не должно «разрушаться», т. е. описание должно сохранять «изобразительный характер изображения» — оно должно задаваться некоторой формальной конструкцией, при синтезе которой последовательно проводятся принципы порождения и иерархичности структуры объекта распознавания и отношения, существующие между описаниями, включенными в эту иерархию, и их непроизводными элементами и характерными объектами как в пределах отдельных морфологических и масштабных уровней, так и между ними;

в) связь (двойственность) процедур порождения формализованных описаний и распознавания, определяющаяся как ролью собственно моделей в процессе распознавания изображений, так и возможностью представления результата распознавания в виде некоторого нового формализованного описания, сопоставляемого анализируемому изображению;

г) реализация процесса распознавания изображений как траектории преобразований формализованных описаний предполагает для управления ходом процесса распознавания использование информации, получаемой при помощи установления соответствий между формализованными описаниями изображения в пределах и между уровнями иерархии.

Концепция, положенная в основу дескриптивной теории распознавания изображений, предусматривает стандартную организацию процедур преобразования и представления информации при решении задач распознавания. Структура процесса является многоуровневой, на каждом уровне выбираются, отыскиваются (обнаруживаются) или вычисляются признаки, непроизводные элементы, характерные объекты и порождающие процедуры, используемые для построения модели изображения на данном уровне. При определении средств формирования описания существенно используются знания — о предметной области, задаче, универсальные физические логические и математические законы, ограничения, отражающие реальности сцены, являющейся объектом изображения. Структура процесса распознавания отражает особенности строения модели алгоритма распознавания изображений (в нее включаются оператор приведения изображения к виду, удобному для распознавания, распознающий оператор и решающее правило), возможность перехода от пространства формализованных описаний в пространство агрегированных оценок и реализацию преобразований в рамках механизма реверсивного алгебраического замыкания.

Важнейшими положениями дескриптивной теории распознавания изображений являются:

1. Принцип порождения.

2. Формализация описания и представления изображения посредством учета многоуровневости конструкции описаний, использования в них синтаксической и реляционной информации, обобщения понятия «признак распознаваемого объекта» применительно к задаче распознавания изображений, параллельного использования порождающих и параметрических моделей.

3. Двойственность процедур синтеза формализованных описаний и процедур распознавания и двойственность формализованных описаний и представлений.

4. Организация процесса распознавания с учетом методологии алгебраического подхода (схема «эвристика — модель — коррекция на множестве моделей»), обобщения понятия «алгоритм распознавания» применительно к задаче распознавания изображений, иерархичности процедур преобразований и реализации их посредством механизма реверсивного алгебраического замыкания.

5. Логическая фильтрация, в основе которой лежит использование знаний об изображаемом объекте, предметной области, универсальных физических, логических и математических законах, целях, методах и средствах анализа.

6. Понятие траектории задачи, обеспечивающее сведение задачи выбора и синтеза преобразований в задаче распознавания изображений к решению задачи распознавания со стандартной информацией.

Вся современная теория распознавания, как и лежащая в ее основе идеологии алгебраического, а применительно к изображениям — дескриптивного подхода, является отражением стремления регуляризировать выбор и синтез при помощи обучения алгоритмических процедур преобразования и анализа информации, *предназначенных для решения таких информационных задач*, для которых соответствующие алгоритмы неизвестны. Эти задачи при всей их функциональной и предметной разнородности обладают рядом принципиальных особенностей, позволяющих рассматривать эти задачи как новый, самостоятельный и исключительно интересный в теоретическом и прикладном отношениях класс задач преобразования информации. В этих задачах процесс решения включает два этапа: приведение исходных данных к некоторому стандартному виду (например, синтез формализованного описания) и переход от этого стандартизированного описания в пространство обобщенных оценок, дающих основу для вынесения суждения об изучаемом объекте или явлении «в целом». В этих задачах имеется возможность вводить понятие подобия между стандартизированными описаниями и оперировать им при принятии решения. В этих задачах имеется возможность задавать прецеденты с тем, чтобы их можно было использовать для настройки алгоритма на задачу в процессе обучения. Эти задачи возникают в тех ситуациях, в которых невозможно (в силу ограниченности теоретических представлений) или нецелесообразно (в силу чрезмерности затрат или ограниченности ресурсов) строить нормальные математические модели. В этих задачах исходная информация плохая в силу природы задачи — улучшение информации приводит к исчезновению задачи, по крайней мере, в такой постановке. Отметим еще раз, что в целом это задачи, в которых известно слишком мало для того, чтобы было возможно или целесообразно пользоваться классическими методами решения (моделями), но все-таки известно достаточно для того, чтобы решение было возможно.

Исторически эти задачи связывают с попытками использовать опыт специалиста высокой квалификации для решения так называемых задач с плохой структурой — в качестве последних обычно выступали диагностические, классификационные или прогностические задачи. В нынешней терминологии задачи такого рода *относят к проблемам разработки экспертных систем*, однако не следует забывать о значительном объеме работ выполненных в связи с этими проблемами в рамках эвристического программирования в 1950—1960-е годы. Работы этого периода, посвященные созданию экспертных систем нулевого поколения, привели к ряду интересных выводов. Оказалось, что, вообще говоря, практически можно решать задачи с «плохой» и даже противоречивой информацией, можно пользоваться эвристическими алгоритмами (эмпирической аксиоматикой), экспериментально убеждаясь в допустимости их использования, и оказалось также, что все попытки создавать универсальные методы решения указанных задач в рамках традиционной математической логики приводят к проблеме алгоритмической неразрешимости, т. е. к переборным методам решения. В сущности, достоинства эвристик и заключаются именно в том, что они позволяют в частных случаях разумно использовать доступную информацию для сокращения или упорядочивания перебора.

В дальнейшем в качестве выхода из этого положения было предложено обратиться к расширениям классической математической логики и математической статистики посредством введения в них средств, позволяющих учитывать ненадежность и противоречивость информации и эвристичность правил вывода. Для этого был использован классический метод математической кибернетики — метод оценивания. Любой эмпирической аксиоме и эвристическому правилу вывода, включаемым в экспертную систему, ставится в соответствие мера (оценка), характеризующая их достоверность; точно так же оценки достоверности ставятся в соответствие цепочкам вывода.

Таким образом, первый способ расширения предусматривает структуризацию исходной информации, выбор систем эмпирических аксиом и эвристических правил вывода, сопоставление необщематематическим аксиомам и правилам вывода оценок достоверности, оценивание эвристических выводов и их результатов в целом и принятие решений по оценкам.

В эти же годы вплоть до середины 1970-х годов интенсивно развивались и другие методы расширения — они подробно рассмотрены в данной главе. Напомним лишь, что методология второго типа расширения предусматривает выбор основных эвристических процедур, построение на основе эвристики соответствующей модели, параметризацию последней и переход к постановке и решению на эвристической модели оптимизационных задач.

В результате всей этой деятельности накопилось значительное количество самых разнообразных эвристических алгоритмов, каждый из которых, как правило формулируется на некотором частном языке, не сопровождается указаниями или рекомендациями об условиях его применения или сведениями о соотношении с другими эвристическими алгоритмами. Это множество алгоритмов можно рассматривать как основу для формирования алгоритмических баз знаний современных экспертных систем. В экспертных системах нулевого поколения алгоритмические базы знаний фиксированы, это же множество эвристических алгоритмов можно рассматривать как набор экспериментальных точек, каждая из которых представляет собой результат применения одного эвристического алгоритма. Такой подход позволяет ввести новый способ структуризации эвристических процедур, обеспечивающий переход к единому языку их описания. Этот переход реализуется при помощи параметризации описания алгоритмов, что, в свою очередь, позволяет натягивать на эвристики — экспериментальные точки алгоритмическую оболочку (иначе — за счет вариации значений параметров описания алгоритмов задавать классы алгоритмов) и решать на подобных параметрических моделях оптимизационные задачи.

После того как в результате введения единого способа описания любой алгоритм можно представлять как процедуру преобразования структурированной (т. е. приведенной к некоторой стандартной форме) информации в некоторою матрицу оценок стандартного вида, оказывается возможной новая структуризация (расширение) — посредством задания операций над матрицами. В этом случае можно отказаться от оптимизации в параметрической модели и синтезировать новые алгоритмы в виде многочленов над базисными алгоритмами, связанными операциями сложения, умножения и умножения на скаляр.

Таким образом, возникает проблематика систематического изучения алгоритмических баз знаний экспертных систем, связанная с автоматизацией синтеза новых алгоритмов по экспертным функционалам качества. Основными элементами этой проблематики являются сжатие пространства исходных алгоритмов, разработка средств единообразного описания алгоритмов, структуризация на основе выделения мажорантных алгоритмов и указания радиуса устойчивости и порядка в локальной окрестности (как, например, это сделано в теории локальных алгоритмов [15]), структуризация на основе алгебраического подхода, представляющая собой континуальное расширение конечного числа базисных алгоритмов. Другую ветвь этой проблематики составляют методы и средства формализации в задачах преобразования информации при расширении пространства допустимых информаций, главным образом для случая задания исходной информации в виде изображений.

ЛИТЕРАТУРА

1. *Айзерман М. А., Браверман Э. М., Розоноэр Л. И.* Метод потенциальных функций в теории обучения машин. М.: Наука, 1970. 240 с.
2. *Бонгард М. М.* Проблема узнавания. М.: Наука, 1967. 320 с.
3. *Вайнцвайг М. Н.* Алгоритм обучения распознаванию образов «Кора» //

Алгоритмы обучения распознаванию образов: Сб. ст. / Под ред. В. Н. Вапника. М.: Сов. радио, 1973. С. 110—116.
4. *Вапник В. Н.* Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979. 448 с.
5. *Вапник В. Н., Червоненкис А. Я.* Теория распознавания образов. М.: Наука, 1974. 415 с.
6. *Горелик А. Л., Гуревич И. Б., Скрипкин В. А.* Современное состояние проблемы распознавания. Некоторые аспекты. М.: Радио и связь, 1985. 162 с. (Кибернетика).
7. *Гренандер У.* Лекции по теории образов: В 3 т. / Пер. с англ. под ред. Ю. И. Журавлева. М.: Мир, 1979—1983. 130 с.
8. *Гуревич И. Б.* Анализ изображений методом реверсивного алгебраического замыкания // Проблемы искусственного интеллекта и распознавания образов: Тез. докл. и сообщ. Науч. конф. с участием ученых из соц. стран (Киев, 13—18 мая 1984 г.). Секция II.: Распознавание образов. Киев: Ин-т кибернетики им В. М. Глушкова АН УССР, 1984. С. 41—43.
9. *Гуревич И. Б.* Алгебраический подход к анализу и распознаванию изображений // Математические методы распознавания образов: Тез. докл. Всесоюз. конф. (Дилижан, 16—21 мая 1985 г.). Ереван: Изд-во АН Арм. ССР, 1985. С. 55—57.
10. *Гуревич И. Б.* Определение класса алгоритмов вычисления оценок по двухмерной информации для задач распознавания изображений // Методы и средства обработки графической информации: Межвуз. сб. науч. тр. / Под ред. Ю. Г. Васина. Горький: Горьк. гос. ун-т им. Н. И. Лобачевского, 1986. С. 47—66.
11. *Гуревич И. Б.* Проблема распознавания изображенией // Распознавание, классификация, прогноз. Математические методы и их применение: Ежегодник / Под ред. Ю. И. Журавлева. М.: Наука 1988. Вып. 1. С. 280—329.
12. *Гуревич И. Б., Журавлев Ю. И.* Минимизация булевых функций и эффективные алгоритмы распознавания // Кибернетика. 1974. № 3. С. 16—20.
13. *Гуревич И. Б., Журавлев Ю. И.* О формализации принципов выбора алгоритмов анализа изображений // Методы и средства обработки сложноструктурированной семантически насыщенной графической информации: Тез. докл. I Всесоюз. конф. (Горький, сентябрь 1983 г.). Горький: Горьк. гос. ун-т, 1983. С. 20—21.
14. *Дмитриев А. Н., Журавлев Ю. И., Кренделев Ф. П.* О математических принципах классификации предметов и явлений // Дискретный анализ: Сб. ст. Новосибирск: Ин-т математики СО АН СССР, 1966. Вып. 7. С. 3—15.
15. *Журавлев Ю. И.* Локальные алгоритмы вычисления информации. I—II // Кибернетика. 1965. № 1. С. 12—19; 1966. № 2. С. 1—11.
16. *Журавлев Ю. И.* Экстремальные задачи, возникающие при обосновании эвристических процедур // Проблемы прикладной математики и механики. М.: Наука, 1971. С. 67—75.
17. *Журавлев Ю. И.* Алгоритмы построения минимальных дизъюнктивных нормальных форм для функций алгебры логики // Дискретная математика и математические вопросы кибернетики / Под ред. С. В. Яблонского, О. Б. Лупанова. М.: Наука, 1974. Т. 1. С. 67—98.
18. *Журавлев Ю. И.* Непараметрические задачи распознавания образов // Кибернетика. 1976. № 6. С. 93—123.
19. *Журавлев Ю. И.* Корректные алгебры над множествами некорректных (эвристических) алгоритмов. I—III // Там же. 1977. № 4. С. 14—21; 1977. № 6 С. 21—27; 1978. № 2. С. 35—43.
20. *Журавлев Ю. И.* Об алгебраическом подходе к решению задач распознавания и классификации // Проблемы кибернетики: Сб. ст. М.: Наука, 1978. Вып. 33. С. 5—68.
21. *Журавлев Ю. И.* Об алгебраических методах в задачах распознавания и классификации // Распознавание, классификация, прогноз. Математические методы и их применение: Ежегодник / Под ред. Ю. И. Журавлева. М.: Наука, 1988. Вып. 1. С. 9—16.

22. *Журавлев Ю. И., Зенкин А. А., Исаев И. В., Кольцов П. П., Кочетков Д. В., Рязанов В. В.* Задачи распознавания и классификации с нестандартной обучающей информацией // Журн. вычисл. математики и мат. физики. 1980. Т. 20, № 5. С. 1294—1309.
23. *Журавлев Ю. И., Никифоров В. В.* Алгоритмы распознавания, основанные на вычислении оценок // Кибернетика. 1971. № 3. С. 1—11.
24. *Ледли Р. С.* Программирование и использование цифровых вычислительных машин / Пер. с англ. под ред. А. И. Китова. М.: Мир, 1966. 644 с.
25. *Леонтьев В. К., Сметанин Ю. Г.* О восстановлении векторов по набору их фрагментов // ДАН СССР. 1988. Т. 302, № 6. С. 1319—1322.
26. *Мазуров В. Д., Казанцев В. С., Белецкий Н. Г., Кривоногов А. И., Смирнов А. И.* Вопросы обоснования и применения комитетных алгоритмов распознавания // Распознавание, классификация, прогноз. Математические методы и их применение: Ежегодник / Под ред. Ю. И. Журавлева. М.: Наука, 1988. Вып. 1. С. 114—148.
27. *Матросов В. Л.* Корректные алгебры ограниченной емкости над множествами некорректных алгоритмов // ДАН СССР. 1980. Т. 253, № 1. С. 25—30.
28. *Матросов В. Л.* Корректные алгебры ограниченной емкости над множеством алгоритмов вычисления оценок // Журн. вычисл. математики и мат. физики. 1981. Т. 21. № 5. С. 1276—1291.
29. *Растригин Л. А., Эренштейн Р. Х.* Метод коллективного распознавания. М.: Энергоиздат, 1981. 80 с.: ил. (Б-ка по автоматике; Вып. 615).
30. *Ту Дж., Гонсалес Р.* Принципы распознавания образов / Пер. с англ. под ред. Ю. И. Журавлева. М.: Мир, 1978. 416 с.: ил.
31. *Тюхтин В. С.* Теория автоматического опознавания и гносеология. М.: Наука, 1976. 192 с.
32. *Чегис И. А. Яблонский С. В.* Логические способы контроля электрических схем // Сборник статей по математической логике и ее приложениям к некоторым вопросам кибернетики / Под ред. С. В. Яблонского. М.: Изд-во АН СССР, 1958. С. 270—360 (Тр. Мат. ин-та АН СССР им. В. А. Стеклова. Т. 51).
33. *Юм Д.* Трактат о человеческой природе / Юм Д. Соч.: В 2 т. М.: Мысль, 1966. Т. 1. С. 203—234.
34. *Becke P. W.* Recognition of patterns using frequences of occurence of binary words. Vienna; N. Y.: Springer, 1978. 120 p.
35. *Pappentin F., Kruger M.* Binary sequences. 3. Complexity versus homogeneity and symmetry // Inform. Sci. 1983. Vol. 31. P. 33—39.
36. *Watanabe S.* Pattern recognition: Human and mechanical. N. Y.: Wiley, 1985. 592 p.