

УДК 62-5:007:621.391:519.2:51:16

МАШИННЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ ЗАКОНОМЕРНОСТЕЙ

Н.Г.Загоруйко

В течение 10 лет в Институте математики СО АН СССР сначала в лаборатории, а затем в отделе моделирования информационных процессов разрабатываются методы обнаружения закономерностей в больших массивах информации. Работа ведется в 2-х направлениях: 1) автоматическое распознавание образов и 2) эмпирическое предсказание.

§ 1. Автоматическое распознавание образов

В задачах распознавания образов исследователь имеет дело с некоторым множеством изучаемых объектов и набором признаков, которыми характеризуются эти объекты. Сами объекты и их признаки выбираются в зависимости от цели исследования. Эта же цель определяет выбор метрики в пространстве признаков и меры близости ("похожести") между объектами. В геометрической интерпретации под образом понимается область в пространстве признаков, в которую отображается множество объектов или явлений, выделяемое в соответствии с определенной целью. Краткое описание этой области называется эталоном, а отдельная точка в ней — реализацией образа. Перечень фиксированных областей, на которое разделено пространство признаков, называется алфавитом образов. В практических задачах распознавания образов количество рассматриваемых объектов, число измеряемых признаков и количество различных значений, которое могут принимать эти признаки, конечно. Количество образов также конечно и всегда не менее двух.

Всякое распознающее устройство состоит из блока измерений параметров ("рецептора" X), блока принятия решения ("классификатора" D) и блока исполнительных устройств ("эффектора" S). Рецептор измеряет значения параметров X реализации образа. Классификатор в соответствии с определенными правилами D указывает, какой области z выборочного пространства принадлежит эта точка X . Результат решения классификатора воплощается эффектом в определенное действие (нажимается клавиша пишущей машинки, включается двигатель и т.д.). Часто распознающий автомат представляет собой цепочку из нескольких таких элементарных распознающих автоматов.

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

1. Какие элементы алфавита (z) автомат должен отличать друг от друга (что распознавать)?
2. На какие свойства (x) предметов или явлений следует обращать внимание (т.е. что измерять)?
3. По каким правилам (D) следует принимать решение о принадлежности реализации к тому или иному элементу алфавита?
4. Какие дополнительные ограничения и гипотезы (N) используются в алгоритме распознавания?
5. Какова стоимость (N) распознающего устройства и потеря из-за ошибок в его работе?

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

Среди задач распознавания образов можно выделить задачи трех основных типов.

1. Заданы все элементы, кроме решающего правила. Требуется найти такое решающее правило D , при котором затраты, связанные с его построением и дальнейшим использованием, достигли бы минимума.

2. Заданы все элементы, кроме признакового пространства. Требуется найти такую систему признаков X , при которой достигли бы своего минимума затраты, связанные с ошибками распознавания и с измерением этих признаков.

3. Заданы все элементы, кроме алфавита образов. Требуется найти такой вариант группировки реализаций в образы S , при ко-

тором минимизировались бы затраты, связанные с дальнейшим использованием этих образов.

При рассмотрении многоуровневых структур распознающих устройств можно обнаружить более сложные задачи комбинированного типа. Они представляют собой комбинацию любых двух или всех трех задач основного типа. Так, встречаются задачи, в которых неизвестными являются одновременно и признаковое пространство, и алфавит распознаваемых образов, или признаковое пространство и правило принятия решений и т.д. [2, 1].

Задачи первого типа. Специфика этой задачи состоит в следующем:

1. Имеется некоторая система A ("учитель"), которая устанавливает классификацию экспериментального массива Z , описанного на языке системы признаков X_A , путем разбиения множества Z на конечное число k непересекающихся подмножеств (образов) S_i ($i=1, 2, \dots, k; k \geq 2$). Тому факту, что система A имеет возможность с помощью эксперимента однозначно установить, к какому классу принадлежит любой элемент массива Z , в литературе по распознаванию часто соответствует выражение "учителю известна классификация генеральной совокупности реализаций".

2. Указанная выше классификация известна обучающему автомату B ("ученику") лишь для части массива эмпирических данных (для "обучающей выборки" $Z_{об}$), а цель обучения состоит в выработке логических методов распространения классификации на большую часть массива (иногда и на весь массив).

3. Реализации обучающей выборки описаны на языке системы признаков X_B , в общем случае не совпадающей с системой X_A .

4. Процедура обучения заключается в построении некоторой ("решающей") функции, зависящей от $Z_{об}$ и X_B , которая на всех реализациях подмножества S_i (в том числе и не входивших в обучающую последовательность) должна принимать значения D_i , а на реализациях подмножества S_j — значения D_j ; причем $D_i \neq D_j$, если $i \neq j$; $i, j = 1, 2, \dots, k$.

5. Решающая функция D , построенная по обучающей выборке, позволяет автомату B принимать на генеральной совокупности решения, совпадающие с решениями автомата A , если выполняются следующие условия: а) закономерность, которая имеется в структуре множества обучающих реализаций, хорошо соответствует закономерности в структуре генеральной сово-

кучности; б) мы правильно угадали (постулировали) характер закономерности; в) наш формальный аппарат точно отражает характер этой закономерности.

Из этого видно, что выбор решающей функции в сильной степени зависит от того, какую закономерность генеральной совокупности мы постулировали. Наибольшее распространение в алгоритмах распознавания образов получил постулат или гипотеза "компактности" [3]. В соответствии с этой гипотезой реализации одного образа "компактны", т.е. близко друг к другу расположены в пространстве признаков и могут быть разделены от реализаций других образов с помощью "не слишком вычурной" границы. Граница, проходящая между областями, принадлежащими разным образам, и есть решающая функция. Отсюда следует, что в соответствии с гипотезой компактности, решающие функции нужно искать среди функций "простого" типа.

Выбор вида решающей функции осуществляется обычно не из бесконечного множества их вариантов, а из некоторого конечного набора. Если функциям, входящим в этот набор, присвоить номера ($v = 1, 2, \dots, V$), то задача поиска наилучшего вида решающей функции сводится к поиску номера v функции D_v , которая удовлетворяет следующему условию:

$$v = \arg \min_{v \in V} N(D_v) / S, X, Z_{00}, H, N_0.$$

Здесь N — предельно допустимая величина затрат N , а $\arg \min$ означает "аргумент, который минимизирует".

В качестве решающих функций можно было бы использовать такие хорошо исследованные статистические критерии, как критерии Байеса, Вальда, Неймана-Пирсона и т.д. Однако эти критерии работают в условиях, когда известны плотности распределения генеральных совокупностей всех k образов $p_1, p_2, \dots, p_k$. Практически такие условия не встречаются. Обычно сведения о генеральных совокупностях представлены в виде оценок $\bar{p}_1, \bar{p}_2, \dots, \bar{p}_k$, полученных по обучающей последовательности длиной $L_{об}$, где $L_{об}$ конечно и часто очень мало. В этих условиях строгие статистические методы можно рассматривать лишь как некую идеальную модель, которую следует иметь в виду, разрабатывая или используя любое решающее правило.

Даже при самых простых исходных гипотезах, таких как гипотеза компактности или гипотеза линейной разделимости, процедура построения решающей функции по обучающей выборке не всегда тривиальна. С большими техническими трудностями, приходится сталкиваться, в частности, при работе с большим числом образов в пространстве большой размерности. Ряд работ отдела был посвящен поиску способов уменьшения этих трудностей. В одной из них описывается алгоритм построения линейных решающих функций, близких к оптимальным [4, I].

Идея этого алгоритма состоит в следующем. Если число распознаваемых образов велико, то обычно некоторая гиперплоскость, разделяющая два образа, попутно разделяет друг от друга и еще некоторые подмножества образов. Следовательно, можно обойтись количеством решающих функций меньшим, чем число сочетаний из K по 2, которое необходимо при разделении всех K образов друг от друга попарно. Задача состоит в том, чтобы найти минимальное число гиперплоскостей, с помощью которых можно было бы отделить друг от друга все образы. Был разработан алгоритм, позволяющий решить эту задачу, и проанализированы условия, при которых ее целесообразно решать.

В других работах [5, I] описан так называемый комбинированный метод принятия решений, при котором процедура состоит из двух этапов. На первом этапе ("вычеркивание") с помощью простых пороговых решений при последовательном использовании отдельных признаков количество образов, участвующих в рассмотрении, существенно уменьшается. На втором этапе ("сравнение") решение принимается с помощью оптимальных решающих функций в пространстве полной размерности. Этот метод позволяет сократить объем вычислений более чем на порядок.

Метод дробящихся эталонов [6, I] предназначен для построения решающих функций в том случае, когда области различных образов не разделяются с помощью гиперплоскостей. Решающая функция представляется при этом в виде иерархии простых геометрических фигур, например гиперсфер. Гиперсферами первого уровня выделяются минимальные области, в которых содержатся все реализации данного образа. Для точек, попавших в область пересечения сфер, описывающих разные образы, строится гиперсфера второго уровня и так до полного разделения образов. Таким простыми средствами можно реализовать решающую функцию любой сложности.

Различные решающие функции отличаются друг от друга как мощностью, так и затратами, связанными с их реализацией. В одной из работ [7, 1] был предложен критерий, который позволяет сравнивать между собой решающие функции разных типов по этим двум параметрам одновременно. Было показано, что если граница между образами представляет собой сложную поверхность, то ее рационально аппроксимировать участками границ более простого типа. Так, если вместо полинома 5-й степени взять равный по мощности набор гиперплоскостей, то эффективность такой решающей функции примерно в 200 раз лучше исходной. Наибольшей эффективностью обладают решающие функции в виде дробящихся эталонов.

В последнее время велись работы по синтезу решающих функций, представляющих собой набор конъюнктивных связей значений признаков. Были разработаны алгоритмы для принятия решений по признакам, имеющим двоичную форму [8]. Более существенным является разработка алгоритмов, позволяющих синтезировать решающие функции в случае работы с признаками, измеренными в разных шкалах [9]. Алгоритмы поиска решающих функций этого типа включают в себя блоки, позволяющие уменьшить неизбежно возникающие при этом вычислительные трудности.

В ряде работ [10, 11, 12] делались попытки продвинуться в решении одной из наиболее трудных задач в области распознавания — задачи представительности выборки.

Большое внимание в работах отдела уделяется проблеме выбора системы информативных признаков. Исходный набор признаков задается обычно на основании опыта и интуиции специалиста в той прикладной области, в которой решается конкретная задача распознавания. Формальные же методы применяются на этапе проверки заданной исходной системы признаков на необходимость и достаточность. Исходное множество признаков (X) конечно. Конечно также и число всех подмножеств этого множества. Если эти подмножества (X_β) пронумерованы ($\beta = 1, 2, \dots, B$), то проблема выбора системы признаков, которая была бы достаточно информативной и вместе с тем минимальной по сложности или стоимости, сводится к решению задачи следующего вида:

$$\beta = \arg \min H(X_\beta) / S, D, H, Z_{00}, N_0.$$

Хотя в практических задачах распознавания B и конечно, тем не менее оно так велико, что простой перебор всех X_β невозможен. Основные усилия в решении этой задачи сводились к поиску способов сокращения перебора. В алгоритме случайного поиска с адаптацией (СПА) для этой цели используется видоизмененный метод Монте-Карло, основанный на некоторых идеях работы персептрона [13]. Вначале предполагается, что вероятности включения различных признаков в искомое признаковое пространство размерности n одинаковы. С помощью датчиков случайных чисел выбирается несколько n -мерных подпространств исходного пространства признаков. По результатам сравнения информативности этих наборов вносится изменение в упомянутые выше вероятности. В результате вероятность выбора тех признаков, которые чаще других входили в состав наиболее хороших подмножеств, увеличивается, а вероятность признаков, попадавших в неудачные наборы, уменьшается. Таким путем выделяется наиболее предпочтительный набор из n признаков. Алгоритм СПА получил дальнейшее развитие и нашел применение при решении разных задач из области планирования эксперимента, подбора параметров в эмпирических формулах и т.д.

В работе [8] описан эффективный алгоритм поиска информативных подмножеств бинарных признаков. В последнее время в этой области внимание привлекают проблемы выбора, сравнения, проверки зависимости признаков, измеренных в разных шкалах [9]. Среди конкретных систем описания распознаваемых объектов исследовались также, как двумерные спектры при распознавании произвольных изображений [14] и специальный язык, предназначенный для распознавания рукописных цифр [15]. При распознавании сигналов, являющихся функциями времени, исследовались различные способы снятия сигналов без потерь информации за счет устранения имеющейся в этих сигналах избыточности [16].

Большой цикл работ связан с проблемой таксономии, или автоматической классификации. Разбиение множества объектов на некоторые классы, или таксоны, связано с потерей информации об индивидуальных свойствах каждого отдельного объекта. Если количество классов сделать большим, то эти потери информации уменьшаются, однако одновременно возрастает потеря, связанная с необходимостью хранить в памяти большое число классов. Естественно стремление исследователей построить такую классификацию, которая представ-

ляла бы собой хороший компромисс между этими противоречивыми требованиями простоты и полноты описания исследуемых объектов. В общем виде этому требованию удовлетворяет решение следующей задачи:

$$\omega = \arg \min_{\omega \in \Omega} N(S_{\omega}) / X, D, H, Z_{00}, N_0.$$

Здесь S_{ω} — искомый вариант таксономии, выбранный из конечного числа Ω вариантов.

В первых алгоритмах таксономии, разработанных в отделе, в качестве гипотезы H использовалась гипотеза компактности. Это определило разработку алгоритмов типа "Форэль" [17], результатом которых являются таксоны простой геометрической (сферической) формы. Позже вопрос о том, какими критериями руководствуется человек в процессе классификации, например, точек, изображенных на плоскости, был подвергнут специальному исследованию. Было обнаружено, что человек при этом руководствуется такими особенностями распределения точек, как их близость друг к другу, удаленность одной группы точек от другой, равномерность распределения количества точек между группами и локальным характером изменения расстояния между точками. Общий критерий качества таксономии, использующий все эти четыре фактора, позволил сконструировать алгоритмы таксономии серии "Крб", с помощью которых можно получать таксоны произвольной конфигурации [18].

Выше уже отмечалось, что в практике встречаются задачи так называемого комбинированного типа. Так, иногда исследователю необходимо найти такое подмножество признаков, в пространстве которых данный набор объектов наилучшим образом разделяется на таксоны. Решение такой задачи состоит в одновременном применении методов таксономии и метода поиска информативной системы признаков [19]. Применение методов распознавания комбинированного типа позволяет решать широкий круг задач, таких как задачи сокращения экспериментальных данных за счет устранения избыточности, задачи заполнения пропусков в эмпирических таблицах и т.д. [20].

Практически все разработанные в отделе алгоритмы распознавания были реализованы в виде машинных программ и нашли применение при решении различных прикладных задач. С участием сот-

рудников отдела наиболее часто решались задачи из области медицины [21], геологии [22, 23, 24, 25] и социологии [26].

Самостоятельное значение для отдела имеют задачи распознавания речевых сигналов [27, 28]. Был разработан комплекс аппаратуры для ввода речевых сигналов в ЭВМ [29]; исследовались статистические характеристики элементов русской речи: фонем, слогов [30, 31, 32]; исследовались различные способы описания речевого сигнала [33, 34, 35]; были разработаны методы повышения надежности распознавания речевых элементов за счет избыточности, содержащейся в речи [36, 37, 38]; исследовались способы автоматического определения места ударения в изолированных словах и в словах слитной речи [39]. Объединение результатов этих исследований позволило получить следующие практические результаты. В 1968 году были проведены эксперименты по распознаванию словаря, состоящего из 168 устных команд [34]. При подстройке под диктора машина распознала этот словарь с надежностью, большей 95%. Такая же надежность была получена вскоре при распознавании словаря, включавшего в свой состав 203 команды [40]. В настоящее время начаты эксперименты по распознаванию словаря из 2000 слов. Во всех этих экспериментах в качестве системы признаков используется энергия на выходе пяти октавных фильтров, измеряемая каждые 10-15 мсек.

В предыдущих экспериментах распознавание производилось путем сравнения эталонной траектории каждого слова с траекторией контрольного слова в пространстве этих 5 признаков. В процессе сравнения использовался алгоритм динамического программирования, с помощью которого делалась нелинейная нормализация слов по темпу. В экспериментах, проводимых в настоящее время, используется многоступенчатая процедура распознавания. На первом этапе распознаются фонемы, а по ним распознаются слова. Увеличение словаря сильно сказывается на увеличении времени принятия решения. Уже при словаре в 200 слов время распознавания одного слова на ЭВМ БЭСМ-6 равнялось 2,5 сек. Впоследствии это время удалось существенно сократить благодаря методу ассоциативного кодирования. В результате при распознавании словаря в 2000 слов время распознавания одного слова равно 0,3 сек. Ведутся также работы по распознаванию фраз слитной речи с предварительным автоматическим выделением ударных гласных. 112 слов, входящих в состав 39 фраз, были распознаны с надежностью 82% [41].

§2. Эмпирическое предсказание

Многие методы, связанные с обработкой больших массивов информации, основаны на использовании предположения о том, что между различными элементами этого массива имеются определенные закономерные связи. В задачах распознавания образов обычно нужно найти закономерности, связывающие свойства объектов и их принадлежность к тому или иному образу. Знание такой закономерности позволяет предсказывать принадлежность к образу и неизвестной реализации. Анализ задач, решаемых в различных естественнонаучных областях, показывает, что все они так или иначе связаны с обнаружением закономерностей и их использованием для предсказания (предсказания хода развития процесса, предсказания результата будущего эксперимента на основании изучения результатов прошлых экспериментов и т.д.). Предметом наших интересов являются процессы предсказания в эмпирических науках ("эмпирическое предсказание"). Целью исследования в этой области является создание машинных методов обнаружения закономерностей в эмпирических данных и методов использования этих закономерностей для предсказания новых фактов.

Проблема эмпирического предсказания может быть разделена на 3 основные части. Первая из них — методы описания, или методы представления исходных данных. На вход вычислительной машины нужно подавать не только экспериментальный протокол pr_0 , представляющий собой результат взаимодействия данных объектов с данным набором приборов. Нужно сообщить машине также и все априорные данные об экспериментальной ситуации, в частности о свойствах, используемых в экспериментах приборов. Эти априорные данные формируются в виде эмпирической гипотезы h_0 . Гипотеза включает в себя перечень символов и тестовый алгоритм, предназначенный для проверки истинности или ложности данной гипотезы. Проблема представления состоит в разработке правил формирования этой пары данных: исходного протокола pr_0 и исходной эмпирической гипотезы h_0 .

Следующая часть проблемы эмпирического предсказания состоит в процедуре обнаружения закономерностей на множестве входных фактов. Обнаруженные закономерности позволяют усилить наши знания об экспериментальной ситуации. Это эквивалентно тому, что на выходе алгоритма, об-

наруживающего закономерности, машина формирует гипотезу h_1 , являющуюся усилением исходной гипотезы h_0 . Тестовый алгоритм в гипотезе h_1 оценивает в качестве недопустимых не только все те протоколы, которые считались недопустимыми и гипотезой h_0 , но также и некоторые другие протоколы. В итоге сокращается число возможных экспериментальных ситуаций, и тем самым уменьшается исходная неопределенность.

Были сформулированы требования, которым должен удовлетворять алгоритм А, предназначенный для усиления эмпирических гипотез [42]. Требования эти вполне естественны с точки зрения современной науки.

1. У н и в е р с а л ь н о с т ь. Алгоритм А должен быть применимым к любой паре $\langle h_0, pr_0 \rangle$ вне зависимости от того, данные из какой прикладной области содержатся в этой паре.

2. Н е т р и в и а л ь н о с т ь. Алгоритм А должен позволять усиливать исходную гипотезу h_0 .

3. К о р р е к т н о с т ь. Если h'_0 и h''_0 и протоколы pr' и pr'' связаны друг с другом взаимно-однозначно (т.е. если одна пара $\langle h'_0, pr'_0 \rangle$ получается из другой $\langle h''_0, pr''_0 \rangle$ путем нетворческой переформулировки), то алгоритм А, при применении его к 1-й и ко 2-й парам, должен выдавать одну и ту же гипотезу h_1 .

Исследование этих условий дало следующий результат [43]. Оказалось, что множество алгоритмов, удовлетворяющих этим условиям, пусто, т.е. что не существует алгоритма, который был бы в указанном выше смысле универсальным, нетривиальным и корректным одновременно. Сейчас исследуются возможности снятия или ослабления указанных выше требований. Основное внимание привлекает 3-е требование, ослабление которого эквивалентно признанию неравноценности языков, предназначенных для описания входных данных. Известно [44, 45], что в методологической литературе имеется много указаний на целесообразность предпочтения языков, которые наиболее часто употребляются в науке и отличаются простотой и изяществом и т.д.

Третья часть проблемы состоит в разработке алгоритма предсказания новых фактов. Этот алгоритм В должен, отправляясь от пары $\langle h_0, pr_0 \rangle$, давать пару $\langle h_1, pr_1 \rangle$, где протокол pr_1 включает в себя все факты из протокола pr_0 и

содержит некоторые новые факты. Если гипотеза h , допускает существование нескольких конкурирующих теорий, то для предсказания новых фактов нужно выбрать одну из них (например, самую простую теорию). Сейчас разработаны некоторые варианты алгоритмов обнаружения закономерностей и предсказания [46], в которых реализовано такое свойство языка описания, как простота. Работоспособность алгоритмов была проверена в экспериментах по обнаружению закономерностей такого типа, как закон Ома или законы механики Ньютона и закон Менделя. Сейчас ведутся эксперименты с данными, содержащими закономерность, описываемую периодическим законом Менделеева. На вход машины в этих экспериментах подаются данные в виде протокола измерения, например тока, сопротивления, напряжения на различных участках пассивных цепей, а также в виде гипотезы о свойствах измерительных приборов: амперметра, омметра, вольтметра и т.д. Машина обнаруживает закономерности, связывающие показания разных приборов, и может отвечать на вопросы такого рода: "Каким будет напряжение, если сопротивление и ток в этой цепи такие?" или "Какое количество розовых цветков будет наблюдаться в пятом поколении чистой линии цветного горошка?".

По нашим представлениям, проблема эмпирического предсказания является главной составной частью более общей проблемы "искусственный интеллект".

Л и т е р а т у р а

1. ЗАГОРУЙКО Н.Г. Методы распознавания и их применение. М., "Сов. радио", 1972.
2. ЗАГОРУЙКО Н.Г., САМОХВАЛОВ К.Ф. Природа проблемы распознавания образов. -В кн.: Вычислительные системы. Вып.36, Новосибирск, 1969, с. 3-12.
3. АРКАДЬЕВ А.Г., БРАВЕРМАН Э.М. Обучение машины распознаванию образов. М., "Наука", 1964.
4. ЗАГОРУЙКО Н.Г. Линейные решающие функции, близкие к оптимальным. -В кн.: Вычислительные системы. Вып.19, Новосибирск, с. 67-76.
5. ЗАГОРУЙКО Н.Г. Комбинированный метод принятия решений. -В кн.: Вычислительные системы. Вып.22, Новосибирск, 1966, с. 77-80.
6. ЗАГОРУЙКО Н.Г. Структура проблемы распознавания слуховых образов и методы ее решения. Распознавание слуховых образов. Новосибирск, "Наука", 1966.

7. ЗАГОРУЙКО Н.Г. Сравнение решающих функций по мощности и затратам. -В кн.: Вычислительные системы. Вып.37. Новосибирск, 1969, с. 10-14.

8. КОТЮКОВ В.И. Оптимизация критерия Фишера-Уилкса и сокращение исходной системы описания в задачах распознавания образов. -В кн.: Вычислительные системы. Вып. 50. Новосибирск, 1972, с. 136-142.

9. ЛБОВ Г.С., КОТЮКОВ В.И., МАНОХИН А.Н. Об одном алгоритме распознавания в пространстве разнотипных признаков. -В кн.: Вычислительные системы. Вып.55. Новосибирск, 1973, с. 108-110.

10. КУРИЛОВ Б.М. Получение достаточных характеристик $P_i(x)$ при распознавании образов. -В кн.: Вычислительные системы. Вып. 22, Новосибирск, 1966, с. 21-37.

11. ЛБОВ Г.С. О представительности выборки при выборе эффективной системы признаков. -Там же, с. 39-58.

12. ЛБОВ Г.С., МАНОХИН А.Н. Об оценке качества решающего правила на основе малой обучающей выборки. -В кн.: Вычислительные системы. Вып.55. Новосибирск, 1973, с. 98-107.

13. ЛБОВ Г.С. Выбор эффективной системы зависимых признаков. -В кн.: Вычислительные системы. Вып.19. Новосибирск, 1966, с. 21-34.

14. ВОЛОШИН Г.Я., ЗАГОРУЙКО Н.Г. Применение двумерных спектров для автоматического распознавания изображений. -Там же, с. 3-11.

15. НУДЮЛЬМАН А.С. Язык для описания плоских рукописных черно-белых фигур. -В кн.: Вычислительные системы. Вып.44. Новосибирск, 1971, с. 141-154.

16. КУРИЛОВ Б.М. Рекурсивный модуляционный анализ функций. -Там же, с. 49-59.

17. ЕЛИКИНА В.Н., ЗАГОРУЙКО Н.Г. Об алфавите объектов распознавания. -В кн.: Вычислительные системы. Вып.22. Новосибирск, 1966, с. 59-76.

18. ЕЛИКИНА В.Н., ЗАГОРУЙКО Н.Г. Количественные критерии качества таксономии и их использование в процессе принятия решений. -В кн.: Вычислительные системы. Вып.36. Новосибирск, 1969, с. 29-46.

19. ЗАГОРУЙКО Н.Г. Одновременный поиск эффективной системы признаков и наилучшего варианта таксономии (алгоритм "sx"). -Там же, с. 47-61.

20. ЗАГОРУЙКО Н.Г. Общие свойства задач распознавания образов. -В кн.: Вычислительные системы. Вып.44. Новосибирск, 1971, с. 3-12.

21. ЛБОВ Г.С., АБДУЛЛАЕВА Н.С. Выбор существенных признаков для диагностики врожденных пороков сердца. -В кн.: Вопросы кибернетики, №51. Ташкент, Изд-во "ФАН", 1972.

22. ЗАГОРУЙКО Н.Г., САМОХВАЛОВ К.Ф. Распознавание ситуаций по динамическим признакам (алгоритм "ДИП"). - В кн.: Вычислительные системы. Вып. 37, Новосибирск, 1969, с. 38-43.

23. ЕЛКИНА В.Н., ЕЛКИН А.Е., ЗАГОРУЙКО Н.Г. О применении методики распознавания образов к решению задач палеонтологий. - "Геология и геофизика", № 9, 1967.

24. ЕЛКИНА В.Н., ЗАГОРУЙКО Н.Г., КУКЛИН А.П., КОМАРОВСКИЙ Э.Д. Типы рудоносных и оловоносных территорий Чукотки. "Колыма", № 4, 1972.

25. КРЕНДЕЛЕВ Ф.П., ЕЛКИНА В.Н., ГАЙ А.Г., ЛУЧКО А.Г. Исследование возможностей использования ЭВМ для разработки классификации непрозрачных минералов. М., ВИНТИ, 1972 (Депонированная работа, № 4074-72).

26. Применение методов распознавания образов в социальных исследованиях. Под редакцией Н.Г. Загоруйко и Т.И. Заславской. Новосибирск, "Наука", 1968.

27. Распознавание слуховых образов. Под редакцией Н.Г. Загоруйко, Г.Я. Волошина, Новосибирск, "Наука", 1968.

28. БОНДАРКО Л.В., ЗАГОРУЙКО Н.Г., КОЖЕВНИКОВ В.А., МОЛЧАНОВ А.П., ЧИСТОВИЧ Л.А. Модель восприятия речи человеком. Новосибирск, "Наука", 1968.

29. ВЛАСОВ В.В. Комплекс аппаратуры для ввода-вывода непрерывных электрических сигналов в ЭВМ БЭСМ-6. - В кн.: Вычислительные системы. Вып. 44, Новосибирск, 1971, с. 132-140.

30. ЕЛКИНА В.Н., ЮДИНА Л.С. Статистика слогов русской речи. - В кн.: Вычислительные системы. Вып. 10, Новосибирск, 1964, с. 58-62.

31. ЕЛКИНА В.Н., ЮДИНА Л.С. Статистика открытых слогов русской речи. - В кн.: Вычислительные системы. Вып. 14, Новосибирск, 1964, с. 55-91.

32. ЕЛКИНА В.Н., ЮДИНА Л.С., ХАЙРЕТДИНОВА А.Г. Статистика двух- и трехфонемных сочетаний русской речи. В кн.: Вычислительные системы. Вып. 37, Новосибирск, 1965, с. 48-74.

33. ЛОЗОВСКИЙ В.С. Анализ и синтез речи на основе Z-операций. - В кн.: Вычислительные системы. Вып. 44, Новосибирск, 1971, с. 102-125.

34. ВЕЛИЧКО В.М., ЗАГОРУЙКО Н.Г. Автоматическое распознавание ограниченного набора устных команд. - В кн.: Вычислительные системы. Вып. 36, Новосибирск, 1969, с. 101-110.

35. ГУСЕВ В.Д., ВОЛОШИН Г.Я. О моделировании спектральных анализаторов для речевых сигналов. - В кн.: Вычислительные системы. Вып. 29, 1967, с. 65-68.

36. ЗАГОРУЙКО Н.Г. Алгоритм распознавания фонем по последовательности сегментных решений. - В кн.: Труды "АКСО-19", Киев-Канев, 1968, с. 135-143.

37. ЗАГОРУЙКО Н.Г., ВЕЛИЧКО В.М., ВОЛОШИН Г.Я., ГУСЕВ В.Д., ЕЛКИНА В.Н., БАХМУТОВА И.В., ХАЙРЕТДИНОВА А.Г., ЮДИНА Л.С. Эксперименты по автоматическому распознаванию речевых сигналов. - Там же, с. 143-147.

38. ВОЛОШИН Г.Я. Об одном использовании языковой избыточности для повышения надежности автоматического распознавания речевых сигналов. - В кн.: Вычислительные системы. Вып. 28, Новосибирск, 1967, с. 21-48.

39. ХАЙРЕТДИНОВА А.Г. Метод автоматического выделения ударения в потоке речи. - В кн.: Вычислительные системы. Вып. 44, Новосибирск, 1971, с. 126-131.

40. ВЕЛИЧКО В.М., ЗАГОРУЙКО Н.Г. Автоматическое распознавание 200 устных команд. - В кн.: Вычислительные системы. Вып. 37, Новосибирск, 1969, с. 75-76.

41. ХАЙРЕТДИНОВА А.Г. Экстремальная проверка алгоритма распознавания слов в слитной речи с использованием информации об ударении. - В кн.: Труды УШ Всесоюзной акустической конференции. М., 1973, с. 147-152.

42. ВИТНЕВ Е.Е., ГАВРИЛКО Б.П., ЗАГОРУЙКО Н.Г., САМОХВАЛОВ К.Ф. Требования к алгоритмам предсказания. - В кн.: Вычислительные системы. Вып. 50, Новосибирск, 1972, с. 100-105.

43. САМОХВАЛОВ К.Ф. О теории эмпирических предсказаний. - В кн.: Вычислительные системы. Вып. 55, Новосибирск, 1973, с. 3-35.

44. GOODMAN N. Fact, Fiction and Forecast (second edition). Indianapolis - New York - Kansas City, 1965.

45. МАМЧУР Е.А., ОВЧИННИКОВ Н.Ф. Принципы простоты и симметрии. - "Природа", 1968, № 6, с.

46. ГАВРИЛКО Б.П., ЗАГОРУЙКО Н.Г. Универсальный алгоритм эмпирического предсказания. - В кн.: Вычислительные системы. Вып. 55, Новосибирск, 1973, с. 134-138.

Поступила в ред.-изд. отд.
5 февраля 1974 года