

РОССИЙСКАЯ АКАДЕМИЯ НАУК
САНКТ-ПЕТЕРБУРГСКИЙ ИНСТИТУТ ИНФОРМАТИКИ
И АВТОМАТИЗАЦИИ

НА ПРАВАХ РУКОПИСИ

ХАРИНОВ
МИХАИЛ ВЯЧЕСЛАВОВИЧ

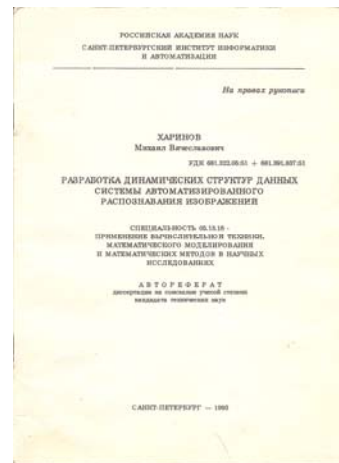
УДК 681.322.05}51 + 681.391.837}51

РАЗРАБОТКА ДИНАМИЧЕСКИХ СТРУКТУР ДАННЫХ
СИСТЕМЫ АВТОМАТИЗИРОВАННОГО РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

СПЕЦИАЛЬНОСТЬ 05.13.16 -
ПРИМЕНЕНИЕ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ,
МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ
И МАТЕМАТИЧЕСКИХ МЕТОДОВ В НАУЧНЫХ
ИССЛЕДОВАНИЯХ

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

САНКТ-ПЕТЕРБУРГ - 1993



Работа выполнена в Санкт-Петербургском институте информатики
и автоматизации РАН

Научный руководитель -
доктор технических наук, профессор В. В. АЛЕКСАНДРОВ

Официальные оппоненты:

доктор технических наук, профессор В. А. ТОРГАШЁВ,
кандидат технических наук И. О. ЛУТОВ

Ведущая организация: Санкт-Петербургский государственный
университет

Защита диссертации состоится " " 1993 года
в часов на заседании специализированного совета Д.003.62.01
при Санкт-Петербургском институте информатики и автоматизации РАН
по адресу: 199178, Санкт-Петербург, 14 линия, д. 39

С диссертацией можно ознакомиться в библиотеке специализиро-
ванного совета Д.003.62.01

Автореферат разослан " " 1993 г.

Ученый секретарь
специализированного совета
кандидат технических наук

В. Е. МАРЛЕЙ

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. В области автоматизации распознавания изображений в последние годы наметилась тенденция интенсивного развития комплексных решений, в которых рассматривается полная цепочка преобразований данных от исходной матрицы яркостей до идентификаторов наблюдаемых объектов. В дополнение к известным новым комплексным решениям имеют самостоятельное теоретическое значение и необходимы для создания автоматизированных программных и аппаратных средств многоцелевого распознавания, настраиваемых посредством обучения на обработку видеообъектов различной природы (аэрокосмических объектов, символической картографической и текстовой информации и др.). Особый интерес представляет разработка программных систем автоматического распознавания, построенных по принципу самоорганизации и обеспечивающих обработку различного типа изображений без использования управляющих параметров, эталонов, ориентированных на узкие прикладные задачи алгоритмов и другой априорной информации, ограничивающей область применения.

Формированию подходов к рассмотрению взаимосвязанных этапов автоматизированного распознавания посвящены работы С.Л. Танимото, М.-К. Ху, А. Розенфельда, А. Клингера, В.В. Александрова, Ю.И. Журавлёва и др. Реализация алгоритмических преимуществ применения различных решений на последовательных стадиях обработки ограничивается применяемым способом организации данных.

В настоящей работе строится общая алгебраическая схема процесса распознавания видеoinформации с учётом морфологической многоуровневости изображения и предлагаются адекватные динамические структуры данных, обеспечивающие согласованное применение элементов перечисленных подходов в сочетании с наглядной интерпретацией.

Цель работы. Целью диссертационной работы является разработка прототипа обучаемой программной системы многоцелевого автоматического распознавания на основе комплексного решения проблемы выделения, признакового анализа и идентификации различного типа объектов полутоновых и двухградационных изображений.

Распознавание в обучаемой системе многоцелевого назначения предполагает перечисление наблюдаемых объектов различных морфологических уровней изображения. В условиях ограниченных ресурсов ПЭВМ описание целых объектов различного масштаба строится в оперативной памяти при локальном (построчном) сканировании изображения, что требует формализации операций с наборами анализируемых точек. Распознавание сводится к выбору из всевозможных наблюдаемых объектов представителей определённого типа и нахождению имён обнаруженных объектов по результатам признакового анализа. Тип идентифицируемых объектов определяется при обучении. Подобно выделению объектов, обучение организуется на основе перечисления.

Общая задача работы сводится к построению алгебраической модели извлечения, преобразования и накопления видеоданных на взаимозависимых стадиях распознавания, что достигается конструктивным решением следующих конкретных задач:

1) Разработка методов многоуровневой сегментации полутонового изображения для адаптивного выделения объектов без искажения формы. Обобщение методов на случай двухградационных изображений.

2) Формализация понятия объекта на основе аппроксимации минимальным числом связных областей различных уровней сегментации изображения.

3) Разработка динамической структуры данных, обеспечивающей по сравнению с матричной эффективное описание операций со связными областями.

4) Построение системы аналитического и структурного (частичного) анализа интегральных геометрических признаков произвольных объектов по свойствам симметрии независимо от ориентации и масштаба.

5) Разработка механизма запоминания и автоматизированной идентификации видеообъектов по повторяющимся значениям признаков.

6) Экспериментальное обоснование разработанных решений.

Научная новизна. В процессе решения поставленных задач получены следующие новые научные результаты, выносимые на защиту:

1) Многоуровневая модель выделения объектов, включающая – метод порождения системы вложенных разбиений полутонового изображения посредством итеративного слияния связных областей близкой яркости;

- метод циклической детализации изображения по локальным яркостным экстремумам, сводящийся к разделению связанных областей на зоны трёх типов без использования управляющих параметров;

- формальное определение объекта, образуемого максимальными связными областями из имеющих характеристики в установленных диапазонах.

2) Динамическая структура данных, в которой для оптимизации расчётов многоуровневое разбиение строится в виде дерева со связями, индексированными номерами уровней, соседство областей

задаётся редуцируемой таблицей связности, а в качестве признаков используются функции от характеристик, аддитивных относительно слияния областей.

3) Система анализа интегральных геометрических признаков объектов, обеспечивающая выбор аналитических определений метрических характеристик невырожденных дискретных фигур на основе преобразования фигур с сохранением симметрии в изотропные по измеряемому в масштабах собственных координатных осей среднеквадратичному линейному размеру.

4) Структурное расширение признаковой системы за счёт разделения фигуры в собственной системе координат на равновесные части, обеспечивающее экспоненциальное возрастание числа признаков при линейном возрастании объёма данных и устойчивости признаков значений при изменении условий съёмки.

5) Способ ассоциативного обучения системы распознавания посредством насыщения списками идентификаторов объектов ограниченного объёма оперативной памяти, предусматривающий автоматизацию обучения на основе интерполяции составных символьных обозначений и сокращение объёма памяти за счёт алгебраической организации обучающей информации.

Практическая ценность работы. Практическая ценность предложенной модели последовательного преобразования видеоданных состоит в том, что она может использоваться для развития специализированных программных и аппаратных реализаций (чтения печатных шрифтов, анализа чертежей, схем, биомедицинских снимков, идентификации лиц и др.), а также для снижения роли интерактивного управления в автоматизированных системах обработки изображений.

Разработанные и проверенные на опыте структуры данных позволяют проектировать, экспериментально исследовать и

реализовывать программные системы автоматического распознавания полутоновых и двухградационных изображений, ориентированные на решение нескольких задач и задач комплексного типа.

Важное значение имеет возможность реализации предложенных решений в условиях ограниченной памяти и быстродействия ПЭВМ, что обеспечивается линейной зависимостью используемых ресурсов от объёма входной информации.

С точки зрения автоматизации исследований имеет самостоятельное значение использование реализованных решений для экспериментального изучения систем признаков, обеспечивающих адаптацию программно-алгоритмического обеспечения к условиям конкретных прикладных задач.

Реализация и внедрение. С целью автоматизации экспериментальных исследований по теме диссертации автором на базе ЭВМ РС/286 реализованы:

- пакет прикладных программ многоуровневой сегментации для выделения объектов на полутоновых изображениях размером 256*256 при 256 градациях яркости;

- прототип системы автоматического распознавания объектов на сегментированных полутоновых и двухградационных изображениях.

Проведённые в СПИИ РАН многочисленные эксперименты с аэрокосмическими снимками и с изображениями, полученными с помощью телекамеры, а также введёнными посредством сканера, обеспечили верификацию разработанных теоретических положений.

Внедрение работы в ПО "Севморгеология" позволило расширить возможности применяемого программно-аппаратного комплекса "Pericolor" и улучшить качество обработки глубоководных снимков.

Разработанные программные системы эксплуатируются в в/ч 41513, 32103.

Представленные к защите результаты являются составной частью работ, проводимых в СПИИ РАН по темам:

- 1) "Создание и ввод в эксплуатацию специализированной автоматизированной системы обработки изображений на основе пирамидально-рекурсивных структур данных" (НТП 0.80.03, задание 07.05).

- 2) "ЛАМПАС-ГКНО" (НИР, выполняемая с в/ч 4153, 32103 согласно договору между научно-техническим центром "Наука" Миннауки России и СПИИ РАН на основании решения Государственной комиссии по ВПВ от 11.09.90 № 264).

3) "Разработка моделирующих инструментальных средств нелинейной системостатики" (НТП "Информатизация России", проект 1.4.1).

4) "Разработка интеллектуальных систем анализа и синтеза визуальной информации на основе рекурсивных структур" ("Перспективные информационные технологии", проект № 539, шифр - "Эволюция").

АПРОБАЦИЯ РАБОТЫ. Основные результаты диссертационной работы докладывались на 18 Дальневосточной математической школе-семинаре (Находка, 1988 г.), 2 и 3 Республиканских семинарах "Проблемы создания систем обработки, анализа и распознавания изображений" (Ташкент, 1989, 1990 г.г.), 3 Всесоюзной конференции "Автоматизированные системы обработки изображений" (Ленинград, 1989 г.), 2 Всесоюзной конференции "Искусственный интеллект-90" (Минск, 1990 г.), 2 совещания по проблемам оцувствления роботов в ИФП АН СССР (Ленинград 1990 г.), 2 Всесоюзной конференции "Оптическое изображение и регистрирующие среды" (Ленинград, 1990 г.).

Программные реализации экспонировались на 2 и 3 Ленинградских научно-технических коммерческих выставках, проведённых ЛИИАН в 1990, 1991 г.г.

ПУБЛИКАЦИИ. По теме диссертации опубликовано 12 научных работ. В том числе научно-технический отчёт.

ОБЪЁМ И СТРУКТУРА РАБОТЫ. Диссертация состоит из введения, 4 глав, заключения, списка литературы из 95 наименований, и приложений. Объём работы составляет 145 страниц основного текста и содержит 29 иллюстраций.

СОДЕРЖАНИЕ РАБОТЫ

Во введении содержится обоснование актуальности темы диссертации, сформулированы основные научные результаты, выносимые автором на защиту, а также практическая ценность полученных результатов.

В первой главе оценивается современное состояние проблемы автоматизированного распознавания изображений, даётся интерпретация постановки задачи, конкретизируется подход к решению. В качестве экспериментального обоснования перспективности исследований приводятся результаты применения базового алгоритма выделения объектов на полутонных изображениях, обобщение которого обеспечило построение многоуровневой модели.

Термин "распознавание" трактуется в применении к объектам, предъявленным впервые.

Изображение представляется в виде некоторого множества предварительно вычисляемых связных областей, из которых строятся объекты. Связные области описываются числовыми признаками. Значения признаков объектов, на распознавание которых настраивается система, фиксируются в процессе обучения на примерах. Диапазоны значений признаков определяют область допустимых значений (ОДЗ), в пределах которой (схема рис.1) производится идентификация объектов и их элементов, наблюдаемых на различных морфологических уровнях.

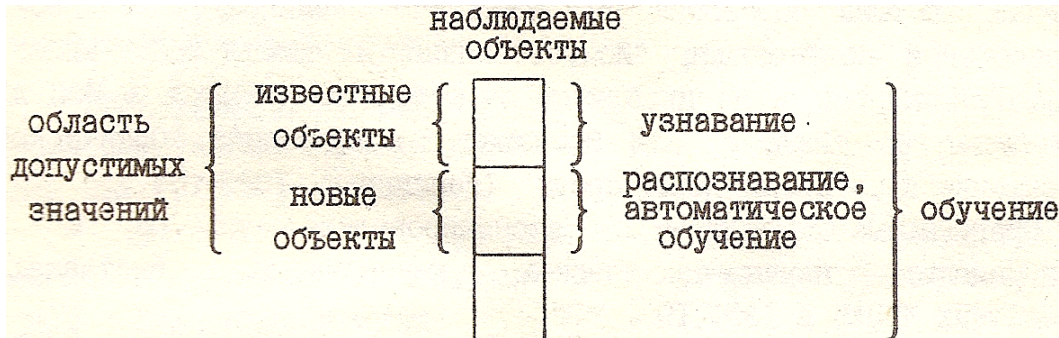


Рис. 1. Задача распознавания видеообъектов

Решение задачи автоматизированного распознавания изображений сводится к последовательному решению трёх подзадач: выделения, структурного анализа и идентификации объектов.

Под выделением объектов понимается выбор максимальных по площади из числа удовлетворяющих ОДЗ. При этом эффективная аппроксимация объектов связными областями различной формы и масштаба достигается благодаря специальной организации данных. После восстановления матричной формы представления производится структурный (частичный) признаковый анализ объекта посредством разделения на равновесные элементы, что необходимо для эффективной детерминированной идентификации. При идентификации находятся неоднозначные обозначения частей объекта в виде списков возможных имён, что отражает характерное для автоматической подготовки данных совпадение признаков. Затем вычисляется пересечение всех обозначений, сопоставляемых различным частям по каждому из признаков.

Ассоциативное обучение без накопления сочетаний значений признаков ограничивается запоминанием имён объектов в сравнительно малом объёме памяти.

Допускается, что обучаемая техническая система может использоваться для распознавания объектов, которые не рассматривались при разработке программного обеспечения. При этом общая стратегия решения состоит в выделении объектов без искажения геометрических свойств, анализе независимо от искажений и детерминированном распознавании (идентификации) по повторяющимся признакам.

В силу большого числа исходных точек изображения по сравнению с количеством наблюдаемых объектов выделение объектов является самой трудоёмкой процедурой и строится на основе простейших из эффективных алгоритмов локальной обработки. В качестве базового в работе предложен алгоритм адаптивной сегментации полутоновых изображений, обеспечивающий выделение объектов, устойчивое к последствиям изменений условий съёмки, в том числе к сдвигу, изменению ориентации и масштаба. Он относится к классу алгоритмов сегментации посредством наращивания областей и сводится к поэтапному слиянию соседних связанных областей, наиболее близких по средней яркости. При этом результат обработки представляется в виде последовательности изображений (уровней) с убывающим количеством областей различной формы и размеров. Число уровней заранее не нормируется. Характерно, что области на том или ином уровне геометрически совпадают с объектами (частями объектов) исходного изображения.

Особенностью базового алгоритма сегментации является то, что он проводится итеративно. При этом все связанные области рассматриваются равноправно, а результат многоуровневой сегментации не зависит от порядка сравнения по яркости пар смежных областей.

Традиционная интерпретация результатов итеративной сегментации, ориентированная на нахождение уровня наилучшей аппроксимации объектов, не позволяет полностью реализовать возможности базового алгоритма, так как в общем случае при обработке реальных изображений оптимальное выделение объектов (элементов объектов) достигается на разных уровнях. Кроме того, в матричной форме представления изображения многократное выполнение относительно долгой операции сегментации неприемлемо без применения спецпроцессора.

Указанные недостатки преодолеваются благодаря формализации понятия объекта, аппроксимируемого сегментами различных уровней, и адекватной организации данных, позволяющей работать со связными областями и их взаимосвязями как с целостными единицами.

Во второй главе описывается многоуровневая модель изображения. Даются основные понятия и определения, описывается динамическая структура данных, алгоритмы её генерации и результаты использования в задаче выделения объектов, рассматривается итеративный метод анализа полутонных изображений по яркостным экстремумам и обобщение модели на случай двухградационных изображений.

Алгебраическое определение объекта в многоуровневой модели обеспечивает его аппроксимацию минимальным количеством составных частей и состоит в следующем.

Рассматривается покрытие изображения связными областями различной формы и размеров в виде совокупности неповторяющихся областей различных уровней, полученных в базовом алгоритме. Пересечение любых двух областей сводится к тому, что одна включает другую.

Из рассматриваемой совокупности областей выделяется набор S , признаки которых лежат в заданной области допустимых значений.

Подмножество областей s набора S , не вложенных в более крупные области из S , считается множеством частей (элементов) искомым объектом. При этом элементы, оказавшиеся смежными, объединяются в сложные объекты, состоящие из нескольких частей.

Формирование составных объектов можно представлять наглядно, последовательно проектируя выбранные области на плоскость при условии, что большие замещают меньшие.

Эквивалентные вычисления организуются на основе представления каждой области в виде некоторого дерева и описания соседства областей друг с другом бинарной симметричной таблицей связности, которая задаётся перечислением для каждой области соседних.

Особенностью многоуровневой модели является то, что слияние областей изображения описывается адекватными преобразованиями таблицы связности и системы деревьев. При этом каждой из N областей исходного изображения сопоставляется дерево, состоящее из единственного корневого узла. Слияние деревьев производится посредством связывания корневых узлов. Преобразование таблицы связности сводится к объединению строк и симметричных им столбцов по логическому закону ИЛИ, причём при объединении каждой пары одна из компонент вычёркивается. Преобразование таблицы в ЭВМ осуществляется в виде эквивалентных операций со списками областей.

Описание слияния областей в виде объединения деревьев и применение редуцируемой таблицы связности обеспечивают быструю

генерацию уровней, которая сводится к поочерёднему слиянию близких по установленному критерию смежных областей (деревьев), определяемых по таблице связности и преобразованию таблицы связности, отражающему слияние областей. При этом предполагается, что смежные области сравниваются по значению средней яркости или другой функции от величин, аддитивных относительно слияния.

В простейшем случае последовательная генерация уровней производится до тех пор, пока все деревья не сливаются в одно совокупное дерево, отвечающее области всего изображения. В процессе генерации для каждой связи i запоминается номер уровня ϵ_i , при формировании которого эта связь была установлена (рис. 2), и любой

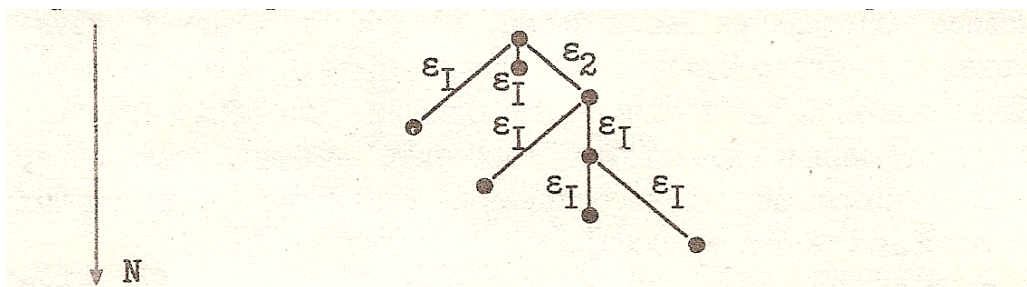


Рис. 2. Описание уровней посредством индексации связей

из них оказывается возможным восстановить без преобразования таблицы связности (за 1-2 сек.). Для этого достаточно разорвать все связи с индексами, превышающими заданный. Тогда результирующее дерево распадается на систему деревьев, задающую искомое разбиение.

Благодаря индексации связей дерева областей достигается компрессированное хранение последовательности уровней изображения и реализуется алгоритм синтеза объектов с подсчётом числа составных частей. Как показывает опыт, это число устойчиво относительно изменений условий съёмки и согласуется с наблюдаемым.

На основе применения разработанной структуры данных в работе развивается метод анализа изображений по яркости, полученный в результате изучения изменения критерия слияния областей. Метод сводится к итеративному наращиванию локальных экстремумов за счёт присоединения к ним новых, которые находятся при условии исключения исходных. В силу того, что новые локальные экстремумы группируются рядом с одноимёнными исходными, процесс наращивания последних продолжается до тех пор, пока все области, не объединённые с исходными экстремумами, не окажутся изолированными друг от друга. При этом изображение разделяется на области максимумов, минимумов и области смешанного типа.

На реальных полутоновых изображениях области смешанного типа, как правило, занимают незначительную долю площади. При этом результаты представления изображения в трёх градациях яркости подобны результатам двухградационного преобразования с локальным выбором яркостного порога. Однако, по сравнению с известными методами обработки с автоматическим вычислением порога, достигается лучшее качество в более широкой области применения.

Метод итеративного анализа изображений по яркости является обобщением идеи разбиения на зоны трёх типов без использования параметров. В силу того, что предложенная структура данных допускает описание изображения любой формы, каждую из областей укрупнённых локальных экстремумов можно рассматривать как отдельное изображение, достигая необходимой детализации. Для обработки всех областей одновременно достаточно упростить таблицу связности, стирая биты информации, регистрирующие соседство элементарных областей различных типов яркости.

Эффективное расширение области применения многоуровневой модели выделения объектов на случай двухградационных изображений и её принципиальное развитие как модели распознавания достигается посредством формализации геометрических признаков объектов в виде дискретных фигур, образованных некоторыми распределениями точек.

В ТРЕТЬЕЙ ГЛАВЕ на основе определения изотропной фигуры и введения связанной с фигурой собственной системы координат строятся признаки, описывающие ориентацию, линейный размер и другие геометрические свойства произвольной дискретной плоской фигуры, рассматривается метод структурного анализа посредством разделения на фрагменты независимо от ориентации и масштаба.

Система признаков вводится на основе рассмотрения линейного Однородного невырожденного преобразования дискретной плоской

фигуры в метрически изотропную.

Фигура, заданная n точками прямой считается вырожденной.

Если среднеквадратичные продольный и поперечный размеры измеряются вдоль собственных осей координат, определяемых по методу наименьших квадратов, то неизотропной считается фигура, длина которой не совпадает с шириной. Всякая фигура из n точек называется изотропной, если среднеквадратичное расстояние от точек до прямой, проведённой через центр инерции, не зависит от наклона прямой.

В двумерном евклидовом векторном пространстве $u \times v$, порождённом линейными комбинациями столбцов одноимённых координат, вычисленных относительно центра инерции, изотропная фигура представляется парой ортогональных векторов равной длины.

Преобразование $W \cdot L$ произвольной неизотропной невырожденной фигуры в изотропную в пространстве $u \times v$ записывается в виде

$$\begin{pmatrix} W\{u\}, W\{v\} \\ W \cdot L\{u\}, W \cdot L\{v\} \end{pmatrix} \cdot \begin{cases} W \cdot L\{u\} = e^\theta \cdot (u \cdot \cos \varphi + v \cdot \sin \varphi) \\ W \cdot L\{v\} = e^{-\theta} \cdot (-u \cdot \sin \varphi + v \cdot \cos \varphi) \end{cases}$$

где посредством преобразования W достигается ортогонализация векторов u, v за счёт поворота исходной плоскости на угол θ , а преобразование L при $\theta \neq 0$ обеспечивает выравнивание длин u, v за счёт деформации вдоль собственных осей фигуры (в том числе - осей симметрии, по числу которых изотропная фигура не уступает исходной).

В силу того, что разложение $W \cdot L$ в произведение поворота V и положительно-определённого симметричного преобразования L почленно совпадает с исходными преобразованиями фигуры, то исходную плоскость и пространство $u \times v$ при рассмотрении преобразования фигуры в изотропную допустимо считать одним множеством векторов.

Упрощение формализма достигается введением в качестве независимых параметров величин $\varphi_1 - \varphi_2$ и γ :

$$\cos(\varphi_1 - \varphi_2) = \frac{(u, v)}{\sqrt{(u, u)(v, v)}}, \quad sh\gamma = -\frac{1}{2} \left\{ \sqrt{\frac{(u, u)}{(v, v)}} - \sqrt{\frac{(v, v)}{(u, u)}} \right\},$$

где $(u, u), (v, v)$ - вторые моменты, часто обозначаемые I_x, I_y, I_z .

Тогда невырожденные фигуры характеризуются $\sin(\varphi_1 - \varphi_2) \neq 0$, а

неизотропные - ненулевой величиной $d = \sqrt{\cos^2(\varphi_1 - \varphi_2) + sh^2 \lambda}$

Для неизотропной фигуры (при $d \neq 0$) из условия ортогонализации

$$\text{векторов находится угол } \varphi : \begin{cases} \sin 2\varphi = \sigma \frac{\cos(\varphi_1 - \varphi_2)}{d} \\ \cos 2\varphi = -\sigma \frac{sh \gamma}{d} \end{cases}.$$

Значения φ берутся в полуоткрытом промежутке от 0 до π , различаются между собой на $\pi/2$ и с точностью до знака задают поворот V и его образ W . Формулы для φ с точностью до угла $\pi/2$ определяют ориентацию собственных осей неизотропной фигуры относительно

исходных, в частности, для симметричной фигуры позволяют установить направление оси симметрии.

Квадраты размеров фигуры независимо от ориентации вычисляются как среднеквадратичные расстояния точек до собственных осей:

$$\text{длина}^2 (\text{ширина}^2) = \frac{\sqrt{(u, u)(v, v)}}{n} (ch\gamma \pm d)$$

Параметр θ , определяемый $\text{th } 2\theta = -\sigma \frac{d}{ch\gamma} \theta$, имеет смысл отношения линейных размеров, измеренных в логарифмическом масштабе:

$$\theta = \frac{1}{2} \sigma \cdot \ln (\text{длина}/\text{ширина}) .$$

Квадрат измеряемого в произвольном направлении линейного размера изотропной фигуры определяет площадь, которая совпадает с площадью исходной фигуры, равной произведению длины и ширины:

$$\text{размер}^2 = \text{площадь} = \frac{\sqrt{(u, u)(v, v)}}{n} \sin(\varphi_1 - \varphi_2) .$$

Наглядное понятие изотропной фигуры и представление фигур в виде пар векторов u, v для промежуточной стадии распознавания заранее неизвестных объектов обеспечивает необходимое уточнение применяемых эвристических определений метрических признаков.

В силу того, что некоторый сдвиг начала координат обеспечивает ортогонализацию и выравнивание длин u, v , а почленное суммирование пар преобразованных векторов не выводит из их множества, представляется перспективным алгебраическое развитие формализма изотропных фигур на основе общих линейных преобразований, учитывающих сдвиг начала координат, индуцируемый объединением фигур. Кроме того, рассмотрение сдвига позволяет уточнить описание особенностей распределения точек вблизи центра инерции фигуры.

Структурное развитие признакового описания строится в рамках пассивного толкования линейного преобразования. При этом благодаря переходу к собственной системе координат с изменением масштабов осей всякая невырожденная фигура представляется изотропной и разделяется на части по образующим полярной системы отсчёта.

Для представления объекта в различном масштабе съёмки неизменным количеством точек предлагается принцип дискретизации изображения, который сводится к итеративному разделению каждой связанной области на заданное число фрагментов, в простейшем случае — на две части, образуемые точками с расстоянием до центра инерции больше и меньше среднего. Процесс дробления каждой образовавшейся

связной области продолжается, если она распадается более, чем на две новые. В противном случае область считается элементарной.

Вариант частичного анализа объектов обеспечивается разделением анализируемой изотропной фигуры на сектора по осям собственной системы координат и биссектрисам координатных углов. В силу того, что смешанный момент для изотропной фигуры обращается в 0, а два остальных совпадают, точки распределяются по секторам равновесным образом. При этом исходная фигура разбивается на октет характерных элементов и, кроме того, представляется восемью точками, отождествляемыми с центрами инерции элементов. Вращение собственной системы координат порождает семейство разбиений, параметризованных значениями угла поворота.

Основу системы признакового описания образуют изотропные образы элементов объекта в виде невырожденных фигур. Различные числовые метрические и другие признаки исходного или вычисляемых дискретных представлений объектов (элементов объектов) в собственных системах координат, тригонометрические и гиперболические параметры преобразований, а также соотношения признаков определяют практические возможности описания объектов. В частности, для решения задач, связанных с юстировкой, полезно использование восьмимерных векторов ориентации элементов относительно объекта в целом. Данные, необходимые для геометрического описания, ограничиваются моментами не выше второго порядка.

Благодаря аддитивности данных относительно объединения областей линейное возрастание их объёма при переходе к частичному описанию обеспечивает экспоненциальное увеличение числа вычисляемых признаков за счёт рассмотрения всевозможных сочетаний элементов объектов, что наряду с параметрическим расширением позволяет полнее использовать применяемые аналитические зависимости.

Для оптимизации расчёты выполняются в целых числах. Необходимые квадраты координат запоминаются после однократного вычисления и используются для нахождения произведений посредством сложения.

Для автоматизации экспериментального исследования систем признаков и создания эффективных средств автоматического распознавания без использования эталонов необходима реализация на ЭВМ библиотеки именованных значений признаков объектов, позволяющей запоминать объекты в виде значений признаков и по признакам вычислять идентификаторы, не прибегая к перебору вариантов.

В ЧЕТВЕРТОЙ ГЛАВЕ описывается способ организации библиотеки объектов и её программная реализация в прототипе обучаемой системы автоматического распознавания.

Библиотека объектов строится на основе требования непротиворечивой автоматической систематизации произвольной априорной информации, поступающей в процессе обучения в виде названий объектов. При адекватной идентификации объектов результаты обучения позволяют судить об эффективности набора признаков, выбранных для описания объектов конкретной предметной области.

Распознавание по повторяющимся признакам организуется при ограничении объёма необходимой оперативной памяти ЭВМ линейной зависимостью от числа признаков, разброса параметров и количества классов идентифицируемых объектов. С насыщением используемого объёма памяти по мере обучения система при предъявлении новых объектов обеспечивает прогноз. Ошибочная идентификация замещается в ходе обучения неоднозначной. Нижняя граница вероятности достоверного распознавания определяется долей распознанных новых объектов, верхняя – долей объектов, идентифицируемых после обучения однозначно. Однозначность идентификации нераспознанных объектов достигается за счёт учёта дополнительных признаков из числа признаков расширенной системы. Циклы выделения дополнительной памяти и её заполнения идентификаторами в процессе обучения продолжают до достижения требуемой достоверности распознавания.

Оценка достоверности по прогнозу отдельных объектов позволяет принимать решение об адекватности выбора признаков на начальных стадиях трудоёмкого процесса интерактивного обучения.

Обозначение объекта определяется в виде списка возможных имён и находится как результат пересечения множеств идентификаторов, ассоциированных со значениями признаков. Все признаки считаются равноправными. В частности, как отдельные признаки рассматриваются компоненты восьмимерных векторов структурного представления объектов. Результирующие обозначения объектов сводятся к суперпозиции пересечений обозначений структурных элементов по всем используемым признакам.

Начальное интерактивное обучение производится по прецеденту. В автоматизированной форме интерактивного обучения используется неоднозначное представление, при котором имя вводится как обозначение некоторого набора объектов. Необходимое представление объ-

екта (элемента объекта) составляют результаты аппроксимации на нескольких последовательных уровнях, либо оно обеспечивается моделированием изменений ориентации, масштаба и др. условий съёмки.

На следующей стадии подготовки системы к автономной работе применяется автоматическое обучение или самообучение.

Самообучение представляет собой моделирование закономерностей распределения имён объектов, характерных для насыщенной памяти, и производится посредством интерполяции составных символьных обозначений (рис.3), причём изображения объектов не используются.

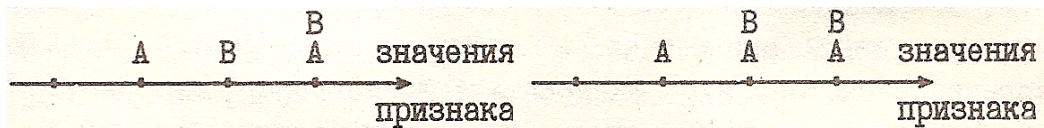


Рис. 3. Интерполяция символьных обозначений объектов.

Автоматическое обучение с использованием обработки изображений производится по критерию однозначности, согласно которому в библиотеку избирательно заносятся имена объектов, обеспечивающие однозначную идентификацию. При интерполяции с учётом критерия однозначности достигается автоматическое формирование диапазонов непрерывного распределения идентификаторов по последовательным значениям признаков (рис.3). При учёте дополнительных признаков полезна идентификация по неполному признаковому набору, применяемая как в режиме автоматического обучения посредством распознавания впервые предъявленных объектов, так и для заполнения идентификаторами известных объектов дополнительно отводимой памяти.

Особый интерес представляет использование библиотек объектов с алгебраической организацией данных, в которых идентификаторы различных сочетаний составных элементов по каждому аддитивному признаку рассматриваются на одном множестве признаков значений, причём сложению признаков отвечает пересечение символьных обозначений (рис.4). При этом достигается дальнейшее сокращение

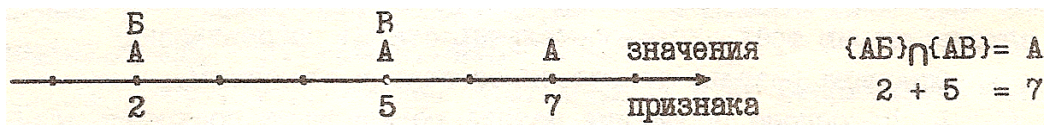


Рис. 4. Алгебраическая организация обучающих данных.

объёма обучающей информации.

Реализация библиотеки сводится к формированию и накоплению списков имён объектов, отвечающих значениям признаков. Списки хранятся в упорядоченном виде (рис.3-4), обеспечивающем оптимизацию вычисления пересечения обозначений за счёт предварительного нахождения пересечения диапазонов идентификаторов.

Как показал опыт, вычисление пересечений обозначений, характеризующих непустыми пересечениями диапазонов имён, при учёте нескольких десятков признаков позволяет проводить идентификацию за время в несколько раз меньшее, чем требуется для выделения и частичного анализа объектов. В работе указывается способ развития структуры данных библиотеки объектов, обеспечивающий дальнейшее ускорение расчётов, что для данных условий решает проблему оптимизации выполнения на ЭВМ теоретико-множественных операций.

Испытание механизмов автоматического обучения и проверка устойчивости системы признаков при изменении условий съёмки проведены в экспериментах по распознаванию печатных шрифтов и рукописных слов.

В перспективе представляет интерес развитие реализованных программных систем в едином комплексе распознавания полутоновых и двухградационных изображений, что при ограниченных ресурсах ЭВМ ЗС/ФЕ-286 связано с использованием вычисляемых дискретных представлений. Принципиальное развитие структур данных обеспечивается обобщённым алгебраическим представлением объектов, которое помимо деревьев и аддитивных характеристик в качестве атрибутов элементов изображения включает составные символьные обозначения (списки формальных символов), преобразуемые при слиянии областей посредством логических операций.

В ПРИЛОЖЕНИЯХ содержится сводка формул и акты внедрения.

ЗАКЛЮЧЕНИЕ

В настоящей диссертации разработана математическая модель изображения и адекватные структуры данных, обеспечивающие автоматизацию трёх последовательных стадий распознавания.

Основные результаты работы состоят в следующем.

1) На основе обобщения метода сегментации построена многоуровневая модель представления изображения, обеспечивающая аппроксимацию наблюдаемых объектов без искажения формы минимальным числом частей в виде связанных областей.

2) Для оптимизации выделения простых и составных объектов разработана динамическая структура данных, в которой многоуровневая сегментация изображения задаётся в виде дерева с индексированными связями.

3) В рамках структуры данных для выделения объектов разработано многоуровневое представление изображения в нескольких градациях яркости методом слияния одноимённых локальных яркостных экстремумов, обеспечивающее улучшение качества пороговой обработки.

4) На основе аналитического метрически изотропного представления невырожденной дискретной фигуры в собственной системе координат формализовано описание интегральных признаков объектов с учётом ориентации и свойств осевой симметрии.

5) На основе разделения фигуры в собственной системе координат на фрагменты построено логическое, параметрическое и комбинаторное расширение признаковой системы, обеспечивающее описание объектов независимо от изменения масштаба и других условий съёмки.

6) Предложен принцип разделения изображения на элементарные связные области, распадающиеся при определённом разбиении в собственной системе координат на минимальное число связных компонент.

7) Для идентификации объектов по повторяющимся признакам предложен способ организации данных в виде ассоциированных с признаковыми значениями списков идентификаторов, предусматривающий автоматизацию обучения согласно критерию однозначности.

8) Показана возможность реализации системы автоматического распознавания видеообъектов при ограниченных ресурсах РС/АТ-286.

9) Результаты работы внедрены в 2 организациях.

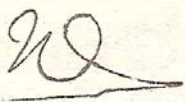
Основные результаты диссертации опубликованы в следующих работах:

1. Александров В.В., Харинов М.В. Представление изображений ранговыми распределениями.- Препринт № 61 - Л.:ЛИИАН, 1988.- 48с.

2. Харинов М.В. Автоклассификация объектов двухградационных изображений на примере вычисления формы // Информационные проблемы распознавания. - Л.: ЛИИАН, 1988.- С.94-107.

3. Искандеров П.М., Слиозберг Н.С., Харинов М.В. Итеративно-рекурсивный метод предварительной обработки полутоновых изображений в задаче выделения конкреций // Там же. - С.107-117.

4. Искандеров П.М.,Харинов М.В. Сегментация полутоновых изображений методом итеративного осреднения // Проблемы создания систем обработки, анализа и распознавания изображений.- Ташкент:ФАН, 1989.- С.25-26.
5. Харинов М.В. Быстрая сегментация полутонового изображения методом итеративного осреднения. // Автоматизированные системы обработки изображений / Тез.докл. 3 Всесоюз. конф.-Л.,1989.-С.85.
6. Харинов М.В. Представление растрового полутонового изображения с помощью однородных по яркости областей // Представление знаний и экспертные системы. - Л.:ЛИИАН, 1989.- С.142-154.
7. Харинов М.В. Сегментация полутонового изображения на основе бескоординатного представления // Проблемы создания систем обработки, анализа и распознавания изображений.-Ташкент:ФАН, 1990.-С.31-42.
8. Александров В.В., Харинов М.В. Представление полутонового изображения с помощью нерегулярного дерева однородных по яркости областей // Искусственный интеллект 90 / Тез.докл. 3 Всесоюз. конф. - Минск: 1990.- т.1 - С.187-190.
9. Харинов М.В. Выделение объектов на полутоновом изображении путём многоуровневой сегментации // Оптическое изображение и регистрирующие среды/ Тез.докл.2 Всесоюз конф.-Л.,1990.,т.2-С.171-172.
10. Харинов М.В. Выявление объектов на основе многоуровневого представления полутонового изображения // Проблемы создания систем обработки, анализа и распознавания изображений.- Ташкент:ФАН, 1990.- С.56
11. Харинов М.В. Представление объектов на основе многоуровневой модели // Структуры данных, алгоритмы и программы обработки сложных изображений.- Препринт № 162 - С.-П.:СПИИРАН,1992. - С. 3-19.



Подписано к печати

Тираж 100, формат бумаги 60x84 1/16, 1.25 печ. л. Бесплатно.

Санкт - Петербург. Типография в/ч

Зак. 221-93