

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

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата технических наук

Научный руководитель
доктор технических наук,
профессор Александров В.В.

Санкт - Петербург
1993

Введение.....	5
Глава 1. Предпосылки комплексного решения проблемы распознавания изображений.....	15
1.1. Теоретические предпосылки комплексного решения.....	15
1.2. Задача адаптивного многоуровневого представления полутонового изображения.....	32
1.3. Базовый алгоритм выделения объектов.....	35
1.4. Цель и задачи работы.....	45
Глава 2. Многоуровневая модель изображения.....	47
2.1. Понятие многоуровневой модели.....	48
2.2. Алгебраическое определение объекта.....	49
2.3. Структура данных для выделения объектов	
2.3.1 Понятие видеоданных.....	52
2.3.2 Представление связанных областей в виде деревьев.....	53
2.3.3 Табличное описание соседства областей.....	55
2.4 Генерация структуры. Обеспечение прямого доступа.....	59
2.5. Формы организации данных. Аппроксимация объектов.....	62
2.6. Метод анализа полутонового изображения по яркости.....	69
Выводы по главе 2.....	75

Глава 3. Геометрическое описание и структурный анализ объектов.....	76
3.1. Аналитическое описание объектов, исходя из свойств осевой симметрии	
3.1.1 Сохранение симметрии при линейных преобразованиях.....	78
3.1.2 Центральные и средние значения координат. Условие координатной симметрии.....	80
3.1.3 Запись в 2-мерном подпространстве n-мерного векторного пространства.....	82
3.1.4 Поворот. Определение ориентации.....	86
3.1.5 Собственные преобразования Лоренца. Приведение к собственным осям.....	89
3.1.6 Интерпретация.....	92
3.1.7 Разложение WWL в произведение ортогонального и симметричного операторов. Отождествление плоскостей.....	98
3.1.8 Характеристика фигуры посредством сдвига системы координат.....	106
3.2. Структурное описание объектов посредством частичного анализа.....	108
Выводы по главе 3.....	118

Глава 4. Распознавание (идентификация) объектов.....	120
4.1. Задача автоматизированной идентификации видеообъектов.....	121
4.2. Библиотека объектов.....	125
4.3. Реализация. Эксперимент. Обучение. Самообучение.....	133
Выводы по главе 4.....	147
Заключение.....	149
Литература.....	153
Приложение.....	164

ВВЕДЕНИЕ

В настоящее время актуальной является проблема создания программных систем и аппаратных комплексов для автоматического распознавания изображений, рассчитанных на решение широкого круга задач при минимальных ограничениях области применения. Наиболее практичными представляются обучаемые программно-аппаратные системы распознавания, обеспечивающие в рабочем режиме выделение, анализ и идентификацию объектов без участия пользователя, а в режиме обучения не требующие от пользователя кроме названия объекта характеристики его свойств как элемента изображения. При этом допускается, что техническая система может использоваться для распознавания объектов, которые не испытывались при разработке.

Проектированию таких систем препятствует недостаточная формализация последовательных этапов распознавания, которая выражается в интенсивном использовании при решении прикладных задач задаваемых извне управляющих параметров, "зашитых" в программы эталонов, алгоритмов, ориентированных на поиск заранее заданных объектов и прочей априорной информации.

С логической стороны наиболее сложным при проектировании универсального типа систем автоматического распознавания является конструктивная формулировка принципа выделения объектов на основе анализа совокупности характеристик, накопленных в виде обучающей информации.

Принципиальное решение достигается в предположении, что для относительно небольшого числа наблюдаемых объектов видеоинформация квантована по ряду признаков (яркости, площади, линейному размеру и др.), причём по некоторым признакам объ-

екты резко различаются между собой. Тогда, при условии независимого решения задачи перечисления всевозможных объектов, для выбора объектов определённого типа достаточно указания диапазонов изменения анализируемых признаков. При этом необходимые оценки диапазонов признаков извлекаются из обучающей информации и являются её простейшими атрибутами.

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 полутонных и двухградационных изображений.

Перечисление видеообъектов, на основе которого производится выделение объектов с характеристиками в определённых диапазонах, достигается благодаря разработке структур данных, адекватных природе изображения.

Кардинальное отличие видеоданных от обычных состоит в избыточном количестве объектов и выражается в морфологической многоуровневости (неоднозначности) изображения, что позволяет накапливать и использовать изображения как самостоятельный источник информации, обеспечивающий решение различных прикладных задач. Неоднозначность интерпретации изображения ограничивается тем, что пересечения объектов визуально не проявляются, наблюдаемые объекты накладываются один на другой. Поэтому элементы видеоданных характеризуются отношением вложения и адекватное исходное представление объектов обеспечивают иерархические структуры.

Основные результаты в области разработки иерархических структур видеоданных принадлежат С.Л. Танимото, А. Розенфельду, А. Клингеру, В.В. Александрову.

В настоящее время наиболее распространены методы выделения объектов (элементов объектов) посредством аппроксимации стандартными аналитическими зависимостями, что в определённой степени связано с искажением формы объектов. Качество аппроксимации заметно улучшается при описании объектов на основе адаптивной многоуровневой сегментации, при которой применяемые для аппроксимации элементы вычисляются в процессе итеративного анализа исходного изображения. Однако, для представления результатов адаптивной многоуровневой сегментации недостаточно разработаны способы адекватной организации данных, что сдерживает её активное применение.

Особенностью разработки структуры данных для адаптивного выделения объектов является необходимость перехода к специальному представлению, отличному от обычной матрицы с элементами, отвечающими точкам или стандартным блокам исходного изображения. Предполагается, что специальное внутреннее представление изображения позволяет сохранить многозначность выделения объектов и обеспечивает оптимизацию расчётов за счёт перехода к операциям с комплексами точек, образующих сегменты различной формы и размеров.

Описание объектов как целостных единиц обеспечивает оптимизацию расчётов на стадии их выделения, но неудобно для последующего признакового анализа в силу того, что общие характеристики наблюдаемых объектов наиболее подвержены влиянию изменений условий съёмки. Поэтому, согласно идее В.В. Александрова, признаковому анализу и идентификации предшествует синтез объектов, в результате которого каждый объект представляется как отдельное изображение. При этом вычисление значений признаков производится после восстанов-

ления объекта в исходном матричном представлении.

В практике признакового анализа объектов в виде плоских дискретных фигур продолжает оставаться актуальным использование интегральных инвариантов, приоритет в области разработки которых принадлежит М.-К. Ху. Примечательно, что эффективное описание объектов обеспечивается при этом средствами стандартной математики на основе рассмотрения линейных преобразований, отвечающих простейшим изменениям условий съёмки.

Разделение распознавания объектов на последовательные этапы перечисления, выделения и идентификации (классификации) выявляет неоднозначность привычных терминов и приводит к необходимости их переосмысления. Такие общепотребительные понятия, как "длина", "ширина", "линейный размер" и др. теряют очевидность, если речь идёт об измерении признаков неидентифицированного объекта. Корректное уточнение метрических и других характеристик дискретных объектов производится из соображений инвариантности симметрии. Интегральные признаки выводятся на основе линейных преобразований, которые в случае симметричных фигур сохраняют осевую симметрию, а в общем случае фигур, отличных от прямых, обеспечивают определение среднеквадратичного линейного размера, единого для всех осей, проходящих через центр инерции. Изотропность размера обеспечивается определённым выбором масштабов связанных с фигурой собственных координатных осей.

Алгебраическое развитие признакового представления объектов посредством введения собственных систем координат достигается в форме структурного (частичного) анализа, основные идеи которого принадлежат Ю.И.Журавлёву. Указанный алгебраический подход широко применяется в практике распознавания об-

разов, в частности, изображений. В области обработки изображений, однако, эффективность результатов применения ограничивается недостаточным использованием наглядной интерпретации, отражающей специфику видеообъектов.

В наглядной интерпретации структурный анализ состоит в представлении видеообъектов наборами некоторых составных частей, вычисляемых независимо от сдвига, поворота, масштаба и других условий съёмки. При этом наблюдаемые объекты представляются посредством элементов адекватной природы, что позволяет на основе сравнения элементов и сочетаний элементов выявить необходимые для идентификации характерные признаки. За счёт комбинирования составных элементов достигается экспоненциальное расширение признаковой системы при возрастании объёма данных пропорционально числу рассматриваемых частей.

Алгебраическое расширение признаковой системы в сочетании с ограничением объёма накапливаемой обучающей информации обеспечивает реализацию идентификации по повторяющимся признакам за счёт насыщения оперативной памяти, в которой систематизируются названия объектов. При этом объекты с совпадающими значениями признаков отождествляются между собой. Ошибочная идентификация новых объектов заменяется неоднозначной. Однозначность достигается посредством использования дополнительных признаков. Организация распознавания на основе насыщения обучающей информации соответствует условиям автоматизированной генерации признакового представления, особенностью которого является повторяемость значений признаков объектов и элементов объектов, что обеспечивает достаточно простую реализацию автоматической адаптивной классификации.

В целом для решения проблемы распознавания применяется

алгебраический подход, который строится на основе наглядной интерпретации и заключается в том, что наряду с формализацией понятия видеообъектов производится также и формализация определённых операций. В качестве элементов представления видеообъектов используются связные области, которые описываются числовыми характеристиками, деревьями, таблицами и наборами символов. Объединению областей сопоставляются адекватные операции с перечисленными компонентами формального описания.

В рамках указанного подхода развивается математическая модель распознавания, представляющая собой общую алгебраическую схему выделения-идентификации видеообъектов и аналитическую систему признаков.

Общая схема описывает механизм последовательного преобразования информации изображения в обозначения объектов, не содержит прямых ограничений области применения и реализуется посредством динамической структуры данных, обеспечивающей выделение объектов, а также структуры данных, предназначенной для запоминания и идентификации объектов на основе обучения по прецеденту. Структуры данных разрабатываются независимо от типа используемых параметров, в том числе, яркостных, геометрических и др. При этом методы, разработанные для конкретных параметров, в рамках предусмотренной схемы обобщаются на другие варианты признакового описания объектов.

Благодаря симметричному рассмотрению признаков аналитическое описание видеообъектов реализуется в виде одного центрального программного модуля, определяющего конкретные результаты обработки на всех стадиях распознавания, а программная реализация системы распознавания в целом обеспечивает автоматизацию экспериментального исследования различных

признаковых представлений.

С целью теоретического и экспериментального решения проблемы выделения, структурного анализа и классификации объектов на полутоновых и двухградационных изображениях в работе решаются следующие конкретные задачи:

1) Построение математической модели многоуровневой адаптивной аппроксимации объектов полутонového изображения связными областями различной формы и размеров, пересечение которых сводится к включению одной области в другую. Обобщение модели на случай двухградационных изображений.

2) Разработка динамической структуры данных, обеспечивающей по сравнению с матричной эффективное описание операций с наборами точек в виде связных областей.

3) Разработка методов структурного анализа изображения по яркостным и геометрическим признакам.

4) Разработка способов организации динамической структуры данных, обеспечивающей запоминание и быструю идентификацию объектов по повторяющимся признакам при линейной зависимости используемого объема оперативной памяти от числа признаков, разброса признаковых значений и количества идентифицируемых классов объектов.

5) Экспериментальное исследование развиваемых технических решений путём обработки реальных изображений на ЭВМ.

В процессе решения поставленных задач получены следующие новые научные результаты, которые выносятся на защиту:

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

- метод циклической детализации изображения по локальным яркостным экстремумам, сводящийся к разделению связных 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 устойчивость признаковых 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.

Благодаря предложенной модели достигается аппроксимация объектов без искажения формы минимальным числом связанных областей по установленным диапазонам значений признаков. Формализация операций с произвольными связными областями обеспечивает оптимизацию расчётов по памяти и по времени. В рамках адекватной организации данных достигается улучшение качества порогового преобразования изображений по яркости и развиваются методы анализа объектов по геометрическим признакам. Для распознавания объектов по повторяющимся признакам реализуется способ непротиворечивой систематизации в памяти ЭВМ произвольной обучающей информации, при котором обеспечивается автоматическое формирование диапазонов признаковых значений, необходимых для управления выделением объектов.

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 решений в условиях ограниченной памяти и быстродействия ПЭВМ, что обеспечивается линейной зависимостью используемых ресурсов от объёма входной информации.

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 СПИИРАН многочисленные эксперименты с аэрокосмическими снимками и с изображениями, полученными с помощью телекамеры, а также введёнными со сканера, обеспечили верификацию разработанных теоретических положений. Внедрение работы в ПО "Севморгеология" позволило расширить возможности применяемого 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: "Создание и ввод в эксплуатацию специализированной автоматизированной системы обработки изображений на основе пирамидально-рекурсивных структур данных" (НТП 0.80.03, задание 07.05), "ЛАМ-ПАС-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), "Разработка моделирующих инструментальных средств нелинейной системостатики" (НТП "Информатизация России", проект 1.4.1), "Разработка интеллектуальных систем анализа и синтеза визуальной информации на основе рекурсивных структур" ("Перспективные информационные технологии", проект № 539, шифр - "Эволюция").

ГЛАВА 1. ПРЕДПОСЫЛКИ КОМПЛЕКСНОГО РЕШЕНИЯ ПРОБЛЕМЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

В аспекте разработки комплексного решения проблемы распознавания и создания обучаемых систем автоматического распознавания даётся оценка состояния современной теории и практики обработки изображений. Конкретизируется постановка проблемы и подход к решению.

Анализируются особенности распознавания видеоданных.

На основе разработки, экспериментального исследования и обобщения интерпретации итеративного алгоритма многоуровневой сегментации предлагается вариант адаптивного объектно - ориентированного представления изображения, в рамках развития которого формулируются основные цели и задачи диссертационной работы.

1.1. Теоретические предпосылки комплексного решения.

В настоящее время автоматизированная обработка изображений применяется во многих областях научных исследований, включая математику, физику и разнообразные области прикладных наук, в которых использование видеоинформации открывает новые возможности для решения задач и создаёт благоприятные условия развития. Удачные приёмы использования обработки изображений, найденные в одной из областей знаний, часто оказываются применимыми в других и трактуются как универсальные, хотя задачи сводятся к рассмотрению изображений объектов сходного вида. При этом многочисленные результаты применения обработки изображений в различных пред-

метных областях сами по себе не составляют достаточно убедительной теории распознавания изображений как самостоятельной научной дисциплины.

Хотя постановка задачи распознавания на первый взгляд предельно проста и сводится к моделированию на ЭВМ способности человека запоминать указанные объекты и затем их узнавать, тем не менее принципы решения этой задачи на современном этапе находятся в стадии формирования. Большинство приложений ориентировано на участие человека, для которого комфортные условия восприятия могут не совпадать с оптимальными формами преобразования информации в ЭВМ. Так, предварительные преобразования фильтрации в виде реставрации, улучшения качества изображений [48,70] могут не иметь принципиального значения при продолжении обработки на ЭВМ. Классические методы распознавания образов [56,57] могут оказаться не оптимальными при автоматизированной подготовке данных признакового описания объектов. Таким образом, совокупность технических решений, повышающих эффективность распознавания человека, не обеспечивает автоматизацию процесса распознавания в целом и для формирования элементов теории необходим комплексный подход, при котором в полной мере проявляется специфика проблемы автоматизированной обработки изображений.

Комплексное рассмотрение процесса распознавания предполагает согласованное решение задач автоматизации на нескольких последовательных стадиях обработки изображения, достигаемое при минимально возможном использовании априорной информации в виде эталонов, задаваемых извне параметров и т.д. В идеальном случае вся поступающая в систему априорная информация сводится к названиям объектов, а необходимые огра-

ничения вычисляются по мере накопления данных при обработке изображений объектов, на распознавание которых настраивается система. Принципиальная возможность выполнения указанного условия до сих пор вызывает споры, что свидетельствует о недостаточном уровне формализации постановки проблемы.

Недостаточная формализация распознавания, вероятно, связана прежде всего с не вполне адекватным сопоставлением математических и реальных понятий. При этом точные математические формулировки задач и решений [9, 35] оказываются относительно громоздкими и существенно менее доступными, чем словесное описание. Таким образом, для моделирования простейших функций распознавания особое значение имеет удачная интерпретация постановки и решения задач, обеспечивающая достижение цели средствами стандартной математики. При моделировании более сложных функций распознавания адекватная интерпретация [52-55] может усилить эффект применения развиваемого нового математического аппарата алгебры алгоритмов [26-28], алгебры образов и теории нечётких множеств [44, 86, 90, 92].

Задачи современной практики обработки изображений связаны с моделированием первоначальных этапов мышления на уровне интуитивного восприятия, что порождает ряд методологических трудностей субъективного характера.

Одна из основных трудностей сводится к осознанию акта наблюдения как процесса, состоящего из нескольких необходимых стадий подготовки, анализа и классификации данных. Для выявления минимального числа этапов обработки оказывается полезным переосмысление интерпретации механизма работы применяемых эвристических алгоритмов в сочетании с моделированием особенностей восприятия [41].

Другая трудность связана с изменением множества обрабатываемых данных в процессе распознавания на ЭВМ. Автоматизированное распознавание характерно тем, что включает по крайней мере две классификации, вложенные одна в другую. Первая классификация состоит в объединении точек в объекты, вторая — в идентификации объектов, что сводится к нахождению их обозначений в виде названий или имён. Группировку точек при выделении объектов и идентификацию или распознавание объектов не всегда разделяют между собой. Например, при поиске объектов методом сравнения с эталоном [4, 5, 72–75, 93] выделение и распознавание объектов производятся одновременно, что достигается благодаря сужению возможной области приложения задачами поиска объектов фиксированного масштаба и влечёт необходимость оптимизации выполнения перебора вариантов. Разделение процедур классификации [77] позволяет строить распознавание на основе анализа информации не об ожидаемых, а об имеющихся объектах, что позволяет расширить рассматриваемую предметную область.

В силу того, что исходных точек на изображении обычно гораздо больше, чем наблюдаемых объектов, группировка точек при выделении объектов по сравнению с идентификацией является более длительной процедурой, особенно, если она связана с выполнением при поточечном анализе сложных вычислений. Идентификация (классификация) объектов строится на основе более сложного логического анализа данных о выделенных объектах.

Необходимая для классификации объектов подготовка данных может достигаться в процессе выделения объектов и этим ограничиваться. Так, при распознавании по эталонам под данными понимаются сами объекты в виде наборов точек. В методах

структурного анализа [3, 21, 49, 50, 58, 59] под данными понимаются образы в виде стилизованных объектов или их признаковых описаний, выраженные в терминах структур достаточно сложной природы (примитивы, таблицы, графы, гистограммы, спектры и т.д.). Указанные структуры являются "полуфабрикатами" обработки, формируются как некоторый эквивалент изображения в терминах стандартных производных элементов и отражают многообразие свойств объектов (отношений объектов). Сложная природа структур данных, применяемых для выделения объектов, связана с необходимостью сохранения информационной избыточности изображения относительно каждого из возможных запросов при условии оптимизации процесса выполнения текущих запросов по времени. При этом оптимизация по времени достигается за счёт того, что специальные структуры данных позволяют оперировать с некоторыми комплексами точек, а информационная избыточность обеспечивает возможность однозначного объектного описания изображения.

Непосредственное использование результатов выделения объектов в качестве входных данных для классификации требует в каждом конкретном случае разработки особого математического и алгоритмического аппарата и ограничивает возможности их систематизации с целью оптимизации классификации. Поэтому структурный анализ имеет смысл рассматривать в качестве самостоятельной промежуточной стадии преобразования данных.

Наиболее удобный подход к структурному анализу используется в методе голосования [29, 30], в наглядной интерпретации которого объект представляется набором составных частей, оценивается сходство каждой части с частями известных объектов и в результате порогового анализа оценок строится

прогноз. При этом объект описывается в терминах компонент адекватной сложности, что упрощает язык описания и позволяет представить объект в виде числового вектора данных.

Согласно идее В.В. Александрова [1,6,8] распознавание изображения представляется непрерывным процессом чередования анализа и синтеза объектов, где под синтезом понимается интерпретация группировки точек в объекты в терминах исходной матрицы, для которой устанавливается принадлежность исходных точек определённым объектам. Если переход к внутренней структуре данных при выделении объектов согласно текущему запросу обеспечивает возможность восстановления адекватного однозначного разбиения изображения, то структурный анализ производится вслед за выделением объектов посредством обработки исходной матрицы. Проведение структурного анализа объектов независимо от внутренней формы организации данных для выделения объектов обеспечивает схему преобразования данных рис. 1, в которой однозначные преобразования составляют логически естественную цепь обработки.

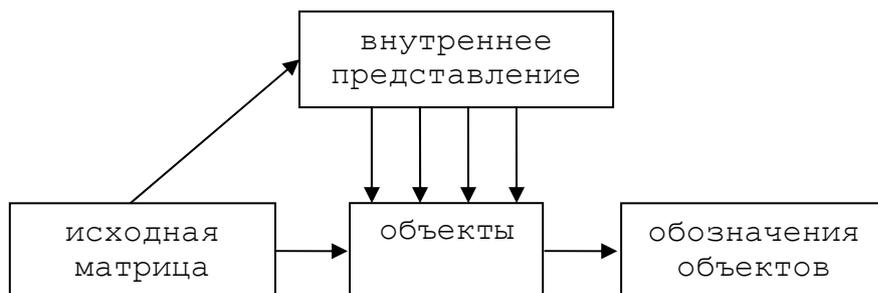


Рис. 1. Схема автоматизированного распознавания изображения

Благодаря реализации структурного анализа как самостоятельной процедуры между операциями выделения объектов и идентификации, признаки при выделении и анализе объектов рассматриваются на основе единой формы представления данных,

что позволяет строить программную реализацию системы, в которой вычисление признаков задаётся единственным центральным программным модулем. При этом программная система в целом может трактоваться как экспериментальная установка, обеспечивающая исследование различных наборов признаков.

Разработка адекватных систем признаков при комплексном решении проблемы распознавания с целью создания эффективных программных систем является одной из важнейших задач, что прежде всего связано с необходимостью получения описания объектов вне зависимости от изменения условий съёмки. При этом в качестве основных факторов, влияющих на изображение, обычно рассматривают поворот и изменение масштаба. Известно, что неинвариантность методов представления и обработки изображения [3, 32, 80] относительно изменения ориентации и масштаба либо ограничивают область применения, либо требуют специальных усилий для их преодоления.

Что касается учёта ориентации объектов, то применяемые методы разделяются на эвристические [11] и аналитические [17, 77, 81, 82]. С принципиальной точки зрения вторые представляются более перспективными, так как позволяют учесть возможность непрерывного изменения ориентации объектов и избежать перебора вариантов. К современным аналитическим методам эффективного анализа различным образом ориентированных объектов следует также отнести методы представления объектов на основе признаков, построенных из соображений инвариантности относительно поворота [16, 76, 85, 89].

По сравнению с учётом ориентации описание объекта вне зависимости от его удалённости при съёмке является более сложной задачей. Реальное увеличение масштаба изображения выража-

ется двумя эффектами—изменением расстояния между точками объекта и заполнением освободившегося пространства дополнительным количеством точек. Обычная нормировка значений признаков компенсирует влияние первого эффекта. Корректный учёт изменения числа точек представляет самостоятельный интерес.

В практике обработки изображений изменение числа точек, представляющих объекты, оказывается связанным не только с изменением масштаба, но и с рядом других закономерных и случайных причин. При этом определённую роль может играть эффект дискретизации, освещение сцены, геометрия наблюдения, искажения при вводе/выводе, а также нехарактерные особенности самого наблюдаемого объекта, что может сводить на нет эффект применения аналитических результатов. Поэтому, в контексте разработки общей модели распознавания двумерных изображений наиболее практичными являются признаки объектов, интегральные по площади [85], которые в отличие от остальных наименее чувствительны к непредусмотренным искажениям. При определённой аккуратности аналитических выкладок и достаточной точности физической интерпретации в числе интегральных признаков могут рассматриваться ориентация и линейные размеры произвольной фигуры.

Учёт ориентации и масштаба имеет немаловажное значение при использовании для распознавания изображений метода [29–30] распознавания образов по частичному прецеденту. Если составление частичного описания трактуется в буквальном смысле дробления исходного объекта, то выбор конкретного способа разделения целого на части представляется проблематичным. Проблема анализа объекта как набора составных частей состоит в том, чтобы способ разделения объекта на составные элементы

определялся самим объектом и был устойчив к варьированию условий съёмки, в частности – не зависел от ориентации изображения. Вариантом решения является описание каждого объекта в некоторой собственной системе координат.

В процессе выделения и структурного анализа формируется некоторое признаковое представление объектов, которое при классификации преобразуется в идентификаторы. На практике результаты автоматизированной идентификации, как правило, не всегда отвечают наблюдаемым объектам. Задача правильной идентификации или распознавания (в узком смысле) заключается в обеспечении истинного соответствия.

Для современного подхода к методам устранения противоречия между имеющимися и желаемыми результатами распознавания характерно пассивное отношение к требуемому результату, который считается достижимым независимо от степени соответствия входных и установленных выходных данных. В общеупотребительном подходе [26–28] к распознаванию образов для согласования признакового описания объектов с их именами используются два механизма адаптации алгоритмов идентификации объектов. Первый сводится к изменению отображения наборов значений признаков в имена объектов посредством варьирования параметров решающих алгоритмов. Второй механизм обеспечивается за счёт расширения множества решающих алгоритмов на основе формального определения алгебраических операций.

Однако, возможности идентификации определяются не только выбором алгоритмов нахождения имён объектов, но также и степенью непротиворечивости признаковых описаний. Внутренняя противоречивость описаний объектов обусловлена повторяемостью как отдельных значений признаков [39], так и повторяе-

мостью рассматриваемых совокупностей значений, что характерно для автоматической подготовки данных, особенно при использовании числовых признаков и применении частичного анализа объектов. В силу присущих неполным описаниям объектов совпадений значений признаков приёмы их классификации могут оказаться искусственными и недостаточно эффективными. В обсуждаемой ситуации оказывается полезным видоизменение постановки задачи идентификации, при которой результат представляется в виде набора возможных имён объекта, а однозначность достигается только при условии непротиворечивости признакового описания и удачном выборе решающих правил. Указанная постановка задачи упрощает разработку и практическую проверку решений и не требует применения аппарата, выходящего за рамки классической математики.

Нельзя не отметить глубокое значение идеи алгебраического замыкания [26–28], которая обеспечивает возможность развития методов идентификации не только на основе расширения множества решающих правил, но также посредством формирования алгебраически упорядоченной системы накапливаемых в виде обучающей информации сведений об объектах. В сочетании с частичным анализом [29, 30] эта идея представляет самостоятельный практический интерес, так как обеспечивает эффективные расчёты на основе использования алгебраических комбинаций результатов признакового описания изображения. При этом для конкретизации подхода применительно к распознаванию изображений представляется необходимым развитие физической интерпретации формальных алгебраических объектов и операций.

Обсуждаемый подход к распознаванию образов обеспечил развитие многих статистических методов обработки. Вероятност-

ные методы оценивания широко используются как в задачах предварительной обработки изображений, так и в задачах идентификации объектов. Тем не менее, на современном этапе в большинстве приложений автоматизированного распознавания рассматриваются видеообъекты, идентификация которых сама по себе не вызывает сомнений. При этом вероятность правильного распознавания является скорее внутренним параметром системы, чем атрибутом объекта, а теоретические оценки ошибочного распознавания не заменяют полученных на опыте и по сравнению с результатами эксперимента имеют второстепенное значение. Поэтому продолжает оставаться актуальным [68] детерминированный подход к разработке самоорганизующихся систем автоматического распознавания. Последний усиливается благодаря более общей постановке проблемы распознавания с обучением, при которой априорной оценки вероятности правильного распознавания не производится, но предусматривается стратегия решения, позволяющая добиться необходимой надёжности.

Если концепция обучения (на примерах) в теории распознавания образов является хорошо изученной [68, 88], то её конкретная интерпретация в случае распознавания изображений нуждается в уточнении, отражающем особенности данной области. В контексте разработки самоорганизующихся систем распознавания видеоинформации задача обучения не сводится к простому частному случаю классической теории распознавания образов, так как помимо поиска адекватного прогнозирования идентифицируемых объектов включает задачу управления выделением объектов. Описание принципа управления выделением объектов представляется необходимым при конструктивном подходе к комплексному решению проблемы автоматического распознавания изображений.

Атрибутами поступающих по мере обучения в ЭВМ сведений об объектах являются значения признаков и идентификаторы. Посредством интерполяции множеств значений признаков известных объектов определяются соответствующие области допустимых значений (ОДЗ), характеризующиеся диапазонами изменений рассматриваемых величин. Численные характеристики ОДЗ в виде диапазонов рассматриваются в системе автоматического распознавания изображений как внутренние управляющие параметры, определяющие адаптацию системы к данной предметной области. Посредством учёта ОДЗ осуществляется выделение объектов.

В режиме автоматического распознавания помимо знакомых объектов в качестве новых объектов данной предметной области рассматриваются объекты, параметры которых лежат в ОДЗ. При этом новые объекты, параметры которых удовлетворяют ОДЗ, но не удовлетворяют множеству значений признаков идентифицируются как незнакомые объекты заданной предметной области и распознаются в процессе автоматического обучения.

В режиме интерактивного обучения рассматриваются все без исключения обнаруженные объекты.

Таким образом, распознавание объектов некоторой предметной области иллюстрируется схемой рис.2.

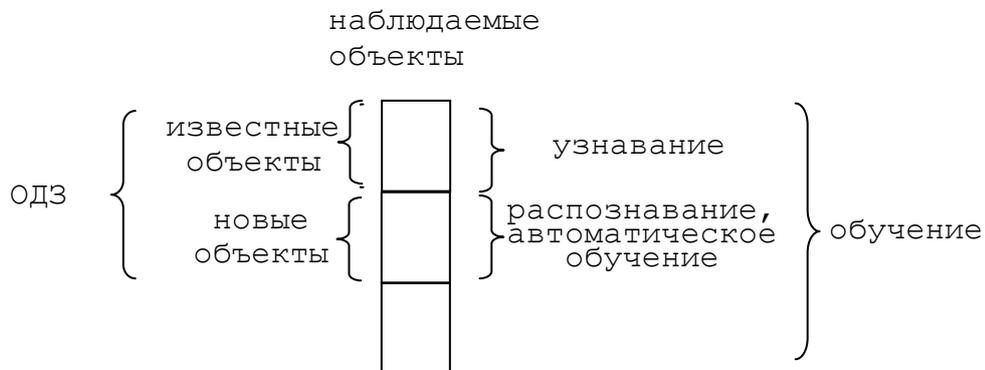


Рис. 2. Задача распознавания объектов

Помимо управления процессом выделения объектов ОДЗ учитывается при фильтрации информации, которая сводится к исключению из текущего рассмотрения объектов (элементов объектов), учёт которых не влияет на результат. В зависимости от контекста задачи из расчёта исключаются объекты, заведомо не являющиеся объектами данной предметной области или при идентификации сужается круг возможных имён объектов.

Таким образом, фильтрация информации при разработке обучаемых систем распознавания трактуется как способ оптимизации вычислений. Применение фильтрации ограничивается также и тем, что в процессе обучения она не используется при выделении объектов, что обеспечивает возможность распознавания любого из доступных в рамках принятой структуры данных.

Следует подчеркнуть, что значение фильтрации информации ограничивается только в контексте научного исследования принципов создания обучаемых систем общего назначения. При этом возможности использования фильтрации несколько ограничиваются. Так неполный учёт признаков в методе голосования обеспечивает эффективную для подавления помех фильтрацию информации, но затрудняет исследование самих систем признаков при использовании модели обучаемой системы в качестве экспериментальной установки. Безусловное рассмотрение признаков наиболее практично в качестве методологии научного исследования, но может требовать модификации при обеспечении внедрения.

Классическая концепция обучения ориентирована на некоторую бесконечную последовательность примеров, что предполагает использование неограниченной памяти ЭВМ. Простейшей реализацией обучения на примерах при обработке изображений является метод использования эталонов, который до сих пор яв-

ляется одним из самых распространённых при решении некоторых прикладных задач [11, 93]. Большой объём оперативной памяти, необходимой для накопления эталонов, влечёт снижение быстродействия в силу резкого возрастания количества данных, анализируемых методом перебора.

Если, в отличие от классического подхода, исходить из заведомо ограниченного объёма используемой оперативной памяти, то вместо эталонов следует применять более общий признаковый анализ. Представление объекта в виде набора числовых признаков и векторов признаков, предполагает сжатие информации с целью экономичного хранения и обработки и обеспечивает упрощение доступа к накопленным данным за счёт исключения перебора посредством упорядочивания. При этом принципиальную возможность обучения обеспечивает неограниченное число признаков, а его практическую достижимость – алгебраическое расширение признаковой системы.

В предположении об ограниченном объёме используемой памяти процесс обучения с накоплением некоторого количества информации об объектах завершается и память переходит в состояние насыщения. При этом необходимый для распознавания объём информации зависит от объекта и насыщение памяти по мере обучения может достигаться для разных объектов в разное время. По достижению насыщения памяти система при предъявлении новых объектов обеспечивает прогноз в виде списка возможных имён. Ошибочная идентификация исправляется в процессе обучения посредством дополнения списка имён. Если прогноз оказывается однозначным, то используемая система признаков считается достаточной. В противном случае она расширяется за счёт включения дополнительных признаков.

Практический смысл концепции ограниченной памяти состоит в том, что её программная реализация при условии симметричного рассмотрения признаков и минимальном использовании управляющих параметров является удобным инструментом прикладного исследования признаковых систем. При этом суждение об адекватности исследуемой системы признаков и заданной предметной области обычно достигается до завершения полного процесса обучения по прогнозу отдельных объектов.

Концепция ограниченной памяти позволяет по-новому представить проблему автоматического обучения, постановка которой до сих пор является предметом обсуждения [15,16,67]. В [67] обучение трактуется как автоматическое расширение накопленной обучающей информации. В [15,16] формулируется задача построения признакового пространства, адекватного объектам.

Как вариант развития [15,16,67] представляется целесообразным развитие идеи формирования признакового описания без непосредственного использования понятия признакового пространства в общепринятом смысле, что достигается отождествлением "близости" объектов по данному признаку со случаем совпадения. При этом распознавание организуется независимо от порядка значений признака, оптимальный выбор которого может оказаться проблематичным [9,61]. Самообучение (автоматическое обучение без использования изображений) трактуется как оптимизация процесса обучения посредством моделирования насыщения обучающей информации. При этом складывающиеся по мере накопления информации закономерности в зависимости от конкретных предметных областей могут влиять на применяемые алгоритмы моделирования. В качестве общего механизма реализации самообучения может использоваться интерполяция, кото-

рую естественно назвать интерполяцией знаний. Благодаря возможности экспериментальной проверки правильности интерполяции самообучение определяется хотя и не вполне строго, но конструктивно, что даёт методологическую основу для алгоритмической организации вычислений.

Эффективное использование аппаратного обеспечения на завершающей стадии распознавания достигается в виде разумного распределения оперативной памяти, занимаемой блоком признакового описания объектов и обучающей информацией (библиотекой знаний). Оба блока информации при распознавании (идентификации) объектов используются одновременно и поэтому в оптимальном случае по суммарному объёму не должны превосходить ресурсов ЭВМ. При ограниченных ресурсах оперативной памяти ЭВМ (типа IBM PC) имеет смысл предположить, что по сравнению с блоком знаний блок признакового описания требует большего оперативного пространства, так как содержит многокомпонентное представление всех имеющихся на изображении объектов в виде некоторых наборов составных частей. Положение об относительно небольшом объёме отводимой под знания оперативной памяти согласуется также и с тем, что в принципе инициализация работы обучаемой программной системы производится при нулевом объёме знаний, на которые по идее должна расходоваться пренебрежимо малая память. Средства поддержки стандартных баз знаний современных экспертных систем [4, 7] сами по себе требуют существенных затрат оперативной памяти, что ограничивает возможности их использования в обучаемых системах автоматического распознавания. Таким образом, при ограниченных ресурсах оперативной памяти организация библиотеки знаний обучаемой системы представляет самостоятельный интерес.

На первоначальном этапе выделения объектов большой объём оперативной памяти занимает многозначное внутреннее представление изображения. При этом однозначное выделение объектов реализуется в соответствии с содержанием блока накопленных знаний. Признакомое описание объектов формируется на промежуточном этапе обработки в процессе структурного анализа, что требует оперативного использования данных о принадлежности исходных точек изображения выделенным объектам. При этом пронумерованные объекты задаются посредством разметки №(.) точек исходного изображения. Таким образом, использование оперативной памяти ЭВМ на трёх последовательных стадиях обработки характеризуется схемой рис.3.

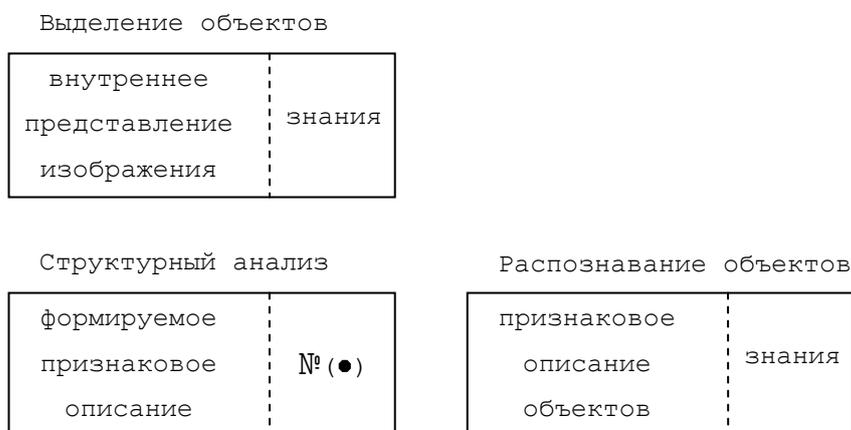


Рис. 3. Схема использования оперативной памяти ЭВМ

Следует обратить внимание, что в схеме рис.3 предусматривается использование оперативного пространства только для определённым образом переработанной, систематизированной информации. При этом для исходного изображения существенного объёма оперативной памяти не отводится. Предполагается, что изображение при генерации многозначного внутреннего представления и при структурном анализе объектов считывается в опе-

ративную память и обрабатывается последовательными порциями ограниченной величины. При этом исходная матрица рассматривается локально, разметка задаётся в компрессированном виде, а извлекаемые сведения об объектах накапливаются в свободном оперативном пространстве.

При непосредственном преобразовании исходных данных в объектно-ориентированные достигается оптимизация расчётов, но одновременно утрачивается наглядность матричного представления, что затрудняет применение эвристических алгоритмов и способствует развитию теоретического обобщения и физической интерпретации. Следует заметить, что принципиальная необходимость формирования объектно-ориентированного описания изображения в процессе локальной (построчной) обработки исходной матрицы возникает не только при разработке программных систем, но также и в случае аппаратных реализаций [36]. Поэтому условие формирования компактного целостного образа в процессе последовательного локального анализа исходного изображения представляется концептуальным.

Применяемое в диссертации конкретное объектно-ориентированное описание полутонового изображения построено в результате обобщения опыта решения практических задач, связанных с адаптивной многоуровневой аппроксимацией.

1.2. Задача многоуровневой адаптивной аппроксимации полутонового изображения.

Аппроксимация полутонового изображения, при котором оно замещается некоторым близким аналогом, используется для оптимизации хранения и передачи видеoinформации и хорошо изучено

[48] в применении к системам, рассчитанных на участие человека. В качестве меры близости обычно принимается среднеквадратичное отклонение по яркости. Наиболее употребительными методами аппроксимации являются "кодирование с преобразованием" и блочное кодирование. Первое [31,83] выполняется на основе подавления спектральных составляющих, с визуальной точки зрения не влияющих на качество изображения, второе [3,71] производится на основе использования набора стандартных блоков различных размеров. Применение как первого, так и второго метода влечёт искажение формы изображённых объектов и с точки зрения возможности автоматизированного распознавания может ухудшать качество изображений, что вызывает необходимость использования адаптивных методов аппроксимации [19,22,33,84,91].

В [33] адаптивная аппроксимация объектов без искажения геометрической формы обеспечивается совокупностью следующих трёх особенностей обработки изображения.

Во первых, описание геометрии объектов не сводится к заранее предусмотренным стандартным аналитическим зависимостям. Вероятностные или иные закономерности распределения яркости не предполагаются известными. Во вторых, все рассматриваемые на изображении области считаются равноправными и в третьих, обработка изображения производится итеративно. При этом обеспечивается также инвариантность аппроксимации объектов относительно способа сканирования и условий съёмки (сдвига, поворота, изменения масштаба и др.).

Следует отметить, что при условии неискажённой аппроксимации объектов оценка качества аппроксимации по среднеквадратичному отклонению теряет самостоятельное значение. Минимум

среднеквадратичного отклонения достигается при простом усреднении яркости в пределах объектов, однако, само по себе объектно-ориентированное представление изображения нуждается в уточнении и связано с проблемой сжатия информации.

В современном контексте создания автоматизированных систем обработки сжатие видеoinформации трактуется [35] как приведение данных к виду, удобному для последующего распознавания, что в конкретной интерпретации сводится к упрощению изображения. Упрощение изображения можно трактовать как укрупнение частей объектов за счёт их взаимного слияния и удаление объектов, в контексте данной предметной области считающихся заведомо ложными. Таким образом, сжатие видеoinформации обеспечивается на предварительной стадии распознавания изображения – при перечислении объектов. При этом специфика состоит в том, что перечисление объектов осуществляется независимо от их идентификации, которая в классической схеме распознавания изображений [77] выполняется на одном из последующих этапов обработки.

Неоднозначная интерпретация изображения, необходимая для обеспечения возможности выделения объектов различных морфологических уровней, достигается применением многоуровневого представления видеoinформации. При этом видеоданные эквивалентны набору матричных представлений исходного изображения с различной степенью детализации. Предполагается, что уровень с максимальной детализацией обеспечивает аппроксимацию объектов с необходимой точностью.

Следует указать, что по сравнению с традиционным использованием [19, 22, 84, 91] адаптивных методов аппроксимации результаты итеративного преобразования изображения [33]

трактуются в более общем смысле. Итеративная аппроксимация объектов областями различной формы и размеров, достигаемая в алгоритме [33], порождает объектно-ориентированное представление, в котором все несовпадающие области рассматриваются равноправно. Считается, что каждая из полученных областей в контексте той или иной задачи может рассматриваться как объект или по крайней мере как часть объекта. При этом вопрос об оптимальной аппроксимации изображения на некоторой итерации снимается, уступая место проблеме эффективного описания объектов посредством элементов, вычисленных в базовом алгоритме выделения объектов [33] на различных итерациях.

1.3. Базовый алгоритм выделения объектов.

Число исходных точек реального полутонного изображения существенно превышает количество объектов. Поэтому, на стадии выделения объектов имеет смысл применять простейшие из эффективных алгоритмов группировки точек, которые при относительно небольшой сложности вычислений обеспечивают аппроксимацию объектов без искажения геометрической формы.

Принятый в работе базовый алгоритм сегментации полутонных изображений [33, 34] обеспечивает выделение объектов, устойчивое к изменению условий съёмки. Он относится к алгоритмам сегментации [18, 23, 45] посредством наращивания областей и сводится к поэтапному слиянию областей, наиболее близких по средней яркости.

Наглядно базовый алгоритм описывается следующим.

Выберем на изображении любую точку и представим, что, начиная с неё, расплывается "чернильное" пятно, последова-

тельно захватывая соседние точки, если тому не препятствует перепад яркости, превышающий некоторый заданный порог ε . В результате формируется определённая область однородной яркости. Рассматривая различные начальные точки, всё изображение можно разбить на области однородной яркости. Разбиение не зависит от выбора начальных точек, а определяется установленной величиной ε .

Сначала изображение разбивается на области при минимальном ε . В пределах каждой области проводится осреднение по яркости и цикл обработки повторяется до тех пор, пока на изображении все локальные яркостные перепады не превысят ε . Затем порог ε наращивается и процедура воспроизводится при новом пороге.

Если под свободными точками понимать точки, не принадлежащие какой-либо области, а под близкими яркостями иметь в виду яркости, модуль разности которых не превышает порог ε , то цикл сегментации изображения при заданном пороге перепада яркости состоит в следующем.

Все точки считаются свободными.

1. Если свободные точки исчерпаны, то обработка заканчивается. В противном случае находим свободную точку.

2. Выбранная точка порождает связную область, которая формируется рекурсивно} при включении точки в данную область к этой же области относим и все свободные соседние точки близкой яркости.

3. Заменяя яркостные значения точек сформированной области их арифметическим средним, возвращаемся к п.1.

Для получения стабильного результата преобразования, имеет смысл при фиксированном пороге повторять обработку

согласно пп.1-3 до получения устойчивой картины.

Преобразуя исходное изображение с порогом 1, затем с порогом 2 и т.д., получаем определенную цепочку в различной степени детализированных представлений исходного изображения. При этом параметр ε определяет яркостное разделение областей между собой.

На практике оказывается допустимым наращивать порог непосредственно с каждой итерацией, определяемой пп.1-3.

Рис. 4 иллюстрирует обработку реального изображения.

Результат обработки представляется в виде последовательности изображений (уровней) с убывающим количеством областей различной формы и размеров. Начальный уровень совпадает с исходной картиной, конечный - представляется однотонной окраской поля изображения. Число уровней заранее не нормируется. Характерно, что области на том или ином уровне геометрически совпадают с объектами (частями объектов) исходного изображения, причём форма объектов не искажается.

Практическое применение базовый алгоритм обработки изображений [33] нашёл при анализе глубоководных снимков (рис.5).

В практике морских геологических работ возникает проблема поиска залежей железомарганцевых конкреций по результатам наблюдения участков морского дна. Районы залежей конкреций прогнозируются на основе анализа большого количества фотоснимков, что требует автоматизации обработки. Конкреции на фотоснимках, получаемых с помощью глубоководных телекамер, хорошо просматриваются визуально, однако, на результаты обработки влияет неравномерность освещения по полю кадра.

Типичные приемы фотоснимков указанного типа используются в системе KMTV (Венгрия). Характерная неравномерность ос-

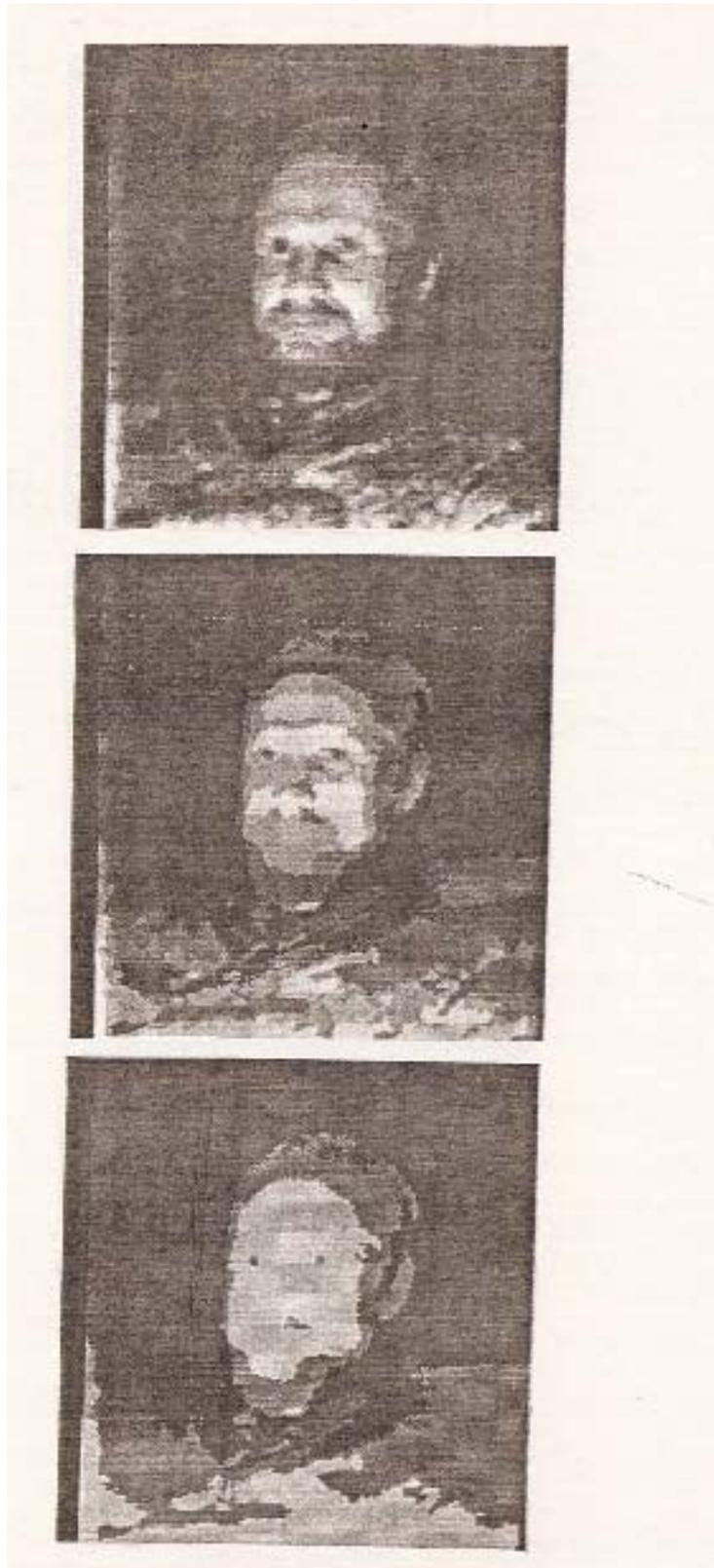


Рис. 4 Многоуровневая сегментация полутонного изображения

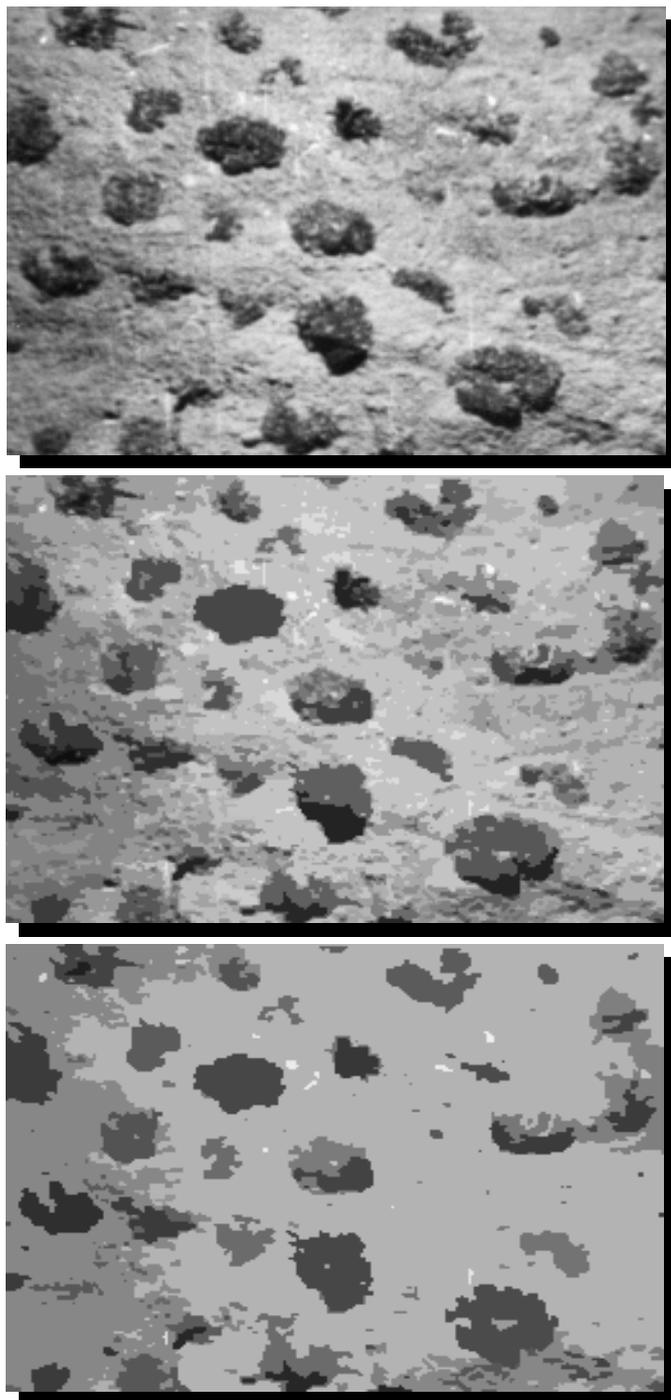


Рис.5. Выделение конкреций на снимке морского дна.

вещения учитывается по нескольким рабочим снимкам. При этом выделение конкреций сводится к пороговому преобразованию [13] с учётом неравномерности освещения. В системе KMTV для поиска конкреций применяется также метод выделения границ по перепаду яркости, что недостаточно эффективно и усложняет расчёты в силу необходимости устранения разрывов при выявлении замкнутых областей ([14]). С целью анализа фотоснимков разрабатываются также и отечественные системы [12].

Опыт использования базового алгоритма обработки изображений для выделения конкреций показал его перспективность. Оказалось, что выделение конкреций достигается на одном из уровней, то есть в наиболее простой стандартной интерпретации результатов итеративной сегментации. Применение базового алгоритма упростило обработку и позволило обойтись анализом отдельных изображений, не проводя сравнения снимков друг с другом. По сравнению с результатами обработки, полученными на основе использования известных методов, достигалось лучшее качество выделения объектов без искажения геометрической формы.

Рис.6 и рис.7 характеризуют выделение конкреций. На рис.6 приведена гистограмма исходного снимка, на рис.7 – гистограмма изображения, полученного при последовательном наращивании порога от 1 до 40. Максимальный пик фона (в центре) четко отделяет группу яркостей, отвечающих конкрециям, от группы яркостей более светлых пятен (бликов). При этом бинаризация результирующего изображения рис.1 сложности не представляет.

Следует подчеркнуть, что итеративное наращивание порога яркостного перепада имеет принципиальное значение. Так, если попытаться обработать исходное изображение конкреций с порогом 20, то конкреции сольются с фоном.

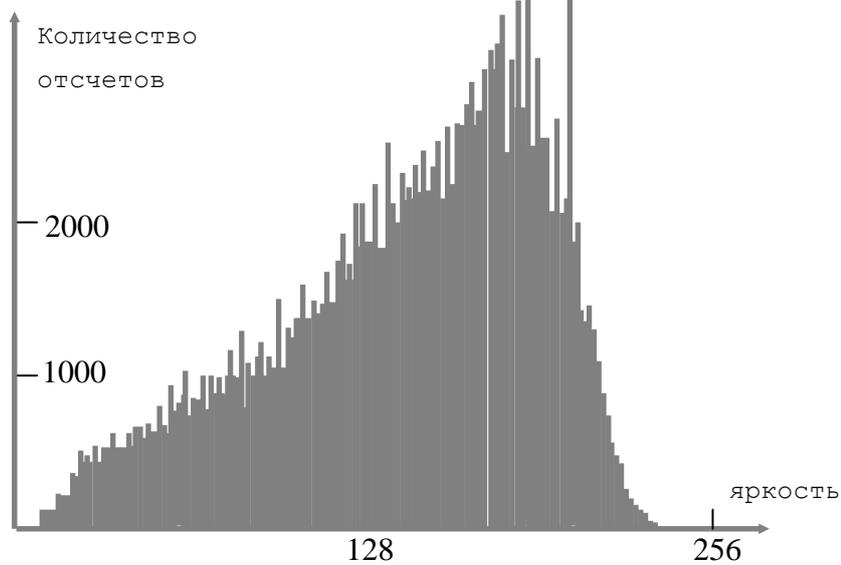


Рис. 6. Гистограмма яркостей исходного изображения.

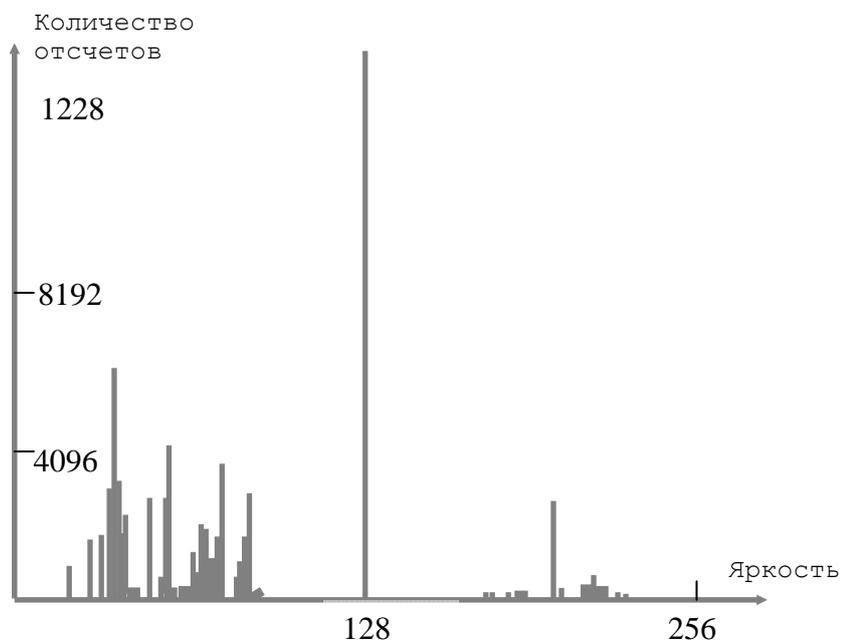


Рис. 7. Гистограмма яркостей сегментированного изображения

Следует отметить, что для обсуждаемого алгоритма характерно симметричное рассмотрение областей, при котором считается, что все области в равной степени представляют объекты (части объектов). Поэтому он применим для выделения на полутоновых изображениях объектов различных типов, что помимо практики обработки глубоководных снимков подтверждено

результатами многочисленных экспериментов с аэрокосмическими изображениями и изображениями, полученными с помощью телекамеры (рис.4). Эксперименты с телевизионными изображениями проводились на примере снимков человеческих лиц. Аэрокосмические изображения изучались в виде снимков земной поверхности и искусственных космических объектов. Обработка изображений лиц показала, что разумная степень сегментации исходных изображений практически не влияет на распознавание, так как глазом различие воспринимается слабо. По сравнению с изображениями лиц снимки земной поверхности геометрически более просты и на них легче рассмотреть результаты укрупнения областей. Как показал опыт, после 10-15 итераций обработки снимков земной поверхности наблюдаемые объекты результирующего изображения представляются равнояркостными областями с резкими границами, что удобно для автоматизации распознавания.

В целом анализ результатов экспериментов с совокупностью изображений показал, что в отличие от задачи выделения конкреций, оптимальное приближение объектов одного вида, а также различных частей данного объекта посредством равнояркостных областей не всегда достигается на одном и том же уровне. Следовательно, стандартный подход к адаптивной сегментации, при котором решается задача поиска оптимального уровня, требует обобщения, допускающего аппроксимацию объектов посредством набора составных частей, вычисленных на разных уровнях сегментации, что требует, в свою очередь, развития адекватного алгоритмического аппарата.

Укрупнение равнояркостных областей в процессе циклического применения базового алгоритма сегментации до

определённого предела можно рассматривать как удаление информационной избыточности исходного полутонового изображения, что обнаруживается явно при применении к сегментированному изображению стандартных программ факсимильного сжатия (pkzip, ise и др.). Как показал эксперимент со снимками искусственных космических объектов, для сегментированных полутоновых изображений с укрупнёнными равнояркостными областями коэффициент сжатия возрастает от 1-1.5 до 10-30. Таким образом, замещение исходных снимков изображениями с укрупнёнными областями решает задачу хранения изображений при ресурсах ПЭВМ.

При интерактивном контроле в ходе итеративной сегментации выбор уровня аппроксимации исходного изображения производится из условия сохранения на сегментированном изображении всей информации, полезной для данной предметной области. Однако, для автоматизированного снижения информационной избыточности необходим численный критерий, который может быть получен [61] на основе анализа площадей областей.

Максимально возможное количество областей различной площади линейно зависит от размеров и для квадратного изображения $N \times N$ оценивается числом $r^2 \cdot N$. Для реального полутонового изображения разброс встречающихся площадей обычно значительно меньше. В процессе итеративной сегментации число областей различной площади растёт, достигая при определённом ϵ_0 максимального значения (рис.8).

Как показали эксперименты, несмотря на существенное (в 7-10 раз) уменьшение общего количества областей, аппроксимация изображения, определяемая ϵ_0 , сохраняет семантическое содержание (объекты не сливаются с фоном и между собой).

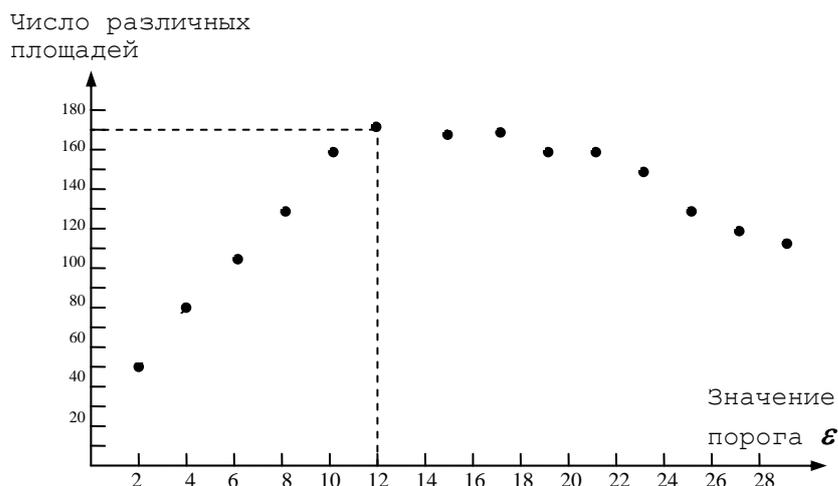


Рис. 8. Изменение числа встречающихся областей различной площади в процессе итеративной сегментации полутонового изображения размером 256×256

Многokратная обработка изображения в базовом алгоритме сегментации позволяет улучшить качество аппроксимации объектов, однако, ведёт к большим временным затратам. Снижение времени достигается без применения спецпроцессоров за счёт оптимизации операций со связными областями произвольной формы и размеров, что требует разработки структуры данных, отличной от матричной. При этом базовый алгоритм сегментации сохраняет практическое значение для отладки программ и в условиях ограниченных ресурсов ПЭВМ может использоваться для редукции изображений по количеству областей с целью уменьшения необходимой оперативной памяти. Необходимо отметить, что редукция изображения по количеству областей согласно базовому алгоритму допускает оптимизацию вычислений с комбинированной структурой данных [60, 61], но выигрыш по времени достига-

ется за счёт неадекватного возрастания сложности алгоритмов. Ресурсов памяти ПЭВМ достаточно для самостоятельной разработки отличного от матричного способа представления изображения.

1.4. Цель и задачи работы.

Совокупность различных сегментов, получаемых на всех уровнях аппроксимации изображения в базовом алгоритме выделения объектов, является вариантом экспериментально обоснованного объектно-ориентированного представления и служит реальной основой комплексного решения проблемы распознавания, которое сводится к выделению, структурному анализу и идентификации (классификации) объектов.

Практической реализацией комплексного решения является создание прототипа обучаемой (самообучаемой) системы автоматического распознавания изображений, настраиваемой на обнаружение и идентификацию объектов желаемого типа при условии, что используемая априорная информация ограничена названиями объектов, вводимыми в ЭВМ в процессе обучения. В условиях исследования реализация сводится к созданию программного ядра, адаптируемого к решению широкого круга прикладных задач. С целью автоматизации исследования прототип системы распознавания рассматривается в качестве экспериментальной установки, обеспечивающей оптимизацию системы признаков объектов.

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 полутонных и двухградационных изображений.

Указанная цель достигается благодаря решению следующих основных задач}

1) Разработка методов многоуровневой сегментации полутонного 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.

Аппроксимация объектов посредством вычисляемых в базовом алгоритме связных областей развивается в многоуровневой модели выделения объектов по яркости методами слияния и дробления сегментов изображения. Оптимизация расчётов достигается благодаря представлению связных областей в виде деревьев и табличному представлению соседства областей.

Возможности многоуровневой модели расширяются благодаря развитию геометрического анализа, необходимого также для реализации идентификации объектов по наборам признаков.

ГЛАВА 2. МНОГОУРОВНЕВАЯ МОДЕЛЬ ИЗОБРАЖЕНИЯ

В главе предлагается компьютерная модель выделения объектов на полутоновом изображении, строится адекватная динамическая структура данных, описываются основные алгоритмы, приводятся результаты применения. Предусматривается обобщение модели на случай двухградационных изображений.

Изображение аппроксимируется последовательностью картин (уровней) с различным количеством яркостных областей. Совокупность уровней задаётся в виде нерегулярного дерева со специальной индексацией узлов. При компрессированном хранении обеспечивается быстрое восстановление в матричной форме. Дётся алгебраическое определение видеообъекта. Без преобразования к матричному виду организуется перечисление объектов с оценкой числа составных частей. Посредством компактной структуры данных обеспечивается также разбиение изображения на зоны трёх типов (тёмные, светлые и переходной яркости), которое производится без использования управляющих параметров. Преобразование изображения в картину с тремя яркостными градациями обобщается в методе получения многоуровневого представления исходного изображения путём последовательной детализации.

Развитие применения модели заключается в использовании наряду с яркостными также и геометрических характеристик.

2.1. Понятие многоуровневой модели.

В простейшем случае многоуровневая модель визуальной информации строится на множестве приближений полутонного изображения посредством областей однородной яркости (рис.4).

Многоуровневая модель – это своего рода алгебраический "конструктор", обеспечивающий аппроксимацию объектов сцены "детальками" различной формы и размеров. Детали способны сливаться друг с другом, что позволяет представить объекты минимальным количеством частей.

Детали конструктора в виде связанных областей задаются как элементы последовательности разбиений изображения (уровней), порождаемой базовым алгоритмом итеративной сегментации. В качестве начального уровня берётся исходное изображение, заданное посредством стандартного разбиения на клеточные поля в виде квадратной решётки, либо используется любое другое разбиение изображения на связанные области. Ограничений на форму и размеры рассматриваемых областей не накладывается. Каждая область характеризуется средней яркостью, которая выражается через характеристики, аддитивные относительно слияния областей – интеграл яркости по площади и площадь. Пары областей характеризуются отношением смежности.

Уровни строятся последовательно. При фиксированном текущем разбиении находятся все смежные области близкой яркости, затем путём слияния найденных областей из рассматриваемого разбиения получают следующее с меньшим количеством элементов. Каждое новое разбиение изображения от

порядка слияния пар смежных областей не зависит. Уровни представляются наглядно, посредством усреднения яркостных отсчётов в пределах укрупнённых областей (рис.4).

В порядке построения уровни индексируются неубывающими значениями параметра ε . В [10,34,62] значение индекса ε отождествляется с верхней границей минимального перепада средней яркости, встречающегося между соседними областями, и приписывается результату нескольких последовательных итераций переходного процесса достижения верхней границы при слиянии областей. Для простоты индекс ε допустимо отождествлять с порядковым номером итерации. Тогда минимальный перепад яркости с ростом ε , вообще говоря, изменяется немонотонно.

2.2. Алгебраическое определение объекта.

В характерных случаях области на том или ином уровне представления геометрически совпадают с объектами. В более общем случае они совпадают лишь с визуально наблюдаемыми частями объектов, как правило, на разных уровнях. Возникает задача моделирования объекта, состоящего из нескольких частей, путём совмещения областей различных уровней. Имеется простое решение [64].

Суть используемого представления об объекте произвольного полутонового изображения поясняется следующим.

Представим вначале, что нам известны некие "кирпичики", из которых состоят искомые объекты. Тогда, выявляя их,

найдём и сами объекты, которые можно разделить между собой по признаку связности, если каждый объект занимает отдельное место. Для того, чтобы составить некоторое макроописание объектов, следует предположить, что "кирпичики" объединяются между собой в укрупнённые «блоки».

Если "кирпичики" заранее неизвестны, то их можно сформировать из некоторых исходных стандартных элементов путём слияния последних по определённому принципу.

Тогда представление об объекте полутонного изображения сводится к следующему:

- из некоторого набора исходных элементов строятся различные "кирпичики";

- из последних отбираются одинаковые по некоторому признаку;

- «блоки» из выбранных "кирпичиков" образуют части объектов сцены;

- объекты разделяются между собой по особым признакам целого.

Следует подчеркнуть, что понятие объекта в многоуровневой модели не связывается с конкретным контекстом изображения, а вводится на основе общих соображений. В точной интерпретации определение объекта формулируется посредством рассмотрения совокупности элементов разбиений различных уровней сегментации изображения. При этом "кирпичики" и "блоки" представляются в виде связных областей, находящихся в отношении "часть"- "целое". Части объекта совпадают с "блоками" максимальной площади и вычисляются посредством объединения пересекающихся связных областей из выбранного набора.

Рассмотрим совокупность областей различной формы и размеров образованную неповторяющимися элементами из всех разбиений многоуровневого представления. Пересечение любых двух областей сводится к тому, что одна из них включает другую.

Из рассматриваемой совокупности областей выделим набор S , признаки которых лежат в некоторой области допустимых значений (ОДЗ), например, имеющих площадь и среднюю яркость в заданном диапазоне. Подмножество областей s набора S , не вложенных в более крупные области из S , рассматривается как множество частей (элементов) различных объектов.

Части s образуют объекты, которые разделяются между собой по признаку связности, исходя из условия, что каждый объект занимает отдельное место. В общем случае разделение объектов может определяться также и другими соображениями.

Сформулированное алгебраическое определение позволяет посредством ЭВМ выявить на произвольном полутоновом изображении объекты заданного типа и определить число составных частей, из которых состоит каждый объект. При этом для выполнения расчёта достаточно указания общих ограничений на параметры областей. Характерно, что в зависимости от установленной ОДЗ любая исходная область может представлять объект (элемент объекта).

Простейшим образом получить структурированные объекты в наглядной форме можно, последовательно изображая области с допустимыми значениями параметров на плоскости. Обработка по уровням производится в порядке возрастания ε . В процессе

формирования частей объектов большие области однородной яркости замещают меньшие, полученные ранее на том же месте.

Как показали вычислительные эксперименты, выявление объектов согласно изложенному принципу с точки зрения человеческого восприятия выглядит естественно. Количество наблюдаемых на изображении составных частей приблизительно совпадает с результатами компьютерных вычислений. Результаты устойчивы к изменению значений входных параметров.

2.3 Структура данных для выделения объектов

2.3.1 Понятие видеоданных.

Матричная форма представления изображения в ЭВМ не позволяет эффективно организовать обработку видеоинформации. Описание в этой форме уровней представления требует слишком много памяти, а анализ областей путём посимвольного сканирования связан с большими затратами времени. Для оптимизации вычислений в рамках многоуровневой модели разработана специальная динамическая структура видеоданных, формализм которой заключается в следующем.

Последовательность разбиений изображения задаётся на начальном отрезке натурального ряда в виде нерегулярной структуры дерева, связи которого индексированы порядковыми номерами уровней, что позволяет хранить и обрабатывать уровни представления в компрессированном виде. При этом обеспечивается быстрое восстановление любого уровня в

матричной форме, необходимое при контроле видеоинформации. Без перехода к матричному представлению путём анализа аддитивных характеристик областей на ЭВМ производится выявление структурированных объектов с подсчётом числа составных частей. Решаются также и другие задачи.

Особенностью аппарата многоуровневой модели составляет описание операций с областями изображения в виде адекватных действий с формальными структурами, представляющими в ЭВМ области и отношения между ними. Основу представлений многоуровневой модели составляет представление связанных областей в виде деревьев.

2.3.2 Представление связанных областей в виде деревьев.

Пусть исходные области полутонового изображения пронумерованы в произвольном порядке целыми числами от 1 до N .

Области любого уровня представления изображения условимся описывать в виде деревьев [10,62], устанавливая связи между обозначениями компонент. Каждой исходной области сопоставляется дерево, состоящее из единственного изолированного узла. Слиянию областей ставится в соответствие слияние деревьев, при котором связываются между собой корневые (минимальные по номеру) узлы (рис. 9). Корневой узел совокупного дерева принимается в качестве обозначения укрупнённой области, а с помощью всего набора узлов её можно получить путём объединения элементов исходного разбиения.

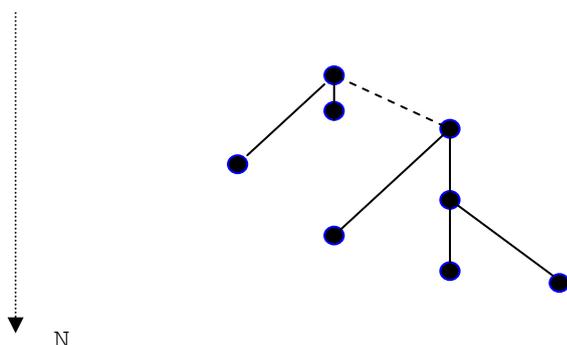


Рис 9. Описание слияния пары областей в форме слияния деревьев. Новая связь обозначена пунктиром.

Таким образом, каждая укрупнённая область представляется в виде дерева исходных. При условии слияния между собой только смежных областей дерево отвечает новой связной области.

В ЭВМ разбиение изображения на области [61] задаётся в виде массива чисел $f[n]$, где n - номера узлов деревьев, отвечающие областям исходного разбиения изображения, значения f обозначают связи, установленные между узлами деревьев.

В исходном состоянии массив $f[n]$ представляет собой ряд натуральных чисел от 1 до N , где N - число элементов начального разбиения изображения на связные области, в котором все значения f совпадают со своими адресами и являются обозначениями элементов начального разбиения.

При слиянии пары областей большее по величине обозначение области замещается меньшим.

Текущее обозначение области совпадает со значением одного из сохранившихся элементов массива и определяется в алгоритме циклической переадресации, который сводится к последовательности переходов от одного элемента к другому.

Значение данного элемента определяет адрес следующего. Переходы производятся до тех пор, пока не найдётся элемент, адрес которого совпадает со значением, записанным по этому адресу, что эквивалентно поиску элемента, сохранившего исходное значение. Значение найденного элемента рассматривается в качестве текущего номера области. Например, область, описываемая массивом 1 1 2 3 4 5 имеет обозначение 1.

В терминах языка СИ алгоритм циклической переадресации выражается оператором:

```
FOR( n=n ; n != f[n] ; n=f[n] ) { } ;,
```

где n - адрес рассматриваемого узла.

Результат n отвечает текущему обозначению, вычисленному на данный момент времени.

Следует отметить, что представление областей в виде деревьев позволяет оптимизировать решение [36] часто возникающей задачи разметки изображения на связные области в условиях построчного считывания.

2.3.3 Табличное описание соседства областей.

Соседство (отношение смежности) областей друг с другом описывается [60,61] в многоуровневой модели бинарной симметричной таблицей связности, в начальном состоянии имеющей размеры $N \times N$ (рис. 10). Подобные таблицы применяются для анализа текстур [95] по локальным характеристикам и в машинной графике [94] для описания трёхмерных объектов. Причём, в [94] рассматриваются операции с таблицами при

описании преобразований конкретных объектов.

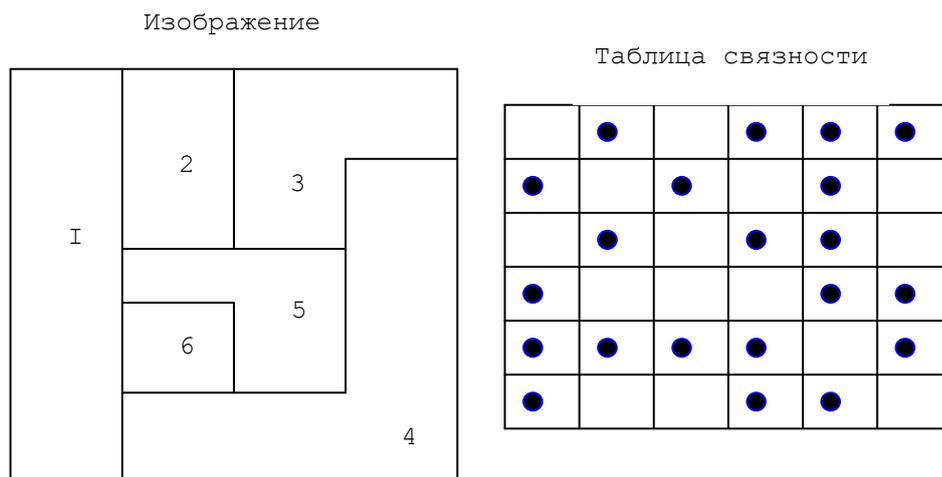
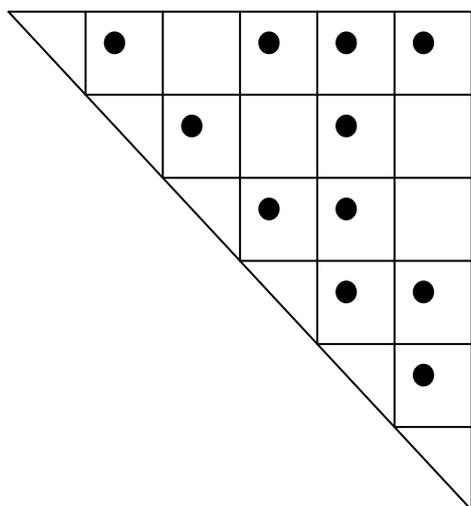


Рис. 10 Описание соседства областей посредством таблицы.

При описании соседства элементов произвольного разбиения областям изображения сопоставляются строки (столбцы) таблицы связности в порядке возрастания номеров. Точка на пересечении i -й строки и j -го столбца вместе с симметричной ей меткой на пересечении j -й строки и i -го столбца свидетельствует, что i -я и j -я области смежные. Благодаря симметрии таблицы связности для описания соседства областей изображения друг с другом достаточно её верхней диагональной части, которая в ЭВМ представляется списками позиций точек в строках (рис. 11). Адрес каждого списка отвечает одной из областей. В списке перечисляются области смежные с данной и превосходящие её по номеру. Порядок перечисления не важен. В случае обычного разбиения исходного изображения с помощью квадратной решётки на $N \times N$ клеточных полей и фиксированной нумерации полей, например, в порядке

телевизионной развёртки, в расчётах может применяться стандартная начальная таблица связности, вид которой известен заранее.

Наглядное представление



Представление в ЭВМ

1: 4 6 5 2

2: 5 3

3: 5 4

4: 5 6

5: 6

6:

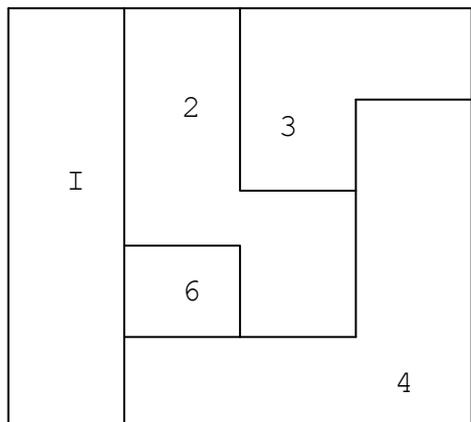
Рис. 11 Представление таблицы связности списками областей.

Слиянию областей изображения отвечает слияние строк таблицы связности и сопутствующее слияние симметричных строк столбцов. Слияние строк (столбцов) сводится к их покомпонентному объединению по закону логического сложения, занесению результата на место строки (столбца) с наименьшим номером и вычёркиванию остальных слагаемых. При этом точки, попадающие на диагональ, стираются.

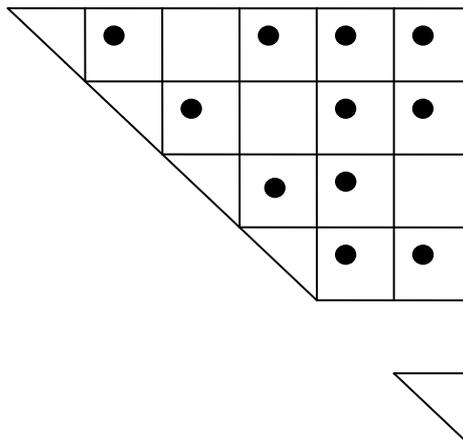
С помощью верхней диагональной части таблицы связности слияние областей описывается почти так же просто (рис.12, ср. с рис.10). Почленное объединение пары строк (столбцов) по закону логического сложения сопровождается вычёркиванием строки (столбца) с большим номером. Точки, попадающие при слиянии столбцов в нижнюю часть таблицы связности путём зеркального отражения относительно диагонали переносятся в верхнюю её часть, а попавшие на диагональ стираются. В ЭВМ

преобразования таблицы рис.12 организуются как операции со списками (рис.11).

Результат слияния сегментов



Слияние строк



Преобразование столбцов

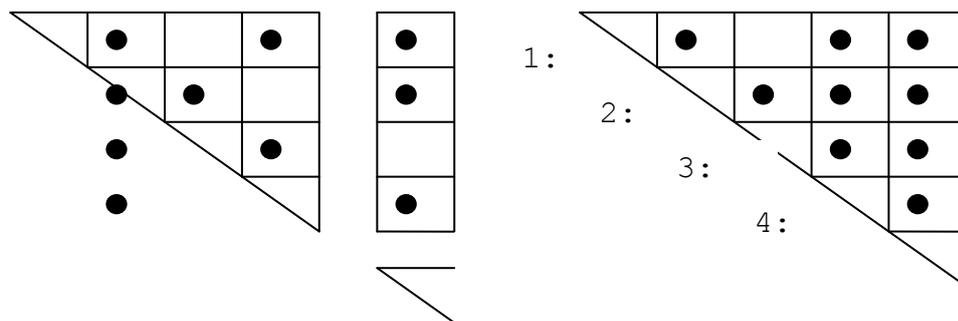


Рис. 12 Преобразование таблицы связности при слиянии областей.

Следует иметь в виду, что таблица связности, даже в усечённом варианте занимает сравнительно большой объём оперативной памяти. При редукции изображения по количеству областей (п.3 первой главы) достигается необходимое сокращение размеров таблицы связности.

Для того, чтобы логическим путём добиться оптимизации

расчётов с таблицей связности по памяти, достаточно предусмотреть перенумерацию областей изображения в порядке возрастания числа соседей. Тогда усечённая таблица связности записывается набором строк минимальной длины. При этом строка, отвечающая фону, оказывается пустой.

2.4. Генерация структуры. Обеспечение прямого доступа.

Преобразование исходного разбиения изображения и последовательная генерация остальных уровней описывается посредством массива f (п.2.1.2), таблицы связности (п.2.1.3) и массивов аддитивных характеристик (площадей и интегральных яркостей).

Получение очередного уровня представления состоит в последовательном установлении связей между корневыми узлами деревьев, отвечающих ближайшим по средней яркости смежным областям. Слияние пары деревьев описывает слияние двух или более областей преобразуемого уровня и производится в том случае, когда узлы, отвечающие смежным областям близкой яркости, в процессе текущего преобразования не оказались узлами одного дерева ранее. По завершению слияния деревьев для укрупнённых областей определяются новые средние яркости.

Очередной цикл обработки при данном значении ε состоит в следующем.

1. Просматриваются имеющиеся строки таблицы связности и находятся все пары близких по яркости смежных областей. При обнаружении каждой такой пары модифицируется массив f :

- для каждой из двух найденных областей в алгоритме циклической переадресации вычисляется текущее обозначение;

- если вычисленные текущие обозначения не совпадают, то значение большего по адресу (и по величине) элемента заменяется значением меньшего. Одновременно суммируются площади и величины интегралов яркости. Результаты суммирования в качестве атрибутов приписываются меньшему элементу массива f .

2. Дважды просматриваются элементы массива f , отвечающие невычеркнутым строкам таблицы связности. По элементам с изменившимися значениями в алгоритме циклической переадресации определяются цепочки объединяемых областей. При одном просмотре производится соответствующее слияние строк таблицы связности, при другом - слияние столбцов (рис.12). В процессе слияния строк для укрупнённых областей вычисляются новые средние яркости.

На выходе каждого цикла обработки в таблице связности остаются строки (столбцы), отвечающие укрупнённым областям. Сопоставляемые этим областям элементы массива f совпадают с исходными. Ассоциированные с сохранившимися элементами массива f площади, яркостные интегралы и средние яркости имеют значения, отвечающие новому разбиению. Алгоритмом циклической переадресации определяется, какие области исходного изображения присоединились к каждой из сохранившихся.

Последовательная во времени генерация новых уровней представления производится до тех пор, пока все деревья не

сольются в одно совокупное дерево, соответствующее области всего изображения. Если для каждой связи i , зафиксировать индекс уровня ε_i , при формировании которого эта связь была установлена, то любой уровень с индексом ε можно восстановить [10, 63]. Достаточно разорвать все связи, для которых

$$\varepsilon_i > \varepsilon.$$

Тогда совокупное дерево распадётся на систему деревьев, задающую искомое разбиение.

Таким образом, многоуровневая сегментация полутонного изображения (рис.4-5) запоминается в виде дерева с индексированными связями рис.13 (ср. с рис.9) и может быть охарактеризована как сегментация с "памятью".

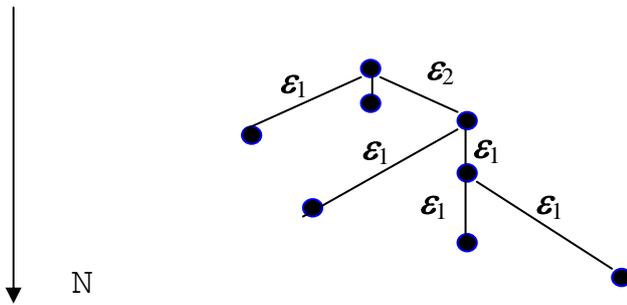


Рис.13. Описание уровней посредством индексации связей

Практическое значение сегментации с "памятью" состоит в том, что она обеспечивает прямой доступ к видеоданным. Под прямым доступом понимается реализация быстрого преобразования видеоданных к матричному виду, что необходимо для интерактивного контроля видеоинформации. Если разбиение

искомого уровня находится по дереву с индексированными связями, необходимые для вычисления средних значений яркости аддитивные характеристики областей определяются посредством суммирования характеристик элементов исходного разбиения, то процесс восстановления любого уровня занимает на ЭВМ IBM/PC-286 две-четыре секунды.

Дерево с индексированными связями (задаваемое массивом f и массивом индексов), а также массивы аддитивных характеристик областей исходного изображения позволяют в бескоординатной форме оперировать с элементами разбиений без учёта отношения соседства (смежности) элементов. Для расширения возможностей обработки полное бескоординатное многоуровневое описание видеоинформации помимо дерева с индексированными связями и аддитивных характеристик включает также таблицу связности, составленную для исходного изображения.

2.5. Формы организации данных. Аппроксимация объектов.

При проведении обработки видеоданных на основе многоуровневой модели следует иметь в виду, что хотя результаты вычислений и их содержательная интерпретация не зависят от способа нумерации областей и от последовательности выполнения операций слияния, тем не менее скорость расчётов, а также их алгоритмическая простота, в определённой степени связаны с особенностями организации информации.

Способ нумерации областей отражается и на свойствах таблицы связности. Если исходное изображение размечено так, что каждое объединение всех областей с номерами, не превышающими некоторого числа K , является связной областью [61], то просмотр таблицы связности можно осуществлять в режиме автосканирования, который сводится к следующему:

- организуется список неповторяющихся номеров строк (номеров областей), содержащий вначале единственную 1;
- последовательно просматриваются строки, номера которых встречаются в списке. При просмотре каждой строки список дополняется новыми номерами областей, соседствующих с областью, которой сопоставляется просматриваемая строка. Так продолжается до тех пор, пока список не будет исчерпан.

С завершением обработки списка просканированной оказывается вся таблица связности.

Автосканирование таблицы связности иллюстрируется схемой рис.14.

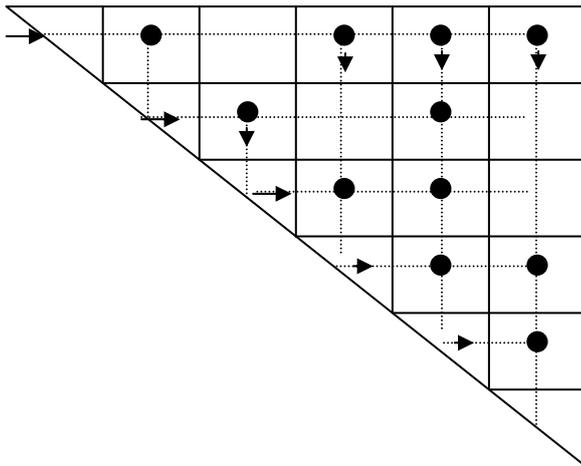


Рис. 14. Схема автосканирования таблицы связности.

Такая схема применима, например, при нумерации исходных областей изображения в порядке обычной телевизионной развёртки, что позволяет в процессе обработки элементов

таблицы связности при расчётах на ЭВМ автоматически обходить освобожденные адреса строк.

Необходимо обратить внимание на то, что одна и та же укрупнённая область изображения в зависимости от установленного порядка слияния компонент может описываться разными деревьями. Деревья с одинаковыми наборами узлов эквивалентны (рис. 15).

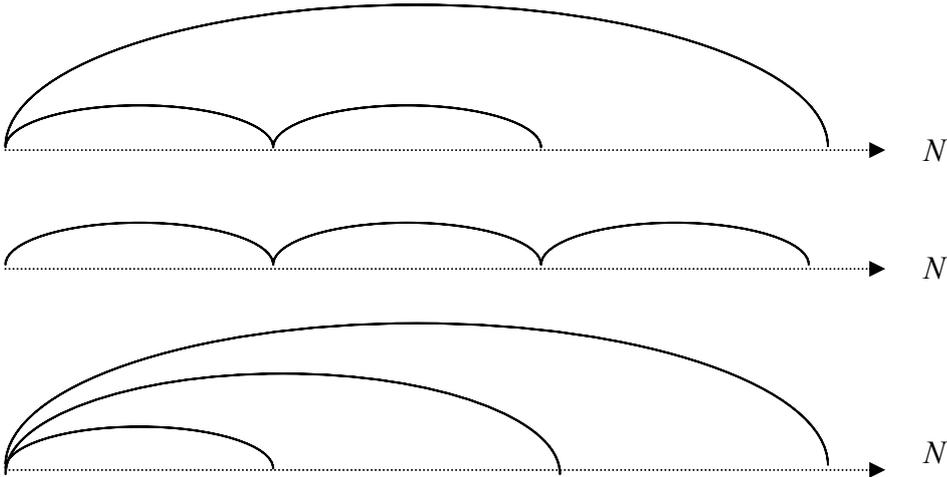


Рис. 15. Эквивалентные варианты описания области в виде дерева, заданного указанием связей между узлами на числовой оси

Выбор определённого варианта описания области обеспечивает строго упорядоченную структуру дерева многоуровневого представления. Если при формировании очередного уровня корневые вершины связываются последовательно (рис.16), то все связи данного узла результирующего дерева с узлами большими по номеру различаются между собой значениями ε_i . В альтернативном варианте (рис.17) связь каждого узла с узлом меньшим по номеру отличается по значению ε_i от остальных.

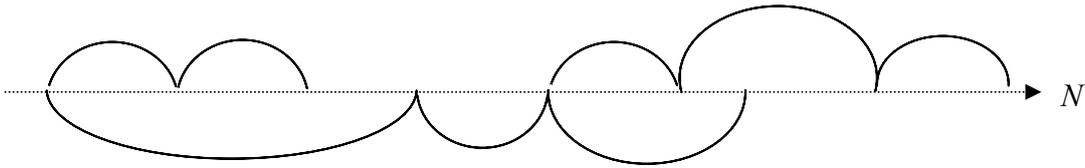


Рис. 16. Дерево последовательных связей. Показаны связи, установленные на уровне ε_1 (верхние) и на уровне ε_2 (нижние), где $\varepsilon_2 > \varepsilon_1$

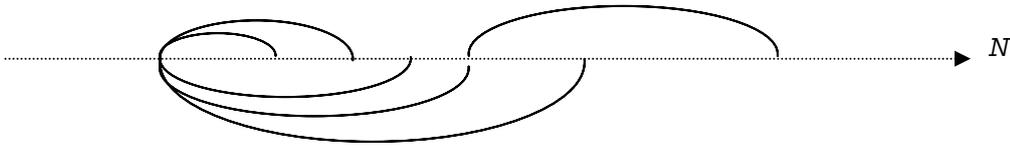


Рис. 17. Дерево прямых связей

При генерации многоуровневого представления имеет смысл не заботиться о форме представления деревьев. Для восстановления уровней представления при расчётах на ЭВМ лучше преобразовать результирующее дерево в дерево прямых связей, а строить приближение объектов изображения областями различных уровней проще с помощью дерева последовательных связей.

Алгоритм аппроксимации объектов (частей объектов) в процессе анализа видеоданных без циклического преобразования изображения описывается следующим.

Рассмотрим набор S тех узлов дерева многоуровневого

представления, каждый из которых хотя бы на одном из уровней обозначает область с заданными признаками.

Включим в указанный набор S дополнительно все промежуточные узлы, встречающиеся на пути по связям дерева от одного из них к другому.

Разорвём связи, последовательности которых устанавливаются при формировании областей каждого уровня, не удовлетворяющих используемым признакам, если на пути по этим связям встречается хотя бы один узел из S .

Из полученного набора S результирующей системы деревьев выделим набор s корневых.

Узлы из s являются корневыми узлами деревьев, представляющих части объектов. Соседствующие части объектов образуют сами объекты, разделение которых в рамках предложенной структуры данных сопряжено с преобразованием исходной таблицы связности.

Приведённый алгоритм в точности соответствует определению объектов п. 2.2, что использовалось практически при отладке программной реализации.

Эффективность применения разработанной структуры данных проверена экспериментально (рис. 18) на ЭВМ типа IBM PC - 286, в программах, рассчитанных на обработку изображений, имеющих до 16000 исходных областей. При обработке более сложных изображений эффективность автоматизированной обработки падает, что связано с ограничениями по быстродействию и по памяти, а также с некоторыми ограничениями применяемой операционной системы DOS.

В результате экспериментального исследования можно заключить, что на базе ЭВМ типа IBM PC - 286 может



Рис. 18. Результат выделения составных объектов на снимке морского дна (справа). Слева-исходное изображение

эффективно исследоваться действующая модель программной системы автоматического распознавания. Для создания реальных систем автоматического распознавания необходимо использование более мощных ЭВМ.

Самостоятельный интерес компьютерная модель распознавания представляет для интерпретации механизмов зрительного восприятия.

Выделение объектов на основе многоуровневой модели заключается в том, что по исходному изображению вычисляются связанные области различной формы и размеров, из которых затем конструируются наблюдаемые объекты. При этом удаётся без искажения формы представить объекты небольшим количеством составных частей. Параметры формируемых элементов зависят от предъявленного изображения. Аналоги связанных областей идентифицируются системой перестраивающихся рецептивных полей [20], ассоциированных с сетчаткой глаза человека.

При выделении объектов на основе многоуровневой модели учитываются все области различных уровней, что согласуется с предположением [47] об информативности промежуточных состояний, достигаемых в процессе формирования рецептивных полей нейронов. Подобно тому, как в зрительной системе человека осуществляется суммация сигналов от отдельных рецепторов [20] формирование связанных областей для описания объектов производится на основе оценки энергии [47] и других аддитивных характеристик.

Множество уровней обсуждаемого представления хранится и анализируется не в матричной форме, а в виде формальной структуры дерева со специальной индексацией узлов,

задаваемой парой массивов чисел. Один массив задаёт связи между узлами дерева, которые могут соответствовать нейронным связям, возникающим в колонках зрительных структур мозга. Второй массив ранжирует эти связи по интервалам времени обработки, что согласуется с гипотезой [47] об упорядоченном во времени восприятии пространственных соотношений и даёт определённую интерпретацию механизма зрительной памяти.

2.6. Метод анализа полутонового изображения по яркости.

При теоретическом и практическом развитии модели особое внимание следует обратить на возможность изменения применяемого критерия слияния областей, который влияет на скорость построения многоуровневого представления и определяет эффективность выявления объектов. Принципиальное изменение критерия состоит в том, что слияние смежных областей производится при условии их принадлежности некоторым установленным классам и автоматически прекращается, когда все области из рассматриваемых классов оказываются не смежными. Тогда последний уровень представления в той или иной мере сохраняет семантическое содержание исходного изображения, а результирующая классификация областей обеспечивает локализацию объектов.

При разбиении полутонового изображения на зоны по экстремальным значениям средней яркости слияние смежных областей определяется итеративным процессом сортировки локальных экстремумов, где под локальным экстремумом

понимается область локального максимума, не имеющая соседних областей с большей средней яркостью, или область локального минимума, в окрестности которой отсутствуют соседние меньшей яркости.

Процесс формирования разбиения изображения на зоны яркостных экстремумов состоит в следующем.

Пусть среди смежных областей рассматриваемого разбиения изображения нет совпадающих по средней яркости.

На первом шаге цикла обработки определяются все исходные локальные экстремумы. Затем среди остальных областей находятся новые, экстремальные по яркости. Новые локальные максимумы, являющиеся одновременно и новыми локальными минимумами (граничащие только с исходными локальными экстремумами обоих типов), из текущего рассмотрения исключаются. Далее производится слияние исходных локальных минимумов со смежными новыми локальными минимумами. Аналогичной процедуре подвергаются локальные максимумы.

Процесс наращивания исходных локальных экстремумов за счёт нахождения и присоединения к ним соседних новых повторяется до тех пор, пока все области рассматриваемого разбиения изображения не разделятся на области трёх типов: локальные минимумы, локальные максимумы и области смешанного типа.

По завершению формирования локальных экстремумов для полученного разбиения вычисляются новые средние яркости, производится слияние совпавших по средней яркости областей, и процесс сортировки локальных экстремумов повторяется сначала.

Окончательный результат локализации объектов наглядно представляется картиной с тремя градациями яркости (рис.19), в которой зоны локальных минимумов, максимумов и зоны экстремумов смешанного типа окрашены, соответственно, в белый, чёрный и серый цвета.

Как показали эксперименты, локализация объектов обычно завершается на третьем – четвёртом уровне. Площадь, занимаемая экстремумами смешанного типа, как правило, незначительна. При этом результаты локализации (рис.19) подобны результатам двухградационного преобразования с локальным выбором значений яркостного порога. Однако, в отличие от традиционных программ двухградационного преобразования [13], программы локализации не требуют указания значений параметра и применимы в более широкой области.

Высокая эффективность способа преобразования изображения в картину с тремя градациями яркости объясняется удачным выбором в рамках многоуровневой модели механизма слияния областей, логически согласованного с тем, что новые локальные экстремумы, в соответствии с определением, группируются рядом с одноимёнными исходными. Особенности организации данных позволяют обобщить полученный результат.

В силу того, что структура данных многоуровневой модели допускает описание изображения любой формы, каждую из областей результирующих локальных экстремумов (рис.19) можно рассматривать как отдельное изображение, достигая необходимой детализации. Для того, чтобы обработку всех областей проводить одновременно, достаточно упростить таблицу связности (рис.11), исключая описание соседства областей различного ти-



Рис. 19. Классификация областей рентгеновского снимка

па яркости. При этом, стирание точек в таблице связности обеспечивает в процессе расчётов игнорирование соседства определённых областей. Вычёркивание строк и симметричных этим строкам столбцов (рис.12) исключает из рассмотрения сами области.

В случае использования варианта с редукцией таблицы связности последовательная детализация изображения методом преобразования к трём градациям сводится к циклическому повторению итеративной сортировки локальных экстремумов, при которой в исходной таблице связности вычёркиваются строки (столбцы), отвечающие пограничным элементам областей каждого из трёх типов яркости. В предельном случае последовательного уточнения исходного изображения как набора отдельных областей различного типа яркости все точки разделяются на три вида: пограничные тёмные, пограничные светлые и пограничные серые. В целом результат детализации представляется в виде нового многоуровневого представления изображения в шести градациях точек, разделяемых по яркости на тёмные, серые и светлые и по геометрическому положению на внутренние и пограничные.

Построенную в результате итеративного анализа многоуровневую детализацию изображения можно описывать посредством индексации узлов некоторого дерева аналогично многоуровневому представлению изображения, достигаемому в результате итеративного слияния смежных областей. При этом роль градаций яркости играют шесть установленных типов областей.

Многоуровневое представление, таким образом, позволяет произвольное полутоновое изображение трансформировать в картину с малым числом градаций яркости, осуществить сегментацию с "памятью" и организовать выявление структурированных объектов с подсчётом числа составных частей.

Необходимо подчеркнуть, что к предложенному формализму многоуровневой модели следует относиться как к принципиальной схеме, в рамках которой допускаются определённые изменения и дополнения.

Развитие формализма обработки видеoinформации в бескоординатной форме заключается в тщательном отборе и обосновании признаков, по которым производится анализ областей. Кроме средней яркости и площади в качестве признаков можно использовать средние значения координат, линейные размеры областей и другие величины. Для современных технических приложений остаётся актуальным развитие методов корректного анализа объектов с учётом изменения ориентации и масштаба. В связи с нестабильностью условий съёмки задача корректного геометрического описания объектов возникает в большинстве технических приложений. Тем не менее, возможности её решения до сих пор не исчерпаны. Поэтому для повышения эффективности обработки изображений представляется перспективным развитие аппарата многоуровневой модели посредством исследования способов формального описания ориентации и масштаба объектов.

По сравнению с характеристиками объектов полутонных изображений признаки объектов двухградационных изображений следует признать более общими. Ориентация является одним из признаков, характерных для изображений обоих типов. Известно [47], что в зрительной системе человека определённую роль играют нейронные системы, opponентные к взаимно-ортогональным ориентациям. Поэтому для компьютерного моделирования механизмов зрительного восприятия развитие аппарата многоуровневого представления посредством исследования способов формального описания ориентации объектов также представляет интерес.

Выводы по главе 2.

Разработана алгебраическая модель изображения, обеспечивающая представление объектов минимальным количеством частей, синтезируемых посредством слияния элементов покрытия, в котором пересечение элементов эквивалентно включению.

Разработана динамическая структура видеоданных в виде дерева с индексированными связями, таблицы связности, набора аддитивных относительно слияния областей характеристик и функций от них.

Исследованы формы представления видеоданных, возможности оптимизации вычислений, способы преобразования видеоданных к матричному виду и аппроксимации структурированных объектов в бескоординатном представлении.

На основе алгоритма разделения изображения на зоны локальных экстремумов трёх типов предложен метод анализа полутоновых изображений по яркости посредством итеративной детализации.

Выбран способ развития и обобщения модели на основе формализации геометрического описания объектов.

ГЛАВА 3. ГЕОМЕТРИЧЕСКОЕ ОПИСАНИЕ И СТРУКТУРНЫЙ АНАЛИЗ ОБЪЕКТОВ

В главе на основе определения метрически изотропной фигуры и введения связанной с фигурой собственной системы координат строится система геометрических признаков, а также рассматривается метод структурного (частичного) анализа произвольных дискретных плоских фигур посредством разделения на фрагменты независимо от ориентации и масштаба.

Под метрически изотропной понимается фигура, линейный размер которой не зависит от направления, где под линейным размером понимается среднеквадратичное отклонение точек относительно оси, проходящей через центр инерции.

Доказывается, что любую фигуру, все точки которой не лежат на одной прямой, можно преобразовать в изотропную. При этом изотропная фигура совпадает с исходной по количеству точек и по площади, а также по сравнению с исходной имеет не меньшее число осей симметрии.

Отображение неизотропной фигуры в изотропную достигается некоторым линейным однородным невырожденным преобразованием и имеет смысл установления масштабов вдоль продольного и поперечного направлений пропорционально размерам фигуры, причём продольное и поперечное направление для неизотропной фигуры находится однозначно.

Выравнивание продольного и поперечного размеров трактуется как собственное преобразование Лоренца, рассматриваемое в двумерном евклидовом пространстве и являющееся одной из

компонент разложения двумерного однородного невырожденного преобразования в произведение положительно определённого симметричного и ортогонального. При этом гиперболический параметр собственного преобразования Лоренца описывает измеренное в логарифмическом масштабе отношение длины фигуры к ширине. Тригонометрический параметр ортогонального преобразования описывает ориентацию фигуры.

Формализация геометрических свойств объектов в виде плоских фигур строится исходя из рассмотрения преобразований, сохраняющих осевую симметрию.

Фигуре, представленной на плоскости набором точек, не лежащих на одной прямой, приписывается ориентация, определяемая поворотом осей координат при условии обращения в ноль смешанного момента, вычисленного относительно центра инерции. Для симметричного распределения точек ориентация с точностью до $\frac{\pi}{2}$ определяется направлением оси симметрии.

Нахождение ориентации фигуры описывается ортогонализацией пары векторов некоторого двумерного подпространства n -мерного векторного пространства, индуцированного n -ками значений одноимённых координат точек фигуры, причём повороту системы координат исходной плоскости сопоставляется, вообще говоря, неортогональное преобразование двумерного подпространства.

Посредством преобразования двумерного подпространства, отвечающего собственному преобразованию Лоренца исходной плоскости, достигается выравнивание длин рассматриваемых векторов, что трактуется в смысле преобразования фигуры в изотропную. Рассматриваются свойства изотропных фигур. Даются эк-

вивалентные определения.

Производится удобная для вычислений параметризация преобразований. Показывается, что суперпозиция поворота и собственного преобразования Лоренца инвариантна относительно перехода к двумерному подпространству.

Исследуются возможности развития геометрического описания фигур на основе обобщения формального приёма ортогонализации и выравнивания длин пары векторов двумерного подпространства за счёт сдвига системы координат исходной плоскости.

Аналитический формализм признакового описания объектов в виде дискретных фигур дополняется описанием варианта структурного (частичного) анализа, который сводится к преобразованию фигур в изотропные и разделению последних на некоторые "равновесные" части. Благодаря частичному анализу обеспечивается расширение признаковой системы и достигается необходимая точность интерпретации аналитического описания дискретных объектов, что служит предметом обсуждения заключительного раздела главы.

3.1. Аналитическое описание объектов, исходя из свойств осевой симметрии

3.1.1 Сохранение симметрии при линейных преобразованиях

Традиционный аппарат современных фундаментальных исследований, основанный на изучении инвариантов относительно

групповых преобразований, в области обработки изображений интенсивно применяется с момента зарождения и по настоящее время [37]. Его сочетание с удачной физической интерпретацией даёт наиболее сильный эффект. Известной методологией физической интерпретации являются построения на основе отображений симметрии.

Рассмотрим линейные невырожденные однородные преобразования, сохраняющие осевую симметрию дискретной фигуры, где под фигурой условимся понимать некоторое конечное множество точек двумерной плоскости. В общем случае точки фигуры могут составлять или не составлять связную область, а также совпадать или не совпадать по местоположению.

Осевая симметрия плоской дискретной фигуры, описываемой в произвольной системе координат, является, очевидно, инвариантом относительно одновременного отражения обеих осей, одинакового изменения масштаба обеих осей, сдвига начала координат и поворота. Кроме того, в системе координат, в которой ось симметрии параллельна координатной, симметрия инвариантна относительно отражения каждой из координатных осей, а также относительно растяжения одной оси при одновременном сжатии во столько же раз другой.

Свойства инвариантности осевой симметрии относительно поворота и растяжения одной оси с одновременным сжатием другой являются основой дальнейших рассуждений.

Как известно, линейное невырожденное преобразование координат точки к новым значениям может трактоваться двумя альтернативными способами. В первом случае считается, что новые координаты описывают новую точку (активная интерпретация),

во втором случае преобразование координат связывают с изменением описания точки при переходе к новому базису (пассивная интерпретация). Для простоты изложения условимся иногда комментировать формулы в терминологии характерной только для одной из интерпретаций, не оговаривая противоположной.

Следует заметить, что пассивная интерпретация преобразований дискретных изображений часто более удобна для вычислений, так как не индуцирует совпадения точек за счёт округления значений координат.

3.1.2 Центральные и средние значения координат. Условие координатной симметрии.

Пусть фигура задана на двумерной плоскости $x \times y$ распределением n точек x_i, y_i .

Центром распределения (фигуры) назовём пару координат \dot{x}, \dot{y} , определяемых как

$$\begin{cases} \dot{x} = \frac{1}{n} \sum_{i=1}^n x_i \\ \dot{y} = \frac{1}{n} \sum_{i=1}^n y_i \end{cases}$$

Средними координатами распределения будем считать следующие величины \bar{x}, \bar{y} :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n y_i}, \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i}$$

которые определены, соответственно, при $\sum_{i=1}^n y_i \neq 0$ и $\sum_{i=1}^n x_i \neq 0$.

Очевидно, если центр распределения не лежит на одной из осей координат и не совпадает с началом координат, то обе средние координаты существуют и лежат на одной прямой,

проходящей через начало координат: $\frac{\dot{x}}{\dot{y}} = \frac{\bar{x}}{\bar{y}}$.

Рассмотрим равенство

$$\sum_{i=1}^n x_i \cdot y_i = \frac{\left(\sum_{i=1}^n x_i\right) \cdot \left(\sum_{i=1}^n y_i\right)}{n}. \quad (1)$$

Если это равенство выполнено для данного распределения из n точек, то будем говорить, что "средние координаты совпадают с центральными".

Легко проверить, что если определено \bar{x} , то (1) эквивалентно $\dot{x} = \bar{x}$, и аналогично, если определено \bar{y} , то (1) эквивалентно $\dot{y} = \bar{y}$. Таким образом, справедливость (1) действительно означает совпадение центральных и средних значений, если последние определены.

Можно проверить, что справедливость равенства (1) не зависит от масштабирования, отражения осей и сдвига системы координат, то есть

$$\sum_{i=1}^n x_i \cdot y_i = \frac{\left(\sum_{i=1}^n x_i\right) \cdot \left(\sum_{i=1}^n y_i\right)}{n} \Leftrightarrow \sum_{i=1}^n \tilde{x}_i \cdot \tilde{y}_i = \frac{\left(\sum_{i=1}^n \tilde{x}_i\right) \cdot \left(\sum_{i=1}^n \tilde{y}_i\right)}{n},$$

S x W S y S x W S y

где $\begin{cases} x_i = \alpha \cdot \tilde{x}_i + a \\ y_i = \beta \cdot \tilde{y}_i + b \end{cases}$ (α, β, a, b – некоторые коэффициенты, причем $\alpha \cdot \beta \neq 0$).

Следствие.

Если рассматриваемое распределение из n точек симметрично относительно хотя бы одной оси, параллельной оси x или оси y , то равенство (1) справедливо.

Для того, чтобы в этом убедиться, достаточно преобразованием сдвига (при $\alpha = \beta = 1$) совместить начало координат с центральными координатами распределения, тогда

$$\sum_{i=1}^n \tilde{x}_i = \sum_{i=1}^n \tilde{y}_i = 0 \quad \text{и}$$

правая часть (1) в новой системе координат обратится в 0 по построению, а левая часть окажется равной 0, вследствие симметрии распределения относительно по крайней мере одной из новых осей \tilde{x} или \tilde{y} .

Будем говорить, что рассматриваемое распределение в данной системе координат обладает координатной симметрией, если оно имеет хотя бы одну ось симметрии, параллельную координатной оси.

Таким образом, равенство (1) является необходимым условием координатной симметрии.

3.1.3 Запись в 2-мерном подпространстве n -мерного векторного пространства.

Для краткости записи удобно использовать представление соотношений с помощью n -го векторного пространства.

Двумерному распределению из n точек $(\tilde{x}_i, \tilde{y}_i) = 0$ на плоскости $X \times Y$ поставим в соответствие пару X, Y векторов в n -ом евклидовом векторном пространстве с ортонормированным базисом

e_1, e_2, \dots, e_n :

$$\begin{cases} X = \sum_{i=1}^n x_i \cdot e_i \\ Y = \sum_{i=1}^n y_i \cdot e_i \end{cases} \quad (e_i, e_k) = \delta_{i,k}$$

Для простоты можно считать, что все точки исходного распределения различаются по местоположению.

Невырожденному линейному преобразованию координат плоскости $x \times y$ сопоставим преобразование векторов X, Y к новым значениям \tilde{X}, \tilde{Y} в n -ом векторном пространстве:

$$\begin{cases} \tilde{X} = \sum_{i=1}^n \tilde{x}_i \cdot e_i \\ \tilde{Y} = \sum_{i=1}^n \tilde{y}_i \cdot e_i \end{cases} .$$

где \tilde{x}_i, \tilde{y}_i - новые координаты точек распределения (координаты точек распределения в новой системе координат, либо, в альтернативном случае, координаты точек нового распределения в исходной системе)

С помощью введённого евклидова n -го векторного пространства соотношение (1) запишется в виде

$$(X, Y) - (X, E)(Y, E) \equiv (X', Y') = 0 \quad , \quad (2)$$

где $\begin{cases} X' = X - (X, E)E \\ Y' = Y - (Y, E)E \end{cases}$ - компоненты n -ых векторов X и Y ,

ортогональные нормированному вектору $E = \frac{1}{\sqrt{n}} \sum_{i=1}^n e_i$, $(E, E) = 1$.

В исходной форме компоненты штрихованных векторов имеют вид

$$\begin{cases} x'_i = x_i - \dot{x} \\ y'_i = y_i - \dot{y} \end{cases} .$$

Значения компонент штрихованных n -мерных векторов определяются координатами точек, вычисленными на двумерной плоскости относительно центра распределения.

$$\text{Координаты центра распределения имеют вид } \begin{cases} \dot{x} = \frac{1}{\sqrt{n}}(X, E) \\ \dot{y} = \frac{1}{\sqrt{n}}(Y, E) \end{cases}.$$

При преобразованиях координат преобразуются и координаты центра распределения.

Условия $\sum_{i=1}^n x_i = 0$ и $\sum_{i=1}^n y_i = 0$ записываются, соответственно, в виде $(X, E) = 0$ и $(Y, E) = 0$.

Преобразованные штрихованные n -ые вектора имеют вид

$$\begin{cases} \tilde{X}' = \tilde{X} - (\tilde{X}, E)E \\ \tilde{Y}' = \tilde{Y} - (\tilde{Y}, E)E \end{cases}$$

С операцией линейного преобразования переход к штрихованным векторам коммутирует.

Далее X' и Y' , рассматриваемые как элементы подпространства, порожденного их линейными комбинациями, обозначаются u и v :

$$\begin{cases} u \equiv \sqrt{(u, u)}(k \cdot \cos \varphi_1 + m \cdot \sin \varphi_1) = X' \\ v \equiv \sqrt{(v, v)}(k \cdot \cos \varphi_2 + m \cdot \sin \varphi_2) = Y' \end{cases}, \quad (3)$$

где k, m - некоторый ортонормированный базис подпространства $u \times v$.

Скалярные произведения (u, u) , (v, v) и (v, v) - это обычные вторые моменты, часто обозначаемые I_x, I_y, I_{xy} .

$$\text{Очевидно, } \cos(\varphi_1 - \varphi_2) = \frac{(u, v)}{\sqrt{(u, u)(v, v)}}. \quad (4)$$

Из неравенства Коши - Буняковского следует, что $\cos^2(\varphi_1 - \varphi_2)$ равен 1 тогда и только тогда, когда u и v линейно зависимы, то есть все n точек исходного распределения лежат на одной прямой и распределение вырожденно [65]. Далее рассматриваются невырожденные распределения, для которых u и v линейно независимы и

$$\sin(\varphi_1 - \varphi_2) \neq 0 \quad (5)$$

В двумерном подпространстве $u \times v$ n -го пространства необходимое условие координатной симметрии (1-2) сводится к условию ортогональности векторов u и v

$$(u, v) = 0 .$$

Для дальнейших выкладок достаточно ограничиться невырожденными линейными однородными преобразованиями. Каждому преобразованию A плоскости исходного двумерного распределения сопоставляется аналогичное преобразование B двумерного подпространства $u \times v$

$$A = \begin{cases} \tilde{x}_i = \alpha_{11}x_i + \alpha_{12}y_i \\ \tilde{y}_i = \alpha_{12}x_i + \alpha_{22}y_i \end{cases} \Rightarrow B = \begin{cases} \tilde{u} = \alpha_{11}u + \alpha_{12}v \\ \tilde{v} = \alpha_{12}u + \alpha_{22}v \end{cases}$$

где $\alpha_{11} \cdot \alpha_{22} - \alpha_{12} \cdot \alpha_{21} \neq 0$.

3.1.4 Поворот. Определение ориентации.

Рассмотрим поворот на угол φ системы координат исходной плоскости относительно исследуемого распределения из n точек. В двумерном подпространстве $u \times v$ он описывается преобразованием W :

$$\begin{cases} W\{u\} = u \cdot \cos \varphi + v \cdot \sin \varphi \\ W\{v\} = -u \cdot \sin \varphi + v \cdot \cos \varphi \end{cases}$$

Преобразование W , вообще говоря, не ортогонально.

Пусть преобразованные вектора ортогональны, то есть удовлетворяют необходимому условию координатной симметрии

$$(W\{u\}, W\{v\}) = (u, v) \cdot \cos 2\varphi - \frac{1}{2}[(u, u) - (v, v)] \cdot \sin 2\varphi = 0. \quad (6)$$

Определим величину d :

$$d = \sqrt{\frac{1}{4} \left(\sqrt{\frac{(u, u)}{(v, v)}} - \sqrt{\frac{(v, v)}{(u, u)}} \right)^2 + \frac{(u, v)^2}{(u, u)(v, v)}} \quad (7)$$

Очевидно, $d = 0$, если u и v ортогональны и совпадают по длине.

При $d = 0$ условимся [65] называть рассматриваемое относительно данной системы координат распределение из n точек изотропным, в противном случае - неизотропным.

Предполагается, что исследуемое распределение неизотропное

$$d \neq 0.$$

Введём в рассмотрение некоторый гиперболический угол γ , определяемый выражением:

$$\operatorname{sh} \gamma = -\frac{1}{2} \left(\sqrt{\frac{(u, u)}{(v, v)}} - \sqrt{\frac{(v, v)}{(u, u)}} \right). \quad (8)$$

из которого следует

$$\operatorname{ch} \gamma = \frac{1}{2} \left(\sqrt{\frac{(u, u)}{(v, v)}} + \sqrt{\frac{(v, v)}{(u, u)}} \right) \geq 1.$$

Инвариантные относительно общего изменения масштаба "безразмерные" тригонометрические и гиперболические функции обеспечивают компактное представление соотношений. Так, величина d (7) выражается через $\cos(\varphi_1 - \varphi_2)$ (4) и $\operatorname{sh} \gamma$ (8) в следующей форме:

$$d = \sqrt{\cos^2(\varphi_1 - \varphi_2) + \operatorname{sh}^2 \gamma} \quad (9)$$

Последнее выражение с учётом (4, 6) можно считать определением d , которое представляется также следующим образом:

$$\begin{aligned} d &= \sqrt{(\operatorname{ch} \gamma \cdot \cos(\varphi_1 - \varphi_2))^2 + (\operatorname{sh} \gamma \cdot \sin(\varphi_1 - \varphi_2))^2} \equiv \\ &\equiv \sqrt{\operatorname{ch}^2 \gamma - \sin^2(\varphi_1 - \varphi_2)} \end{aligned}$$

Условие координатной симметрии (6) удобно записать в виде:

$$(W\{u\}, W\{v\}) = \sqrt{(u, u)(v, v)} \cdot d \cdot \sin 2(\varphi_0 - \varphi) = 0, \quad (10)$$

где φ_0 находится из соотношений

$$\begin{cases} \sin 2\varphi_0 = \sigma \frac{\cos(\varphi_1 - \varphi_2)}{d} \\ \sin 2\varphi_0 = -\sigma \frac{\operatorname{sh} \gamma}{d} \end{cases} \quad (11)$$

Здесь параметр σ принимает значения ∓ 1 .

Значения φ_0 берутся в полуоткрытом промежутке от 0 до π , различаются между собой на $\frac{\pi}{2}$ и в исходной плоскости $X \times Y$ определяют поворот, описываемый в двумерном подпространстве $U \times V$ преобразованием W , с точностью до знака.

Следствие.

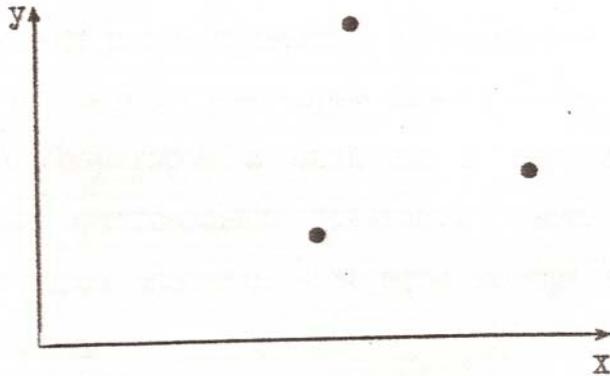
Все распределения конечного числа точек, имеющие неортогональные оси симметрии, изотропные, то есть описываются нулевой величиной d .

Можно проверить, что для случая вырожденного неизотропного распределения точек, не лежащих на координатной оси, по формулам (11) находятся две ортогональные прямые, одна из которых проходит через точки распределения. Таким образом, смысл (11) сохраняется.

Пример.

Пусть дано симметричное распределение из 3-х точек:

№	x_i	y_i
1	8	3
2	14	15
3	9	10



Из (8-9) и (11) имеем:

$$\cos(\varphi_1 - \varphi_2) = -3/\sqrt{1209}; \quad \text{sh } \gamma = 4/\sqrt{1209}; \quad d = 5/\sqrt{1209};$$

$$\begin{cases} \sin 2\varphi_0 = -\sigma \frac{3}{3} \\ \sin 2\varphi_0 = -\sigma \frac{4}{5} \end{cases} \Rightarrow \varphi_0 = \arctg \frac{3}{4} \quad \vee \quad \varphi_0 = \frac{\pi}{2} + \arctg \frac{3}{4};$$

Таким образом, для симметричной неизотропной невырожденной фигуры формулы (11) с точностью до угла $\frac{\pi}{2}$ позволяют установить направление оси симметрии.

Для несимметричной неизотропной невырожденной фигуры формулы (11) задают определённую ориентацию относительно системы координат. При этом следует указать, что соотношения (11) определяют оси координат, совпадающие с прямыми, вычисляемыми [38] методом наименьших квадратов при условии минимизации суммы квадратов расстояний до искомой прямой.

3.1.5 Собственные преобразования Лоренца.

Приведение к собственным осям.

В плоскости $x \times y$ рассмотрим двумерные собственные преобразования Лоренца.

В отличие от традиционного представления [51] собственные преобразования Лоренца рассматриваются в евклидовом пространстве, в котором описываются некоторым симметричным (вещественным эрмитовым) оператором и сводятся к растяжению вдоль одного из взаимно ортогональных "световых" векторов с одновременным сжатием вдоль другого. При этом коэффициенты сжатия и растяжения совпадают.

Если на исходной плоскости $x \times y$ собственные преобразова-

ния Лоренца задаются в базисе собственных векторов, то в двумерном подпространстве $u \times v$ они описываются преобразованием L :

$$\begin{cases} L\{u\} = (\operatorname{ch} \theta + \operatorname{sh} \theta) \cdot u \\ L\{v\} = (\operatorname{ch} \theta - \operatorname{sh} \theta) \cdot v \end{cases}$$

где θ - параметр, введённый Минковским.

Преобразование L , очевидно, сохраняет скалярное произведение собственных векторов.

Рассмотрим преобразование $W \cdot L$:

$$WL: \begin{cases} WL\{u\} = e^\theta (u \cdot \cos \varphi + v \cdot \sin \varphi) \\ WL\{v\} = e^{-\theta} (-u \cdot \sin \varphi + v \cdot \cos \varphi) \end{cases} \quad (12)$$

Очевидно, $W \cdot L$ сохраняет скалярное произведение векторов $W\{u\}$ и $W\{v\}$:

$$(WL\{u\}, WL\{v\}) = (W\{u\}, W\{v\}) \quad . \quad (13)$$

Потребуем, чтобы $W \cdot L\{u\}$ и $W \cdot L\{v\}$ оказались ортогональны и совпали по длине, что отвечает переходу на исходной плоскости $X \times Y$ к новым координатам, для которых справедливо условие координатной симметрии, с последующим взаимно обратным изменением масштабов новых координатных осей. При этом по-прежнему предполагается, что u и v описывают некоторое неизотропное невырожденное распределение.

В силу (13) условие ортогональности $(W \cdot L\{u\}, W \cdot L\{v\}) = 0$ сведётся к (10), а условие нормирования $(W \cdot L\{u\}, W \cdot L\{u\}) = (W \cdot L\{v\}, W \cdot L\{v\})$ с использованием (4) и (5)

запишется в виде:

$$\operatorname{th} 2\theta_0 = \frac{\cos 2\varphi_0 \cdot \operatorname{sh} \gamma - \sin 2\varphi_0 \cdot \cos(\varphi_1 - \varphi_2)}{\operatorname{ch} \gamma},$$

где φ_0 , θ_0 - искомые значения параметров φ и θ .

Значения φ_0 , определяющие $W \cdot L$ с точностью до знака, даются (11). Исключая с помощью (11) из последнего равенства $\cos 2\varphi_0$ и $\sin 2\varphi_0$ и используя (9), после деления числителя и знаменателя на d получаем окончательное выражение для θ_0 :

$$\operatorname{th} 2\theta_0 = -\sigma \frac{d}{\operatorname{ch} \gamma} \quad (14)$$

Как следствие (14) можно получить выражения для $\operatorname{sh} 2\theta_0$ и $\operatorname{ch} 2\theta_0$:

$$\begin{aligned} \operatorname{sh} 2\theta_0 &= -\sigma \frac{d}{|\sin(\varphi_1 - \varphi_2)|} \\ \operatorname{ch} 2\theta_0 &= \frac{\operatorname{ch} \gamma}{|\sin(\varphi_1 - \varphi_2)|} \end{aligned}$$

Нетрудно убедиться, что квадрат длины векторов $W \cdot L\{u\}$ и $W \cdot L\{v\}$, полученных в результате преобразования $W \cdot L$ с параметрами φ_0 , θ_0 выражается в виде:

$$(W \cdot L\{u\}, W \cdot L\{u\}) \equiv (W \cdot L\{v\}, W \cdot L\{v\}) = \sqrt{(u, u)(v, v)} \cdot |\sin(\varphi_1 - \varphi_2)|, \quad (15)$$

Проще всего проверить (15), если решить систему (12) относительно u, v :

$$\begin{cases} u = e^{-\theta} \cdot \cos \varphi \cdot WL\{u\} - e^{\theta} \cdot \sin \varphi \cdot WL\{v\} \\ v = e^{-\theta} \cdot \sin \varphi \cdot WL\{u\} + e^{\theta} \cdot \cos \varphi \cdot WL\{v\} \end{cases} \quad (16)$$

и затем вычислить пропорциональную $\sin^2(\varphi_1 - \varphi_2)$ величину $(u, u)(v, v) - (u, v)^2$, учитывая ортогональность и совпадение

длин $W \cdot L\{u\}, W \cdot L\{v\}$.

В исходной плоскости $X \times Y$ прообраз преобразования $W \cdot L$ означает переход к некоторым новым координатам, которые назовём собственными. Относительно собственных координат рассматриваемое распределение является изотропным.

Таким образом, произвольное неизотропное невырожденное распределение можно преобразовать в изотропное или описать как изотропное в собственной системе координат.

3.1.6 Интерпретация.

Для того, чтобы уяснить смысл параметра θ_0 , рассмотрим общий случай изотропного невырожденного распределения из 3-х точек с центром в начале координат.

Координаты точек удовлетворяют системе уравнений:

$$\begin{cases} x_1 + x_2 + x_3 = y_1 + y_2 + y_3 = 0 \\ x_1 y_1 + x_2 y_2 + x_3 y_3 = 0 \\ x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2 \end{cases},$$

где соотношения в первой строке выражают условие совпадения центра распределения с началом координат, а остальные эквивалентны требованию $d = 0$.

Методом исключения неизвестных находится общее решение:

$$\begin{cases} x_1 = \frac{\varepsilon}{\sqrt{3}}(t_2 - t_1); & y_1 = -t_1 - t_2 \\ x_1 y_1 + x_2 y_2 + x_3 y_3 = 0; & y_2 = t_1 \\ x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2; & y_3 = t_2 \end{cases},$$

где t_1, t_2, t_3 - действительные параметры, а $\varepsilon = \pm 1$.

Используя общее решение, нетрудно установить, что выполняются соотношения:

$$\begin{aligned} x_1^2 + y_1^2 &= x_2^2 + y_2^2 = x_3^2 + y_3^2 = \frac{4}{3}(t_1^2 + t_1 t_2 + t_2^2) \\ (x_1 - x_2)^2 + (y_1 - y_2)^2 &= (x_1 - x_3)^2 + (y_1 - y_3)^2 = (x_2 - x_3)^2 + (y_2 - y_3)^2 = \\ &= 4(t_1^2 + t_1 t_2 + t_2^2)^2 \end{aligned}$$

Следовательно, три произвольные точки, не лежащие на одной прямой и образующие неизотропное распределение, в собственной системе координат составляют равносторонний треугольник.

Таким образом, в собственных координатах исходные свойства осевой симметрии фигуры усиливаются. Значение θ_0 можно трактовать как меру отличия распределения от изотропной фигуры некоторой "правильной" формы.

Для того, чтобы показать логичность названия изотропной фигуры следует рассмотреть метрические соотношения, характеризующие линейные размеры.

Основными метрическими соотношениями, характеризующими размеры произвольной фигуры, являются длина и ширина, которые как интегральные характеристики представляются в виде среднеквадратичных отклонений, измеряемых относительно взаимно-ортогональных осей. В качестве указанных осей имеет смысл выбрать оси координат, определяющие ориентацию фигуры. Тогда, в соответствии с п.3.1.4 длина и ширина фигуры определяются в виде:

$$\begin{aligned} \text{длина} &= \max \left(\sqrt{\frac{(W\{u\}, W\{u\})}{n}}, \sqrt{\frac{(W\{v\}, W\{v\})}{n}} \right) \\ \text{ширина} &= \min \left(\sqrt{\frac{(W\{u\}, W\{u\})}{n}}, \sqrt{\frac{(W\{v\}, W\{v\})}{n}} \right) \end{aligned}$$

При этом, как уже отмечалось, продольное и поперечное направления отождествляются с осями, проведёнными по методу наименьших квадратов [38].

Воспользовавшись явным выражением для $W\{u\}$ и $W\{v\}$, а также соотношениями (8-9, 14) получим [65]:

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

где длина и ширина фигуры вычисляются при альтернативном выборе знака.

Из выражений для $\text{sh } 2\theta_0$, $\text{ch } 2\theta_0$ и $\text{sh } 2\theta_0$ (согласно следствиям из (14)) с учётом выражения d через $\sin(\varphi_1 - \varphi_2)$ (4) и $\text{ch } \gamma$ (согласно следствию из (9)) следует:

$$e^{2\theta_0} = \sqrt{\frac{\text{ch } \gamma - \sigma \cdot d}{\text{ch } \gamma + \sigma \cdot d}} .$$

Сравнивая два последних соотношения, получаем выражение для θ_0 :

$$\theta_0 = -\frac{\sigma}{2} \cdot \ln \left(\frac{\text{длина}}{\text{ширина}} \right) .$$

Непосредственно из выражений для длины и ширины следует:

$$d = n \cdot \frac{(\text{длина})^2 - (\text{ширина})^2}{2 \cdot \sqrt{(u, u) \cdot (v, v)}} .$$

Таким образом, θ_0 характеризует измеренное в логарифмическом масштабе отношение длины фигуры к её ширине, а величина d с точностью до множителя совпадает с разностью квадратов размеров фигуры.

Длина и ширина изотропной фигуры, для которой $d = 0$, очевидно, совпадают. Непосредственно по любой из формул (12) легко проверить, что в случае изотропной фигуры, благодаря ортогональности и совпадению длин векторов u и v , среднеквадратичное отклонение точек относительно произвольной оси, проведённой через центр инерции, не зависит от направления оси (Заключение об инвариантности линейного размера изотропной фигуры можно вывести также из (16))

Таким образом, название изотропной фигуры не случайно.

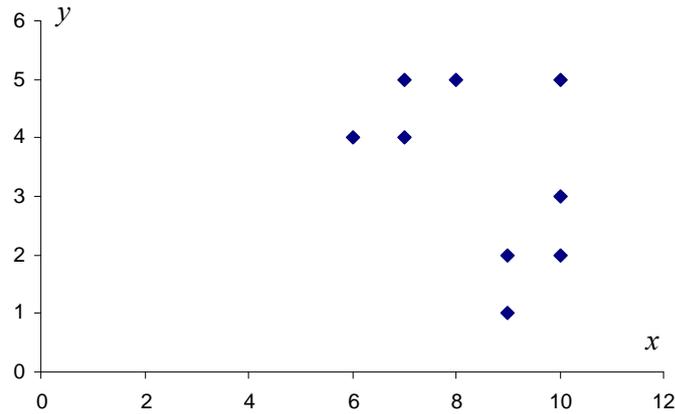
Параметр θ_0 в случае изотропной фигуры имеет нулевое значение. Верно и обратное: если параметр θ_0 равен 0, то фигура изотропна.

Следует заметить, что само по себе совпадение среднеквадратичных отклонений относительно осей координат не означает изотропности фигуры.

Пример.

Пусть дано распределение из 9 точек, расположенных симметрично относительно оси, параллельной биссектрисе координатного угла:

№	x_i	y_i
1	6	4
2	7	4
3	7	5
4	8	5
5	10	5
6	10	3
7	10	2
8	9	2
9	9	1



Имеем:

$$\cos(\varphi_1 - \varphi_2) = -35/82, \quad \text{sh } \gamma = 0, \quad d = 35/82.$$

$$\begin{cases} \sin 2\varphi_0 = -\sigma \\ \cos 2\varphi_0 = 0 \end{cases} \Rightarrow \varphi_0 = \frac{\pi}{4} \vee \varphi_0 = \frac{3\pi}{4},$$

$$\text{th } 2\theta_0 = -\sigma \neq 0.$$

Условимся вычислять площадь, как произведение длины и ширины. Тогда, из выражения для длины и ширины получим с учётом выражения d через $\sin(\varphi_1 - \varphi_2)$ и $\text{ch } \gamma$ выражение для площади произвольной фигуры:

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

Сравнивая последнее выражение с результатом (15), можно заключить, что площадь при преобразовании невырожденной фигуры в изотропную остаётся неизменной [65].

Средняя плотность точек фигуры, приходящихся на единицу площади, имеет вид:

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

Квадратный корень из последнего выражения определяет масштабный множитель, необходимый для преобразования исходной фигуры в фигуру со стандартной плотностью точек.

Резюмируя обсуждение основных метрических свойств плоских дискретных фигур на основе понятия изотропной фигуры, можно заключить следующее.

Изотропная фигура эквивалентным образом определяется:

- нулевым значением d ;
- ортогональностью и нормировкой на одну величину векторов u и v ;
- единым среднеквадратичным линейным размером, не зависящим от направления измерения;
- совпадением длины и ширины при условии измерения линейных размеров вдоль осей, вычисленных методом наименьших квадратов.

Произвольная невырожденная фигура преобразуется в изотропную посредством сжатия вдоль продольной оси и растяжения в перпендикулярном направлении. Для неизотропной фигуры продольное и поперечное сечение определяется однозначно.

Характерно, что изотропная фигура по площади совпадает с исходной и по сравнению с исходной имеет не меньшее число осей симметрии. При этом средние и центральные координаты изотропной фигуры совпадают независимо от ориентации координатных осей.

По сравнению с известным способом [69] классификации объектов на основе интегральных метрических соотношений, в

котором используется особенность равностороннего треугольника и тетраэдра сохранять "свойство суммы расстояний из любой точки внутри фигуры до её оснований", предложенное геометрическое описание произвольных фигур посредством преобразования в изотропные представляется теоретически более развитым. Тем не менее, идея трёхкомпонентного представления объектов представляет самостоятельный интерес [54].

3.1.7 Разложение $W \cdot L$ в произведение ортогонального и симметричного операторов. Отождествление плоскостей.

Следует отметить, что ввести понятие изотропной фигуры и применять его для описания свойств произвольных дискретных плоских фигур, можно, избегая нетрадиционной формы интерпретации аппарата векторных пространств. Однако, такая форма представления оказалась наиболее удобной при разработке исходных положений и, по мнению автора, может представлять самостоятельное значение для их развития.

Благодаря введению n -мерного пространства суперпозиции $V \cdot L$ поворота и собственного преобразования Лоренца исходной плоскости $X \times Y$ сопоставляется однородное линейное преобразование $W \cdot L$ двумерного подпространства $u \times v$. $W \cdot L$ можно разложить в произведение некоторого ортогонального преобразования (поворота V') и положительно определённого симметричного (собственного преобразования Лоренца L'). При этом возникает идея исследовать цепочку преобразований $V \cdot L, V' \cdot L', \dots$ которую можно получить, если менять

плоскости $x \times y$ и $u \times v$ ролями и продолжать отображение преобразований. Оказывается, указанная цепочка обрывается на первом же шаге, что утверждает следующая теорема.

Теорема.

Пусть u, v - пара линейно независимых векторов n -мерного пространства, характеризующихся ненулевой величиной d (9) и порождающих двумерное подпространство $u \times v$, пусть $W \cdot L$ - линейный однородный оператор преобразования пары векторов u, v в ортогональные вектора $W \cdot L\{u\}, W \cdot L\{v\}$ равной длины, в явном виде задаваемый (12) со значениями параметров $\varphi = \varphi_0$, $\theta = \theta_0$, которые в свою очередь определяются (11) и (14).

Пусть далее в двумерном подпространстве $u \times v$ рассматриваются также следующие операторы.

V - оператор поворота двумерного подпространства $u \times v$, задаваемый действием обратного оператора V^{-1} на ортогональные вектора $W \cdot L\{u\}, W \cdot L\{v\}$:

$$V^{-1} : \begin{cases} V^{-1} \cdot W \cdot L\{u\} = W \cdot L\{u\} \cdot \cos \varphi_0 - W \cdot L\{v\} \cdot \sin \varphi_0 \\ V^{-1} \cdot W \cdot L\{v\} = W \cdot L\{u\} \cdot \sin \varphi_0 + W \cdot L\{v\} \cdot \cos \varphi_0 \end{cases} . \quad (17)$$

L - оператор двумерного собственного преобразования Лоренца подпространства $u \times v$, который с помощью ортонормированных векторов k, m из подпространства $u \times v$ задаётся в виде:

$$L\{w\} = \text{ch } \theta_0 \cdot w + \text{sh } \theta_0 \cdot [(w, k) \cdot m + (w, m) \cdot k] , \quad (18)$$

где W - произвольный вектор из подпространства $u \times v$; k - орт, направленный вдоль "временной" ([52]) оси; m - орт "скорости".

Тогда при определённой ориентации векторных параметров k, m собственных преобразований Лоренца относительно исходных u, v , такой, что

$$\begin{cases} \sin(\varphi_1 + \varphi_2) = \sigma \frac{\text{ch } \gamma}{d} \cos(\varphi_1 - \varphi_2) \\ \cos(\varphi_1 + \varphi_2) = -\sigma \frac{\text{sh } \gamma}{d} \sin(\varphi_1 - \varphi_2) \end{cases} \quad (19)$$

достигается:

$$W \cdot L = V \cdot L \quad . \quad (20)$$

При этом условие (19) согласно (3) задаёт в базисе k, m положение биссектрисы угла между векторами u и v .

Доказательство.

Используя разложение u, v по ортогональным $W \cdot L\{u\}$, $W \cdot L\{v\}$ (16), нетрудно найти скалярные произведения $W \cdot L\{u\}$, $W \cdot L\{v\}$ на u, v , затем из (17) вычислить $(V^{-1} \cdot W \cdot L\{u\}, u)$, $(V^{-1} \cdot W \cdot L\{v\}, v)$, $(V^{-1} \cdot W \cdot L\{u\}, v)$, $(V^{-1} \cdot W \cdot L\{v\}, u)$ и убедиться, что последние два совпадают. Вычисляя согласно (18) аналогичные скалярные произведения для L , можно свести доказательство (20) к проверке тождеств:

$$\begin{aligned} -\text{sh } \theta_0 \cdot \sin 2\varphi_0 |\sin(\varphi_1 - \varphi_2)| &= \text{ch } \theta_0 \cdot \cos(\varphi_1 - \varphi_2) + \text{sh } \theta_0 \cdot \sin(\varphi_1 + \varphi_2) \\ (\text{ch } \theta_0 - \text{sh } \theta_0 \cdot \cos 2\varphi_0) |\sin(\varphi_1 - \varphi_2)| &= \sqrt{\frac{(u, u)}{(v, v)}} (\text{ch } \theta_0 + \text{sh } \theta_0 \cdot \sin 2\varphi_1) \\ (\text{ch } \theta_0 + \text{sh } \theta_0 \cdot \cos 2\varphi_0) |\sin(\varphi_1 - \varphi_2)| &= \sqrt{\frac{(v, v)}{(u, u)}} (\text{ch } \theta_0 + \text{sh } \theta_0 \cdot \sin 2\varphi_2) \end{aligned} \quad (21)$$

Для того, чтобы проверить справедливость (21), достаточно представить тождества посредством (4), (8), (9) (11), (14) в виде системы трансцендентных уравнений

относительно $\varphi_1 - \varphi_2$ и γ . При этом удобно преобразовать функции от θ_0 в функции от $2\theta_0$, домножая все равенства на $2 \operatorname{ch} \theta_0$, и вместо двух последних рассматривать полуразность и полусумму. Тогда первое тождество сведётся к сумме двух нулевых слагаемых:

$$\begin{aligned} \operatorname{ch} 2\theta_0 \cdot \cos(\varphi_1 - \varphi_2) + \operatorname{sh} 2\theta_0 \cdot \sin(\varphi_1 + \varphi_2) &= 0 \\ \cos(\varphi_1 - \varphi_2) + \operatorname{sh} 2\theta_0 \cdot |\sin(\varphi_1 - \varphi_2)| \cdot \sin 2\varphi_0 &= 0 \end{aligned} \quad (22)$$

а два другие тождества из (21) запишутся в виде:

$$\begin{aligned} (1 + \operatorname{ch} 2\theta_0) \cdot |\sin(\varphi_1 - \varphi_2)| &= \alpha \cdot \operatorname{ch} \gamma - \beta \cdot \operatorname{sh} \gamma \\ -\operatorname{sh} 2\theta_0 \cdot |\sin(\varphi_1 - \varphi_2)| \cdot \cos 2\varphi_0 &= -\alpha \cdot \operatorname{sh} \gamma + \beta \cdot \operatorname{ch} \gamma \end{aligned} \quad (23)$$

$$\text{где } \begin{cases} \alpha = 1 + \operatorname{ch} 2\theta_0 + \operatorname{sh} 2\theta_0 \cdot \sin(\varphi_1 + \varphi_2) \cdot \cos(\varphi_1 - \varphi_2) \\ \beta = \operatorname{sh} 2\theta_0 \cdot \cos(\varphi_1 + \varphi_2) \cdot \sin(\varphi_1 - \varphi_2) \end{cases}.$$

Наконец, чтобы убедиться в справедливости системы (21)–(23), достаточно функции с аргументами $2\varphi_0$, $2\theta_0$, $\varphi_1 + \varphi_2$ заменить эквивалентными выражениями через γ и $\varphi_1 - \varphi_2$, воспользовавшись (11), следствиями (14), а также (19) и (9).
Следствие.

Произведению поворота и собственного преобразования Лоренца, заданному на плоскости $x \times y$, отвечает произведение одноимённых преобразований двумерного подпространства $u \times v$:

$$V(\varphi_0) \cdot L(\theta_0) \Big|_{x \times y} \Rightarrow V(\varphi_0) \cdot L(\theta_0) \Big|_{u \times v}.$$

В отношении суперпозиции поворота и собственного преобразования Лоренца плоскость $x \times y$ и подпространство $u \times v$

не различаются. Поэтому допустимо считать, что исходная плоскость $x \times y$ совпадает с двумерным подпространством $u \times v$ n -мерного пространства, порожденного набором точек рассматриваемой фигуры.

Следует заметить, что это заключение не является очевидным, так как для каждой из компонент суперпозиции оно, вообще говоря, не справедливо.

Содержание представления $W \cdot L$ в виде произведения $V \cdot L$ сводится к тому, что исходные u, v посредством собственного преобразования Лоренца преобразуются в пару ортогональных векторов равной длины, которые затем посредством поворота совмещаются с собственными векторами преобразования Лоренца.

Если ввести ортонормированные вектора e_a, e_b :

$$e_a = \frac{W \cdot L\{v\}}{\sqrt{\sqrt{(u,u)} \cdot (v,v)} \cdot |\sin(\varphi_1 - \varphi_2)|}} ,$$

$$e_b = \frac{W \cdot L\{u\}}{\sqrt{\sqrt{(u,u)} \cdot (v,v)} \cdot |\sin(\varphi_1 - \varphi_2)|}}$$

то, используя (3), (16), (19), можно показать, что k, m представимы в виде:

$$k = \frac{-e_a + \varepsilon_s \cdot e_b}{\sqrt{2}} ,$$

$$m = \frac{e_a + \varepsilon_s \cdot e_b}{\sqrt{2}}$$

где $\varepsilon_s = \frac{|\sin(\varphi_1 - \varphi_2)|}{\sin(\varphi_1 - \varphi_2)}$.

При этом V и L задаются посредством e_a, e_b в виде:

$$V: \begin{cases} V\{e_a\} = e_a \cdot \cos \varphi_0 + e_b \cdot \sin \varphi_0 \\ V\{e_b\} = -e_a \cdot \sin \varphi_0 + e_b \cdot \cos \varphi_0 \end{cases} . \quad (24)$$

$$L\{w\} = e^{-\theta_0} \cdot (w, e_a) \cdot e_a + e^{\theta_0} \cdot (w, e_b) \cdot e_b$$

Безразмерные коэффициенты разложения u, v в базисе собственных векторов e_a, e_b выражаются через длину и ширину фигуры (определённые в п. 3.1.6) в виде:

$$\begin{aligned} \frac{(u, e_a)}{\sqrt{(u, u)}} &= -e^{\frac{\gamma}{2}} \cdot \sqrt{\operatorname{ch} \gamma - \sigma \cdot d} \cdot \sin \varphi_0 \\ \frac{(u, e_b)}{\sqrt{(u, u)}} &= -e^{\frac{\gamma}{2}} \cdot \sqrt{\operatorname{ch} \gamma + \sigma \cdot d} \cdot \cos \varphi_0 \\ \frac{(v, e_a)}{\sqrt{(v, v)}} &= e^{\frac{\gamma}{2}} \cdot \sqrt{\operatorname{ch} \gamma - \sigma \cdot d} \cdot \cos \varphi_0 \\ \frac{(v, e_b)}{\sqrt{(v, v)}} &= -e^{\frac{\gamma}{2}} \cdot \sqrt{\operatorname{ch} \gamma + \sigma \cdot d} \cdot \sin \varphi_0 \end{aligned}$$

Тригонометрические коэффициенты $\sin \varphi_1, \cos \varphi_1, \sin \varphi_2, \cos \varphi_2$ разложения u, v в базисе k, m находятся посредством соответствующих линейных комбинаций выписанных.

Следует обратить внимание, что используемое векторное представление (12), (24) (эквивалентное (17), (18)) от вида интерпретации не зависит. Конкретный выбор активной либо пассивной интерпретации необходим для согласования векторного представления с координатным при описании в явном виде совмещения плоскостей $x \times y$ и $u \times v$.

Альтернативные интерпретации в общем случае описываются следующим.

Пусть A - линейный однородный оператор в произвольном n -мерном вещественном пространстве векторов W с базисом $\{e_j\}$, "+" - операция эрмитового сопряжения.

Тогда активная интерпретация аналитически определяется представлением $A\{w\}$ в следующих двух эквивалентных формах:

$$A\{w\} = \sum_j (w, A^+\{e_j\}) \cdot e_j = \sum_j (w, e_j) \cdot A\{e_j\} \quad .$$

а пассивная интерпретация определяется эквивалентными представлениями w :

$$w = \sum_j (w, A^{+1}\{e_j\}) \cdot A\{e_j\} = \sum_j (w, e_j) \cdot e_j \quad ,$$

в которых скалярные части задают адекватные преобразования координат.

Отождествление плоскости $x \times y$ и двумерного подпространства $u \times v$ сводится к тому, что в подпространстве $u \times v$ n -мерного пространства вводятся точки фигуры W_i :

$$w_i = x_i \cdot e_x + y_i \cdot e_y \quad ,$$

где e_x, e_y - некоторый ортонормированный базис $u \times v$.

Считается, что координатное представление фигуры индуцирует преобразование $W \cdot L$ (12).

e_x, e_y в случае активной интерпретации определяются:

$$\begin{cases} e_x = e_b \\ e_y = e_a \end{cases} .$$

Тогда x_i, y_i согласуются с координатным представлением векторного равенства (12):

$$\begin{cases} x_i = (w_i, V^{-1} \cdot L_1 \{e_x\}) = e^{\theta_0} (x_i \cdot \cos \varphi_0 + y_i \cdot \sin \varphi_0) \\ y_i = (w_i, V^{-1} \cdot L_1 \{e_y\}) = e^{-\theta_0} (-x_i \cdot \sin \varphi_0 + y_i \cdot \cos \varphi_0) \end{cases} .$$

где $L_1 = V \cdot L_1 \cdot V^{-1}$:

$$L_1 \{w_i\} = e^{\theta_0} \cdot (w_i, \cdot V^{-1} \{e_x\}) \cdot V^{-1} \{e_x\} + e^{-\theta_0} \cdot (w_i, \cdot V^{-1} \{e_y\}) \cdot V^{-1} \{e_y\}$$

В случае пассивной интерпретации:

$$\begin{cases} e_x = e_a \\ e_y = e_b \end{cases} .$$

При этом выполняется:

$$\begin{cases} x_i = (w_i, V \cdot L^{-1} \{e_x\}) = e^{\theta_0} (x_i \cdot \cos \varphi_0 + y_i \cdot \sin \varphi_0) \\ y_i = (w_i, V \cdot L^{-1} \{e_y\}) = e^{-\theta_0} (-x_i \cdot \sin \varphi_0 + y_i \cdot \cos \varphi_0) \end{cases} .$$

Таким образом, пассивное и активное представление в координатной записи векторного соотношения (12) формально совпадают, что достигается за счёт выбора для каждой интерпретации своего базиса. Однако, в силу

некоммутативности V и L активная интерпретация координатной записи по сравнению с пассивной описывает собственное преобразование Лоренца, получаемое преобразованием подобия и отличающееся поворотом векторных параметров k, m .

3.1.8 Характеристика фигуры посредством сдвига системы координат.

Идея сопоставления плоских фигур и пар n -мерных ортогональных векторов равной длины даёт возможность дополнить набор признаков, выведенных на основе рассмотрения однородных преобразований. При этом геометрические признаки дополняются параметрами, которые вводятся на основе рассмотрения сдвига начала координат и позволяют улучшить учёт особенностей распределения точек фигуры вблизи центра инерции.

Оказывается, что линейное преобразование векторов u, v в ортогональные равной длины обеспечивается не только однородным преобразованием $W \cdot L$, но также и смещением начала координат исходной плоскости из центра инерции $0, 0$ в некоторую новую точку $\delta x, \delta y$.

Преобразование вторых моментов при сдвиге определяется соотношениями:

$$\begin{aligned}(\tilde{u}, \tilde{u}) &= (u, u) + n \cdot (\delta x)^2 \\(\tilde{v}, \tilde{v}) &= (v, v) + n \cdot (\delta y)^2 \\(\tilde{u}, \tilde{v}) &= (u, v) + n \cdot \delta x \cdot \delta y\end{aligned},$$

где $(\tilde{u}, \tilde{u}), (\tilde{v}, \tilde{v})$ - вторые моменты относительно нового начала

координат, а (\tilde{u}, \tilde{v}) – новый смешанный момент.

Рассматриваемое преобразование, очевидно, инвариантно относительно одновременного изменения знака δx и δy .

Потребуем, чтобы в результате сдвига начала координат вторые моменты по x , по y совпали и одновременно смешанный момент обратился в 0. Тогда, решая систему трёх уравнений относительно δx и δy , получаем, учитывая условие неотрицательности значений $(\delta x)^2$, $(\delta y)^2$:

$$\begin{aligned} \delta x &= \varepsilon_x \cdot \sqrt{\frac{\sqrt{(u, u) \cdot (v, v)}}{n} \cdot (d + \operatorname{sh} \gamma)} \\ \delta y &= \varepsilon_y \cdot \sqrt{\frac{\sqrt{(u, u) \cdot (v, v)}}{n} \cdot (d - \operatorname{sh} \gamma)} \quad , \quad (25) \\ \varepsilon_x \cdot \varepsilon_y &= -\operatorname{sign}\{\cos(\varphi_1 - \varphi_2)\} \end{aligned}$$

где $\varepsilon_x, \varepsilon_y = \pm 1$.

При этом достигается:

$$\begin{aligned} (\tilde{u}, \tilde{v}) &= 0 \\ (\tilde{u}, \tilde{u}) &= (\tilde{v}, \tilde{v}) = \sqrt{(u, u) \cdot (v, v)} \cdot (\operatorname{ch} \gamma + d) \quad . \end{aligned}$$

Из последнего соотношения следует, что псевдодлина, определяемая при сдвиге среднеквадратичным отклонением относительно новых осей координат, совпадает с длиной фигуры, вычисляемой согласно п.3.1.6.

Из (25) и п.3.1.6 следует:

$$(\delta x)^2 + (\delta y)^2 = (\text{длина})^2 - (\text{ширина})^2 = 2 \cdot \frac{\sqrt{(u, u) \cdot (v, v)}}{n} \cdot d$$

Таким образом, величина d с точностью до множителя

совпадает с характерным сдвигом системы координат, равно как и с разностью квадратов размеров фигуры.

Следует обратить особое внимание на то, что сдвиг начала координат индуцируется объединением фигур. Поэтому в перспективе алгебраического развития аналитического описания геометрических свойств рассмотрение сдвига в сочетании с однородными преобразованиями представляется необходимым.

3.2. Структурное описание объектов посредством частичного анализа.

Понятие изотропной фигуры обеспечивает аналитическое описание геометрических свойств объектов, которое имеет адекватную наглядную интерпретацию. В качестве главных метрических характеристик при этом используются длина и ширина фигуры, определяемые как интегральные параметры на основе введения собственной системы координат. Длина, ширина, линейный размер, площадь, ориентация, компоненты сдвига по осям x и y и другие параметры составляют аналитическую систему геометрических признаков, принятую в многоуровневой модели.

Помимо геометрических признаков в аналитическую систему входит также средняя яркость точек фигуры, которая в случае спектрального анализа представляется несколькими компонентами. В зависимости от типа обрабатываемых изображений в систему могут включаться и другие

характеристики. В общем случае под признаками, удовлетворяющим ограничениям многоуровневой модели, понимаются функции от величин, аддитивных относительно слияния областей (объединения фигур). В исследуемой системе набор аддитивных величин составляют: число точек, интегральная яркость, а также первые и вторые моменты, которые вычисляются для каждой из рассматриваемых областей (фигур) и образуют блок данных, описывающих целостные характеристики (абсолютные признаки) объектов.

Необходимая детализация интегрального описания объектов достигается посредством разделения их на части, в результате которого каждый исходный объект представляется в виде набора новых объектов. При этом исходное интегральное описание объекта по каждому признаку замещается рядом компонент.

Можно считать, что при использовании эталонов применяется предельная степень детализации, при которой каждая точка рассматривается как самостоятельный элемент, характеризуемый своим положением в системе центра инерции. Остальные характеристики в этом случае теряют информативность.

По сравнению с описанием целостного объекта возможности признакового описания объектов посредством нетривиального разделения на некоторое число частей возрастают не только за счёт пропорционального увеличения количества вычисляемых значений абсолютных признаков, но также и благодаря формированию новых.

В дополнение к абсолютным признакам оказывается полезным рассматривать относительные, вычисляемые с учётом абсолютных. Простейшими примерами относительных признаков являются процент распределения точек между компонентами объекта, ориентация частей относительно объекта в целом, соотношения линейных размеров объекта и его частей и др.

Существенное расширение признакового описания объектов достигается за счёт рассмотрения комбинаций частей объектов. Если объект разделяется на k частей, то каждому из $2^k - 1$ возможных объединений частей отвечает своё значение признака, возможность вычисления которого обеспечивается аддитивностью характеристик из блока данных. Таким образом, при возрастании объёма данных в k раз достигается почти 2^k -кратное увеличение числа признаков.

Изменение интерпретации представления объекта в виде набора составных частей также расширяет возможности признакового описания.

Исходное представление объекта в виде последовательности новых объектов эквивалентно его описанию посредством набора векторов одинаковой размерности. Каждый вектор отвечает одному из признаков, а каждая компонента вектора описывает соответствующую часть объекта.

При альтернативной интерпретации части, на которые разделяется объект, замещаются точками, совпадающими с центрами инерции и признаки образованной точками фигуры рассматриваются в качестве дополнительных.

Таким образом, рассматриваемое числовое представление объекта индуцируется его разделением на k частей и включает

по каждому аналитическому признаку базовый k -мерный вектор данных, а также производные признаки, описывающие всевозможные объединения частей. При этом базовые и производные элементы объекта описываются как самостоятельные объекты или замещаются точками.

В последнем случае, точки представляют определённые элементы объекта, размеры которых зависят от масштаба изображения. Если расстояния между центрами инерции элементов с изменением масштаба меняются пропорционально, то аналитическое представление объекта данным количеством точек согласуется с реальностью. В противном случае интерпретация аналитических соотношений не точна. Например, при использовании исходной аппаратной дискретизации изображения, значения безразмерных признаков, получаемые посредством нормировки, отвечают представлениям объекта различным числом точек и, вообще говоря, не инвариантны относительно изменения масштаба. Поэтому разделение объекта на части как способ дискретизации изображения представляет самостоятельный интерес.

Структурный анализ объектов посредством разделения на элементы в многоуровневой модели производится независимо от ориентации и обеспечивает формирование признакового описания, устойчивого по отношению к изменению условий съёмки [65-66].

Разделение объекта на части достигается посредством преобразования анализируемых фигур в изотропные и разделения последних на сектора по осям собственной системы координат и

биссектрисам координатных углов (рис.20). При этом исходная фигура разбивается на октет характерных элементов, рассматриваемых как новые фигуры, а также представляется восемью точками, отождествляемыми с центрами инерции элементов.

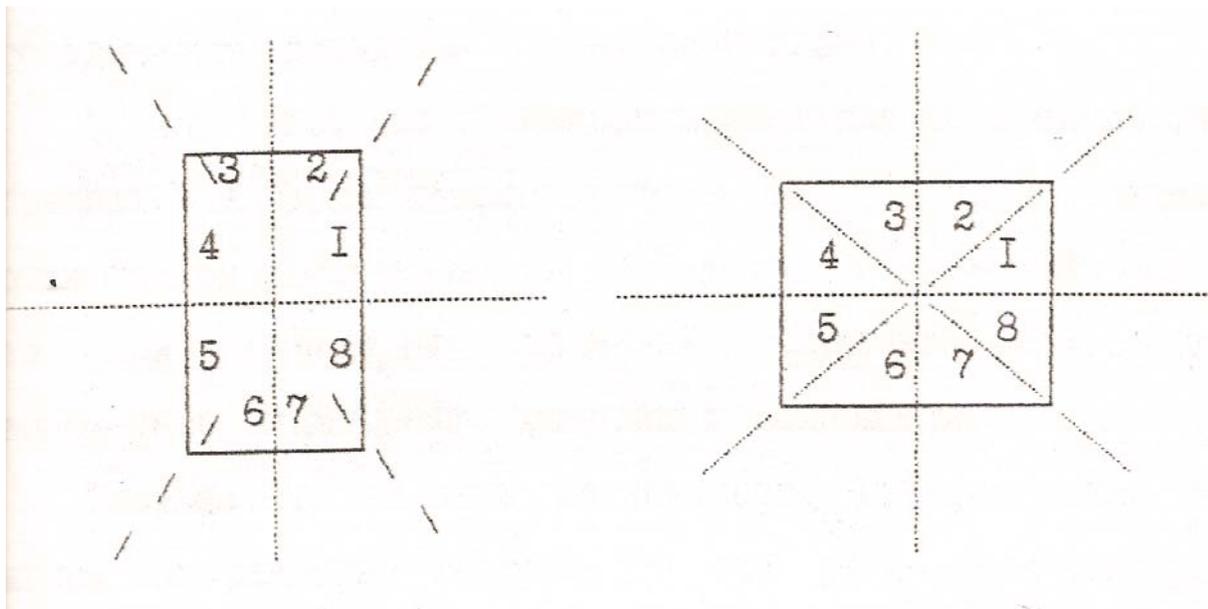


Рис. 20. Разделение фигуры на части посредством преобразования в изотропную

Разделение фигуры на восемь частей порождает алгебраическое расширение признакового описания в 255 раз. Алгебраическое расширение признаков включает базовые признаки элементов октета, на которые разбивается исходная фигура, а также признаки всевозможных объединений базовых элементов, в число которых входят признаки исходной фигуры в целом, признаки восьми половин фигуры, получаемых её разделением на две вдоль каждой из установленных осей и др.

При структурном анализе геометрических свойств симметричные фигуры разделяются на части вдоль осей симметрии и перпендикулярных осей (рис. 20). Так как свойства симметрии инвариантны относительно одновременного

изменения масштаба по осям координат, то способ представления объектов в виде октета элементов обеспечивает устойчивые результаты не только в отношении изменения ориентации, но также и в отношении изменения масштаба, что подтверждается на опыте. Результаты экспериментального исследования приводятся в следующей главе.

В силу того, что смешанный момент для изотропной фигуры обращается в 0, а вторые моменты по x и по y совпадают, точки фигуры распределяются по секторам равновесным образом, что при вычислении значений признаков обеспечивает равномерное подавление случайных искажений.

Разумеется, условие равновесного распределения точек фигуры по секторам сохраняется при разделении изотропной фигуры на элементы в системе координат, повернутой относительно собственной на произвольный угол. При этом поворот собственной системы координат порождает множество разбиений фигуры на октеты фрагментов, что обеспечивает дополнительную возможность многократного параметрического расширения признакового описания фигур, состоящих из достаточного количества точек. В качестве параметра, определяющего выбор разбиения, в этом случае принимается угол поворота координатной системы.

Следует обратить внимание, что принятый способ структурного анализа в общем случае не обеспечивает сохранения связности областей (рис. 21). Новые фигуры, получаемые в результате деления исходной на части, могут содержать более одной связной области (компоненты

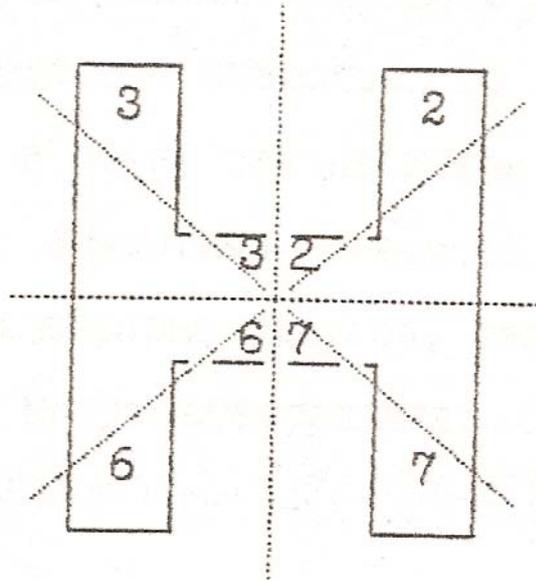


Рис.21. Преобразование компонент связности

связности), что допускается формализмом геометрического описания. При использовании структуры данных, ориентированной на описание связных областей, возможность образования в процессе структурного анализа дополнительных компонент связности следует учитывать особо.

Если разбиение фигуры на октет новых фигур недостаточно для описания геометрических свойств с требуемой точностью, то процесс структурного анализа следует продолжить, рассматривая полученные элементы объектов в качестве исходных. Итеративное повторение структурного анализа элементов изображения индуцирует некоторое многоуровневое представление картины в виде последовательности разбиений исходного множества точек на фигуры, непустое пересечение которых эквивалентно вложению одной в другую. Указанную последовательность разбиений удобно задавать в виде дерева с индексированными связями, описанного в главе 2. При условии корректного расчёта преобразований компонент связности (рис.21) для систематизации и использования результатов поэтапного дробления изобра-

жения остаётся применимой структура данных многоуровневой модели в целом, разработанная в главе 2 для случая аппроксимации объектов посредством итеративного объединения областей.

Таким образом, в общем случае итеративный процесс обработки изображения по яркостным и геометрическим характеристикам производится как в форме синтеза, так и в форме анализа, которые в сочетании могут обеспечивать самостоятельные технические решения.

Аналогично синтезу уровней представления посредством слияния областей дробление (дискретизация) изображения в общем случае производится локально. Исходное изображение на очередном этапе рассматривается в виде набора фигур, по характеристикам которых принимается решение о целесообразности дальнейшей детализации каждой фигуры.

Критерии локального прекращения детализации варьируются в зависимости от контекста задачи. При этом результат дискретизации может определяться:

- условием деления фигуры на связные области;
- максимальным числом компонент различной площади (глава 1, рис.8);
- условием деления фигуры на соразмерные компоненты приблизительно равной площади;
- требованием допустимости значений, при котором дробление фигуры производится до тех пор, пока площадь превосходит пределы значений площадей, характерных для распознаваемых объектов;

- из условия близости дискретной и исходной фигуры по определённому параметру, например, ориентации, где под дискретной имеется в виду фигура, образованная центрами инерции частей;

- из условия семантической однородности частей, при котором дробление производится до тех пор, пока не останутся фигуры, близкие к вырожденным ($\sin(\varphi_1 - \varphi_2) \cong 0$) либо к изотропным ($d \cong 0$), представляющих "отрезки" и "точки" согласно разделам 3.1.3-3.1.4 настоящей главы.

В последнем варианте исходное изображение замещается аналогом с меньшим числом точек.

Предложенное аналитическое описание геометрических свойств позволяет развить стандартный способ дискретизации на основе использования клеточного разбиения. При этом клеточное разбиение каждой фигуры осуществляется в собственной системе координат, а размер элементарной ячейки находится в результате масштабного преобразования, обеспечивающего заданную плотность точек (п. 3.1.6).

Как показывает эксперимент, при разбиении объектов на восемь частей октет центров инерции для простых фигур (символы "Г", "С" и др.) достаточно точно воспроизводит форму фигуры. Для более сложных (символы "Ж", "Ш" и др.) центры инерции образуют овал и самого по себе октета точек оказывается недостаточно. По-видимому, степень отличия распределения центров инерции от эллиптического также может служить критерием прекращения дробления фигуры.

Следует подчеркнуть, что обсуждение возможностей "интеллектуальной" дискретизации изображения нельзя считать исчерпывающим. Развитие этого вопроса связано с решением задачи идентификации и формализацией понятия сложности фигуры, которая представляется одной из перспективных задач.

В качестве меры сложности разбиения изображения на связные области может быть принято суммарное количество элементов, на которые распадаются области в определённом алгоритме дробления.

Вариантом дробления связной области на элементарные является итеративный алгоритм, при котором точки данной области, рассматриваются в собственной системе координат и классифицируются по удалённости от центра инерции, которая сравнивается со среднеквадратичным отклонением. При этом точки разделяются на два класса только при условии, что рассматриваемая связная область распадается более, чем на две компоненты. В противном случае область считается элементарной.

Процесс дробления каждой новой связной области повторяется до тех пор, пока все области не оказываются элементарными.

Приведённый алгоритм очевидным образом обобщается для других способов разбиения области в собственной системе координат на различное количество фрагментов. При заданном способе дробления необходимое уточнение сводится к определению элементарной связной области.

В общем случае область считается элементарной, если количество связных областей, на которые она распадается в собственной системе координат, не превышает числа классов, на которые разделяются точки области.

Совокупность центров инерции элементарных областей, которым приписаны средние значения яркости и других характеристик, представляет собой некоторым образом стилизованное изображение. Замещение изображения наглядным стилизованным аналогом имеет самостоятельное значение для разработки эффективных способов распознавания, особенно в условиях ограниченных ресурсов ЭВМ.

Выводы по главе 3.

На основе рассмотрения однородных линейных преобразований, сохраняющих осевую симметрию плоских дискретных фигур, изучены возможности аналитического описания геометрических свойств объектов независимо от изменения условий наблюдения.

Произвольные плоские дискретные фигуры разделены на вырожденные, метрически изотропные и неизотропные. Построено отображение неизотропных фигур в изотропные, которые совпадают с исходными по площади и имеют не меньшее число осей симметрии. На основе преобразования фигур в изотропные и введения собственной системы координат построено аналитическое описание ориентации и метрических свойств (линейного размера, длины, ширины и др.) объектов независимо от ориентации исходной координатной системы.

Исследованы формальные свойства преобразования фигур в изотропные, рассмотрены характерные примеры. Найдена параметризация аналитических выражений с помощью гиперболических и тригонометрических функций от $\varphi_1 - \varphi_2$ и γ , позволяющая производить теоретические и практические расчёты в упрощённой форме (см. сводку формул в приложении). Дана наглядная интерпретация параметров.

Рассмотрена возможность геометрического описания особенностей распределения точек вблизи центра инерции фигуры на основе сдвига системы координат. Намечены перспективы алгебраического развития формального аппарата геометрического описания посредством рассмотрения объединения фигур.

Разработан метод структурного анализа фигур посредством разделения на фрагменты в собственной системе координат, обеспечивающий описание объектов независимо от ориентации и масштаба, благодаря которому при k - кратном увеличении объёма данных достигается $(2^k - 1)$ -кратное увеличение числа признаков, обеспечивается многократное параметрическое расширение системы признаков за счёт поворота собственной системы координат, а также расширение признакового описания объектов за счёт введения относительных признаков и признаков фигур, образованных центрами инерции фрагментов.

Сформулирована задача дискретизации исходного изображения в зависимости от сложности. Дано определение элементарной связной области, на основе которого предложен принцип решения задачи посредством итеративного дробления связных областей из некоторого разбиения на элементарные компоненты.

ГЛАВА 4. РАСПОЗНАВАНИЕ (ИДЕНТИФИКАЦИЯ) ОБЪЕКТОВ

В главе, исходя из специфики признакового представления видеообъектов, анализируются недостатки стандартных способов классификации и предлагается методологическое решение проблемы идентификации на основе концепции насыщения обучающей информации. Разрабатывается адекватная структура данных, обеспечивающая использование для обучающей информации ограниченного объёма памяти. Обсуждаются особенности организации структуры данных, обеспечивающие при учёте большого количества признаков оптимизацию

идентификации объектов по времени. Рассматриваются алгоритмы обучения и идентификации, при которой наряду с распознаванием предусматривается выделение неопознанных объектов заданной предметной области. Самообучение трактуется как форма автоматизации процесса обучения на основе интерполяции знаний. Описывается минимальный набор независимых функциональных модулей, необходимых для выполнения обучения и автоматического обучения. На примерах распознавания печатных шрифтов и рукописных слов производится верификация механизмов обучения (автоматического обучения), устанавливаются критерии оценки соответствия применяемой признаковой системы и данной предметной области. Предлагается способ алгебраической организации обучающей информации, обеспечивающий снижение требуемого объема памяти.

4.1. Задача автоматизированной идентификации видеообъектов.

При комплексном решении проблемы распознавания особенности организации идентификации видеообъектов определяются условием автоматической подготовки данных, на базе которых производится классификация. В отличие от подготовки данных с участием человека результаты машинного вычисления исходных величин не подвергаются проверке "на правдоподобие" и не гарантируют возможность обеспечения желаемой классификации.

Принципиальное ограничение, препятствующее обеспечению требуемой классификации на заданной выборке данных, выражается в повторяемости значений числовых характеристик различных объектов [39]. Использование признакового представления объектов, связанного со сжатием объема информации, усугубляет эффект повторяемости, что вызывает необходимость применения адекватных способов систематизации данных.

Если при формировании аналитического описания решается задача разработки признаковой системы, в рамках которой достигается максимальное разделение объектов по каждому признаку, то на уровне разработки структуры данных требуемая классификация наблюдаемых объектов обеспечивается возможностью расширения числа используемых признаков. Реальная возможность выбора конкретных систем признаков

достигается за счёт алгебраического расширения, при котором благодаря разбиению объектов на элементы достигается экспоненциальное возрастание количества признаков.

Традиционные способы классификации на основе анализа расстояний между объектами, представленными в виде точек многомерного признакового пространства, в случае автоматизированного распознавания изображений оказываются недостаточно эффективны, особенно при использовании большого числа признаков. Классификация посредством перебора расстояний между точками требует продолжительного времени. Заполнение рабочего объёма признакового пространства идентификаторами объектов позволяет избежать перебора вариантов и оптимизировать процесс распознавания по времени, однако, с ростом числа признаков влечёт экспоненциальное возрастание необходимой оперативной памяти.

При табличном задании объектов в виде сочетаний значений признаков объём накапливаемых при обучении данных возрастает практически неограниченно в силу небольшой вероятности повторения всего набора признаковых значений. Накопление сведений об объектах в табличной форме позволяет воспроизвести распознавание при отсутствии входных изображений, что не относится к числу необходимых условий реализации и представляется излишним, так как приводит к неэффективному использованию оперативной памяти ЭВМ.

Разумно предположить, что повторяемость значений признаков позволяет обойтись при запоминании различных объектов ограниченным объёмом памяти. Целью настоящей главы

является обсуждение способа накопления сведений о видеообъектах [65-66], обеспечивающего идентификацию без использования прямого перебора вариантов при линейной зависимости необходимого объёма памяти ЭВМ от числа используемых признаков и от количества различных идентифицируемых объектов (классов, имён объектов).

При заведомо ограниченном объёме используемой памяти возможность распознавания объектов обеспечивается простым запоминанием признаков. Если, в отличие от табличного описания объектов, значения признаков систематизируются независимо друг от друга, то по мере обучения на примерах они начинают повторяться, что свидетельствует о насыщении обучающей информацией рабочего объёма памяти. Тогда система обеспечивает достоверную идентификацию объектов, которая может быть однозначной или неоднозначной.

Указанная логическая схема идентификации объектов удовлетворяет требованию непротиворечивости и допускает обучение распознаванию произвольных объектов при любой системе признаков и произвольной обучающей информации. Предполагается, что признаковую систему, обеспечивающую неоднозначную идентификацию объектов, можно преобразовать в полную, включив в неё некоторое число дополнительных признаков. Полной считается такая система признаков, которая даёт адекватную и однозначную идентификацию новых объектов с требуемой вероятностью правильного распознавания. Под новыми понимаются объекты, распознавание которых производится впервые [25, 68].

Таким образом, методологическое решение задачи идентификации достигается посредством разработки структуры данных [65–66], обеспечивающей использование ограниченного объема памяти, и сводится к способу построения полной признаковой системы посредством включения достаточного количества дополнительных признаков, выбираемых из числа признаков расширенной системы.

На практике, при организации распознавания на основе концепции ограниченного объема используемой оперативной памяти возникает проблема автоматизации процесса обучения. Обучение представляет собой процедуру заполнения рабочего объема памяти ЭВМ идентификаторами объектов, связанными со значениями признаков. Трудоёмкость обучения, проводимого в интерактивной форме, зависит от предметной области и при большой ёмкости памяти может оказаться достаточно высокой. Принципиальная возможность снижения трудоёмкости заключается в моделировании насыщения рабочего объема памяти посредством интерполяции значений признаков объектов, накопленных на промежуточном этапе обучения.

Для эффективных приложений предложенного технического решения автоматизация обучения имеет решающее значение. С исследовательской точки зрения представляют интерес формы реализации в виде алгоритмов и программ. Следует подчеркнуть, однако, что особенности предлагаемого способа организации идентификации объектов определяются постановкой и принципом решения задачи, которые обеспечивают возможность экспериментальной проверки результатов автоматизации.

4.2. Библиотека объектов.

Структура данных, обеспечивающая идентификацию при условии использования логически ограниченного объёма памяти, реализуется на ЭВМ в виде библиотеки определённым образом упорядоченных имён (названий) объектов, посредством которой по значениям признаков находится обозначение объекта, задаваемое в процессе обучения.

Библиотека объектов [65-66] представляет собой набор обозначений объектов, ассоциированных со значениями признаков, где под обозначениями понимаются некоторые множества имён. При этом обозначения задаются явно в виде последовательностей (списков) имён, а значения признаков отождествляются с адресами обозначений (рис. 22).

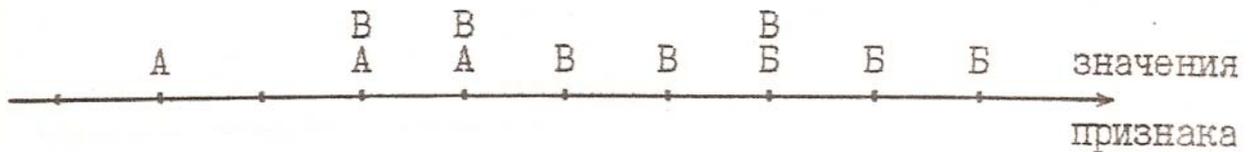


Рис. 22. Организация символьных обозначений объектов по одному из признаков

Составляющие обозначение идентификаторы, рассматриваются как набор предполагаемых объектов, характеризующихся данным значением признака. Так, 'BB' (следует читать: @B или B@) на рис. 22 описывает как объект 'B', так и объект 'B', которые не всегда различаются по значению рассматриваемого признака.

Следует заметить, что в [53] обосновываются более сложные правила идентификации, в которых элементы обозначений рассматриваются как множества. При этом использование обобщённой концепции обозначений [53] представляется необходимым в перспективе развития структуры данных обсуждаемой библиотеки объектов.

Предполагается что для идентификации используются по крайней мере несколько признаков. Все признаки считаются независимыми и рассматриваются симметрично, в равной степени влияя на результат. Различные аналитические признаки и признаки разных элементов объекта в рамках обсуждаемой структуры данных используются одинаково.

Результирующее обозначение объекта определяется как пересечение обозначений, ассоциированных со значениями набора используемых признаков (рис.23).

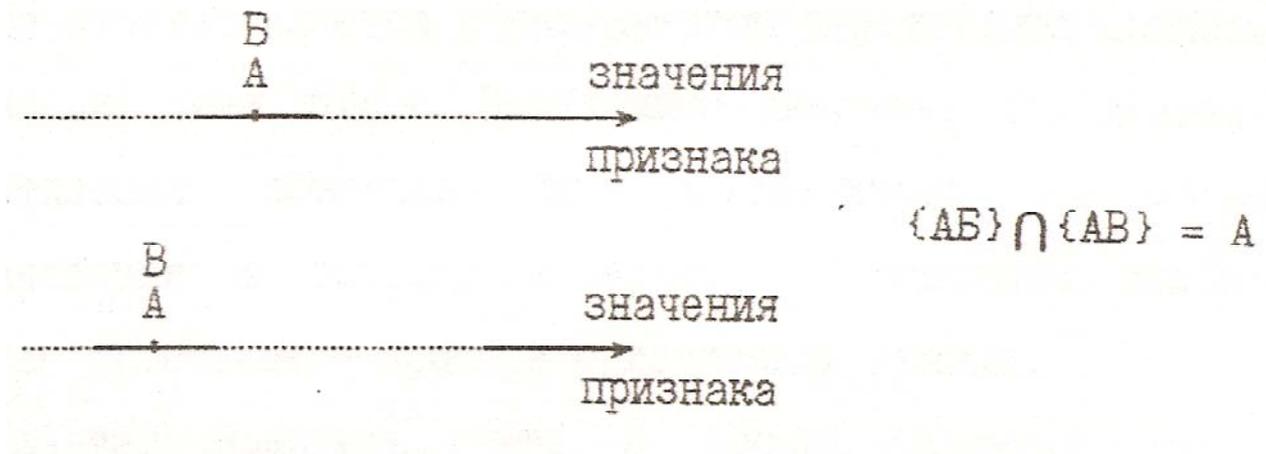


Рис. 23. Идентификация объекта на основе пересечения обозначений, отвечающих различным признакам

Наполнение структуры данных в процессе обучения производится по закону логического сложения и сводится к внесению идентификатора каждого рассматриваемого объекта в

обозначения, ассоциированные со значениями признаков.

Очевидно, величина объема обучающей информации, накапливаемой по одному признаку, не превышает произведения количества значений признака на число идентифицируемых классов объектов.

Ошибочная идентификация объектов в процессе обучения замещается неоднозначной. Однозначность достигается посредством адаптации признаковой системы к заданной предметной области. Адаптация признаковой системы сводится к расширению числа используемых признаков.

В предельном случае, который соответствует использованию эталонов, обучающая информация накапливается в виде символьных обозначений точек объектов. При этом в обозначение каждой точки входят имена всех объектов, у которых данная точка встречается в качестве элемента. Имя объекта отождествляется с результатом пересечения символьных обозначений всех точек. Необходимо указать, что предельная дискретизация объектов не обеспечивает оптимального распознавания и поэтому интересна в основном только в качестве простейшего примера организации данных.

Следует заметить, что в особых случаях описания объектов обсуждаемая структура данных имеет ограниченное применение. Так, координатное описание распределения цифр в латинском квадрате (рис. 24) приводит к объединению объектов в один класс.

1	2	3	4	5	6	7
2	3	4	5	6	7	1
3	4	5	6	7	1	2
4	5	6	7	1	2	3
5	6	7	1	2	3	4
6	7	1	2	3	4	5
7	1	2	3	4	5	6

Рис. 24. Распределение объектов по закону латинского квадрата

Однако, при описании реальных объектов указанный недостаток не имеет решающего значения и компенсируется возможностью использования разнообразных признаков, в том числе индуцируемых алгебраическим расширением признаковой системы. Разумеется, в рамках предложенной структуры данных не исключается применение стандартного способа формирования разделяющихся признаков посредством образования прямых произведений, который в силу недостаточного уровня формализации представляет не теоретический, а преимущественно практический интерес.

Основная функция структуры данных накапливаемой обучающей информации состоит в обеспечении достаточно быстрого нахождения обозначения объекта при большом количестве признаков, что по аналогии с человеческим восприятием позволяет рассматривать библиотеку объектов как модель "быстрой" памяти.

Оптимизация процесса идентификации достигается благодаря тому, что обозначения объектов представляют собой списки некоторым образом упорядоченных идентификаторов, как

на рис. 22-23, где для указанной цели используется алфавитный порядок.

Установленный порядок чередования идентификаторов в обозначениях позволяет определить по первому (начальному) и последнему (конечному) диапазон вероятных имён объекта. Посредством сравнения диапазонов выполняется фильтрация объектов, относимых к рассматриваемой предметной области, что позволяет при обработке большинства имеющихся на изображении объектов избежать непосредственного выполнения операции пересечения множеств.

Алгоритм идентификации данного объекта состоит в следующем:

- по значениям признаков находится верхняя граница начальных и нижняя граница конечных вероятных имён объекта;

- если установленное значение нижней границы превышает значение верхней, то объект считается не относящимся к объектам рассматриваемой предметной области и его идентификация заканчивается;

- в противном случае объект считается объектом рассматриваемой предметной области, предполагаемое имя которого принадлежит промежутку идентификаторов между верхней и нижней границей. Результирующее обозначение объекта находится посредством вычисления пересечения множеств идентификаторов, ассоциированных со значениями признаков. В случае пустого пересечения объект идентифицируется как незнакомый объект данной предметной области.

Как показывает опыт, время, затрачиваемое на идентификацию объектов, по сравнению с продолжительностью всего процесса распознавания незначительно. Каждая из операций выделения и частичного анализа объектов двухградационных изображений выполняется со скоростью порядка 20 килобит/сек. Продолжительность классификации непосредственно с объёмом исходного изображения не связана, а зависит от сложности изображения и пропорциональна общему количеству объектов. Тем не менее, классификация 1000 объектов изображения типа текста при использовании трёх десятков признаков, занимает примерно 40 сек., что в пересчёте на единицу объёма исходной информации даёт в 4-5 раз большую скорость обработки и для данных условий решает проблему оптимизации выполнения на ЭВМ операции пересечения множеств.

При необходимости скорость классификации может быть увеличена посредством усложнения структуры данных, в которой значения признаков помимо обозначений объектов сопоставляются значения перечисленных в обозначениях идентификаторов, отложенные на отрезках формальной оси имён объектов. Использование обозначения объекта как в форме списка, так и в виде последовательности значений, отмеченных на некоторой оси, позволяет оптимизировать выполнение необходимого поиска идентификаторов в заданных множествах, что связано, однако, с возрастанием объёма используемой оперативной памяти и усложнением структуры данных.

Таким образом, в предложенной структуре данных библиотеки объектов в качестве обучающей информации хранятся только множества имён. При этом по сравнению с табличным представлением теряется возможность воспроизведения признакового описания встретившихся ранее объектов без обработки изображения, но достигается эффективное использование памяти. Оптимизация расчётов по времени обеспечивается за счёт учёта диапазонов упорядоченных имён объектов.

Следует отметить возможность алгебраического развития структуры данных [66], применяемой для организации библиотеки объектов, которое строится на основе расширения системы признаков за счёт всевозможных объединений элементов объектов. При этом объединение объектов A , B представляется пересечением обозначений D_A , D_B :

$$A \cup B \Leftrightarrow D_A \cap D_B \quad .$$

Операция пересечения рассматривается на множестве отвечающих объектам сочетаний значений признаков.

С практической точки зрения особого внимания заслуживает вариант, в котором компоненты и результат пересечения обозначений рассматриваются на одном и том же множестве значений признака. При этом в случае аддитивных признаков формируется алгебраическая структура данных, в которой пересечения обозначений компонент совпадают с обозначениями объединений, а адреса обозначений объединений находятся посредством суммирования адресов обозначений

компонент (рис. 25).

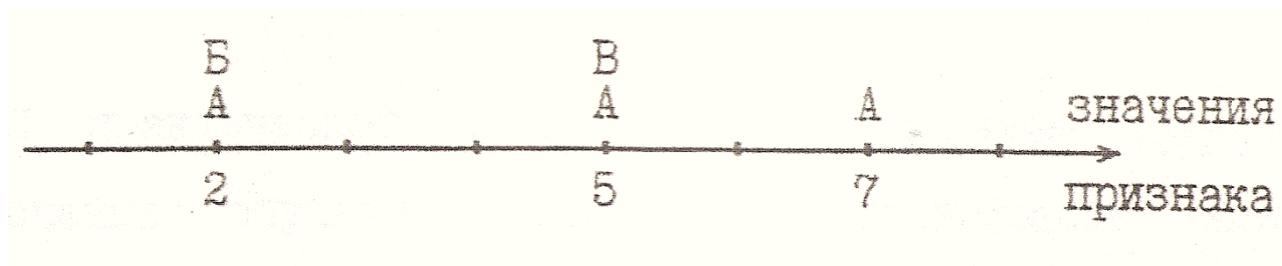


Рис. 25. Алгебраическая организация символьных обозначений в случае аддитивных признаков

Идея алгебраической организации обозначений объектов логически завершает концепцию использования при идентификации ограниченного объёма памяти, так как обеспечивает замещение нескольких признаков одним, но с более высоким уровнем организации данных.

По мере обучения системы при фиксированном наборе признаков число идентификаторов в результирующих обозначениях объектов увеличивается, а с расширением системы используемых признаков - уменьшается. Границы имён, соответствующие встречающимся наборам значений признаков, смещаются. Таким образом, в результате обучения при достаточном количестве признаков обеспечивается возможность формирования библиотеки объектов, адекватной предметной области.

Необходимым признаком успешного формирования библиотеки объектов на практике является устойчивое выделение объектов, принадлежащих требуемой предметной области. Критерием завершения процесса обучения в результате адекватного отражения предметной области библиотекой объектов является однозначная идентификация новых объектов с требуемой вероятностью правильного распознавания.

4.3. Реализация. Эксперимент. Обучение. Самообучение.

С практической точки зрения создание программного обеспечения обучаемой системы распознавания составляет основную, но далеко не полную часть реализации ряда технических решений. Настройка системы на распознавание объектов заданной предметной области посредством ввода в процессе обучения необходимой информации, как любое интерактивное взаимодействие человека с компьютером, представляет собой процесс, эффективность которого прямо связана с простотой выполнения необходимых действий.

Основной смысл использования программной реализации обучаемой системы как на стадии проектирования, так и на стадии внедрения сводится к замене труда квалифицированных постановщиков и программистов трудом операторов. Усилия квалифицированных специалистов сосредотачиваются на модификации центрального звена в виде модуля вычисления признаковых значений. При этом на стадии проектирования за счёт получения практических результатов насыщения обучающей информации обеспечивается экспериментальное исследование систем признаков и разработка алгоритмов автоматического обучения. На стадии внедрения пользователь обеспечивается возможностью выбора признаков в рамках предусмотренной системы, а также снабжается программами, реализующими алгоритмы обучения и самообучения. При оформлении модуля вычисления признаков в открытом виде пользователь получает дополнительную возможность самостоятельного формирования признаковой системы на множестве подготавливаемых данных.

Библиотека объектов реализована на ЭВМ РС/286 в пакете прикладных программ автоматического распознавания полутоновых и двухградационных изображений [66]. Программы выполнены на языке СИ. В простейшей форме обучение вручную обеспечивается указанием объектов и имён объектов, в результате которого идентификаторы объектов заносятся в библиотеку по адресам, отвечающим значениям признаков.

Хотя логическая схема систематизации данных в библиотеке допускает произвольную обучающую информацию, тем не менее, ошибочная идентификация объектов приводит к необходимости использования избыточной признаковой системы и снижает эффективность моделирования характерных для насыщения закономерностей распределения имён. Основную опасность представляют ошибки, связанные с искажениями диапазонов распределений идентификаторов по значениям признаков.

Избежать такого рода ошибок позволяет дополнение структуры данных списками диапазонов значений признаков, сопоставляемых каждому классу идентифицируемых объектов. Тогда при вводе объекта с экстремальными значениями признаков система предупреждает пользователя о возможной ошибке. При этом проверка "на правдоподобие" производится вслед за вводом имени объекта и сводится к условию, что значения признаков данного объекта не выходят за пределы значений признаков объектов того же класса (с тем же именем), встретившихся ранее.

Необходимым средством автоматизации формирования библиотеки объектов в процессе экспериментального исследования, сравнительного анализа и оптимизации систем признаков является реализация в системе распознавания возможности переобучения (дообучения) [66]. Переобучение (дообучение) сводится к генерации для различных систем признаков новых библиотек объектов на основе некоторой исходной. При этом исходная система признаков служит для нахождения имён объектов, которые затем используются в качестве входных данных при формировании (дополнения) библиотеки, отвечающей новой признаковой системе.

Реализация возможности переобучения (дообучения) системы позволяет ограничиться однократным проведением интерактивного обучения в случае многократного использования одних и тех же изображений, в частности при исправлении ошибочной идентификации. Следует обратить внимание, что обсуждаемая структура данных обеспечивает ограниченные возможности "забывать" предъявленные экземпляры объектов. Поэтому для чистоты эксперимента ошибки ввода имеет смысл исправлять посредством переобучения системы с исключением неправильно идентифицированных объектов, что в автоматизированном виде не является трудоёмкой процедурой.

Конкретный набор признаков, используемый для распознавания объектов, их количество заранее не нормируется, а в зависимости от предметной области устанавливается пользователем в процессе работы с системой распознавания.

Вычисление признаков оформлено в виде отдельной

подпрограммы. Остальные программные модули оперируют с формальными обозначениями признаков, что обеспечивает возможность расширения и иной модификации реализованной признаковой системы. При этом на стадии проектирования система распознавания используется в качестве экспериментальной установки для исследования признаков.

Общая признаковая система включает базовые сведения об объектах: среднюю яркость, площадь, координаты центра инерции, первые и вторые моменты относительно центра инерции, вычисленные для компонент каждого объекта, представленного октетом составных элементов (рис.20). Помимо базовых в системе согласно определениям предыдущей главы реализуется вычисление абсолютных и относительных признаков элементов объектов, рассматриваемых в различных сочетаниях, а также вычисление признаков фигур, образованных центрами инерции составных элементов.

Необходимо заметить, что возможность выбора признаков не следует ограничивать условиями типа соображений инвариантности, так как при этом ограничиваются практические возможности формирования признаков систем. Не следует, например, пренебрегать использованием в качестве промежуточных признаков координат центров инерции, обеспечивающих в процессе обучения (переобучения) выделение областей расположения объектов, а также необходимую идентификацию объектов в специальных библиотеках, рассчитанных на обработку экспериментальных данных, которые удобно формировать для решения конкретных прикладных задач. При решении вопроса о включении "неинвариантных" признаков в

число рабочих имеет смысл ориентироваться на возможность их наглядной интерпретации, что при подходящих условиях упрощает переход к автоматизированной обработке.

Для сформированных признаков систем, обеспечивающих распознавание в автоматическом режиме, характерно использование признаков, устойчивых к изменению условий съёмки, в частности к сдвигу, вращению и изменению масштаба изображения. Результаты обучения контролируются по способности технической системы распознавать объекты, предъявленные впервые.

В таблице рис.26 приводятся результаты последовательного обучения программы распознавания по изображениям заглавных букв русского алфавита, снятых с различным разрешением. Общее количество символов – 851.

Разрешение (ТЧК/дюйм) №	0	1	2	3	4	5	6	7
250	842							
255	4	830						
260	4	40	839					
265	6	41	111	841				
270	12	56	132	210	851			
280	1	36	116	202	267	846		
290	4	48	124	209	279	329	840	
297	5	61	148	240	302	372	425	841

Рис. 26. Таблица числа опознанных символов различного масштаба

Использовалось 27 признаков, вычисленных согласно определениям главы 3 :

- яркость;
- вектор распределения точек по секторам, компоненты которого заданы в % от количества точек фигуры;
- ориентация октета центров инерции относительно ориентации фигуры;
- отношение ширины к длине, вычисленное для октета центров инерции;
- восьмимерный вектор ориентации частей относительно фигуры в целом;
- вектор отношений ширины к длине, вычисленных для частей фигуры.

В столбцах таблицы указано количество опознанных объектов (символов) на каждом из 8 шагов обучения. Первое число в каждом столбце описывает предельное количество однозначно идентифицируемых объектов, достигаемое непосредственным обучением на изображении с данным разрешением. Каждое из этих чисел или минимальное из нескольких последовательных, отнесённое к общему количеству объектов, для вероятности правильного распознавания даёт оценку сверху. Остальные числа совпадают с количеством распознанных новых объектов. Вычисление доли распознанных новых объектов позволяет оценить вероятность правильного распознавания снизу.

По мере обучения доля распознанных новых объектов постепенно возрастает. Помимо развития в процессе обучения способности технической системы узнавать объекты, предъявленные впервые, адекватность системы признаков и

объектов данной предметной области проявляется в успешном формировании множества допустимых значений параметров (рис. 2), обеспечивающих устойчивое выделение объектов, а также в отсутствии фактов ложного распознавания.

Характерным признаком насыщения рабочего объёма памяти является тенденция к уменьшению объёма сжатой обучаемой информации, достигаемого при упаковке посредством стандартных программ (ice, pkzip и др.), что связано с ростом числа повторений сочетаний идентификаторов в библиотеке объектов.

Рис. 27 иллюстрирует содержание реальной библиотеки объектов по одному из признаков.

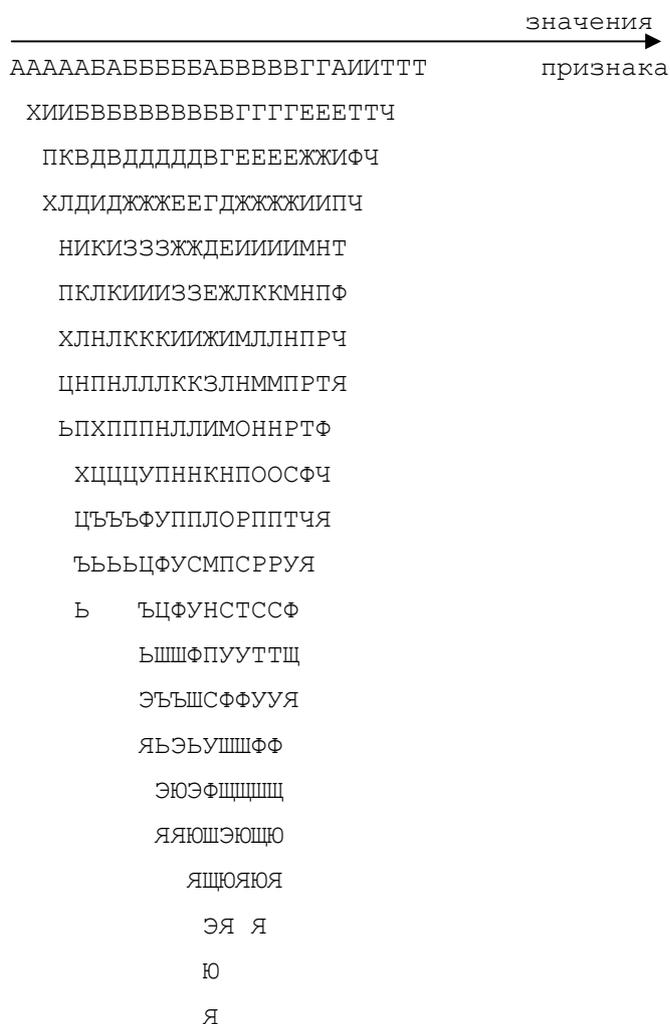


Рис. 27. Обозначения объектов в виде наборов имён, ассоциированных со значениями признака при распознавании символов различного масштаба

Очевидно, большая фактическая ёмкость памяти, отведённой для библиотеки объектов (рис.27) в зависимости от установленной точности измерения признаков, обеспечивает большую возможность различения объектов, которая, однако, реализуется в процессе более длительного обучения. Так, для обеспечения распознавания с вероятностью превышающей 50% в рассматриваемом случае букв различного масштаба потребовалось обучение системы более чем на 6000 (!) примерах объектов. Поэтому обучение вручную имеет смысл проводить главным образом с целью разработки способов моделирования его результатов, обеспечивающих существенное снижение трудоёмкости обучения за счёт автоматизации.

Под автоматическим обучением понимается [66] моделирование насыщения обучающей информации (библиотеки объектов), которое производится без участия человека на основе некоторого начального объёма накопленных данных.

Способы автоматического обучения, реализуемого на основе предложенной структуры данных, имеет смысл разделить на два типа - с использованием и без использования исходных изображений.

Под самообучением понимается [66] автоматическое обучение без использования обработки исходных изображений.

Автоматическое обучение с использованием обработки исходных изображений на основе предложенной модели обеспечивается наборами аппроксимаций объекта на нескольких последовательных уровнях. Стандартным образом оно может организовываться посредством моделирования изменений условий

съёмки, при котором моделируются и отображаются в библиотеке различные искажения объектов. Однако, точное воспроизведение результатов комплексных изменений условий получения изображения представляет собой достаточно сложную задачу [40, 78].

Более просто автоматическое обучение организуется посредством обработки реальных изображений по принципу идентификации на основе критерия однозначности. Указанный принцип состоит в том, что посредством имеющейся в библиотеке информации тем или иным способом строится прогноз названия объекта, который при условии однозначности отождествляется с искомым идентификатором. При этом прогноз может обеспечиваться по неполной системе признаков, строится на основе использования более грубых признаковых шкал или достигаться посредством интерполяции знаний.

Интерполяция знаний представляет собой метод моделирования насыщения библиотеки объектов, который может применяться как с использованием, так и без использования обработки исходных изображений.

На рис.28 приводится карта заполнения библиотеки объектов, в наглядной форме представляющая информацию рис.27, накопленную по одному из признаков.

Интерполяция знаний (рис. 28) сводится к занесению в библиотеку идентификаторов объектов, обеспечивающих заполнение промежутков между имеющимися идентификаторами. Для подавляющего большинства объектов, снятых с различным разрешением, идентификатор каждого объекта заполняет единственный диапазон последовательных значений данного

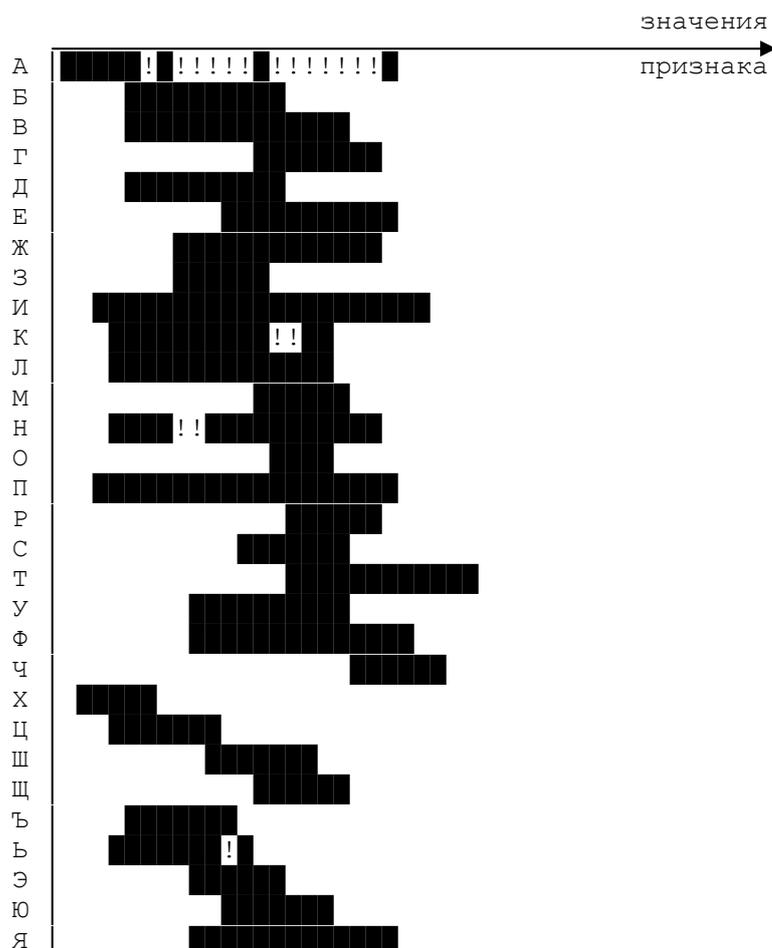


Рис. 28. Карта заполнения памяти. Места интерполяции обозначены "!"

признака. В более сложных случаях распознавания каждому объекту по данному признаку может сопоставляться несколько диапазонов значений признака. Поэтому в общем случае алгоритмическая реализация интерполяции знаний в зависимости от предметной области может быть различной.

Эффективность самообучения определяется точностью воспроизведения законов распределения идентификаторов в библиотеке объектов, складывающихся естественным образом. Использование самообучения в сочетании с дообучением позволяет расширить возможности применения.

Следует обратить внимание, что программная реализация самообучения посредством интерполяции символьных обозначений без дополнительного программирования обеспечивает возможность интерполяции знаний согласно критерию однозначности. Выполнение последней сводится к тому, что посредством алгоритмов самообучения создаётся вспомогательная библиотека объектов, которая затем используется для дополнения исходной в режиме дообучения.

Таким образом, интерполяция знаний согласно критерию однозначности на основе экспериментальных данных обеспечивает фильтрацию обучающей информации, заносимой в библиотеку в процессе самообучения, что позволяет снизить требования к разработке для самообучения алгоритмической базы.

Помимо экспериментов с буквами, снятыми в различном масштабе, ставились также эксперименты по распознаванию различных печатных шрифтов. При ограниченном интерактивном обучении оптимальные результаты достигались в случае распознавания букв простой формы (типа @Г@, @X@, @С@ и др.), форма которых достаточно хорошо передаётся фигурой, образованной октетом центров инерции фрагментов. Для обеспечения распознавания букв сложной формы требовалось существенно более длительное обучение, что позволяет сделать вывод о необходимости стилизации объектов при промышленной реализации алгоритмов. При этом стилизация (дискретизация) изображения может осуществляться по принципу деления связных областей на элементарные, описанному в предыдущей

главе, либо выполняться на основе одного из многочисленных известных эвристических алгоритмов.

Наиболее убедительные экспериментальные результаты проверки механизмов обучения были получены [66] в случае распознавания рукописных слов (рис. 29), каждое из которых рассматривалось как единый символ. При этом исходная обучающая выборка включала около 100 экземпляров каждого слова. Использовался набор из 27 признаков, перечисленных выше. Результаты интерактивного обучения тестировались посредством предъявления нескольких сотен новых экземпляров слов, написанных от руки, а также путём предъявления множества объектов иного типа (печатных символов).

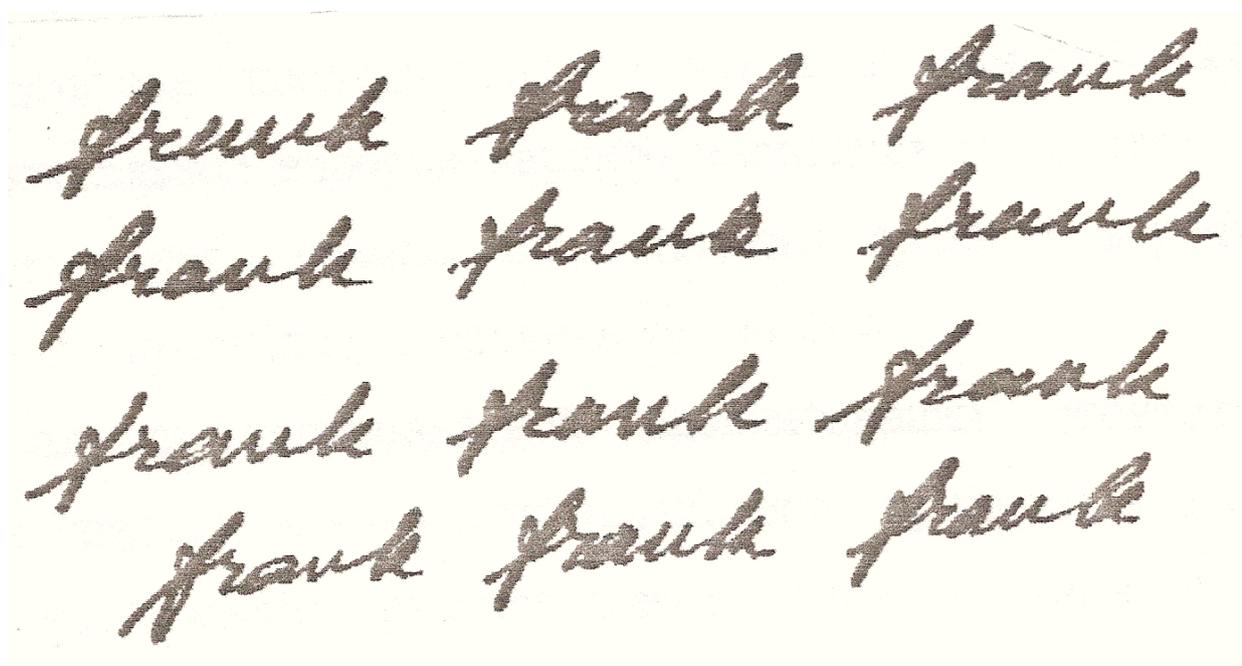


Рис. 29. Пример изображения, изготовленного для тестирования результатов обучения (автоматического обучения)

Формирование библиотеки объектов продолжалось в режиме автоматического обучения по неполной системе значений признаков с использованием критерия однозначности.

Автоматическое обучение по неполной системе признаков в данном случае производилось в простейшем варианте без предварительного выполнения процедуры интерполяции, что сводилось к следующему:

- посредством переобучения создавалась промежуточная библиотека, обеспечивающая распознавание единственного слова;

- в процессе обработки изображений, вводимых со сканера, промежуточная библиотека дополнялась обозначениями данного слова, ассоциированными со значениями признаков однозначно идентифицированных объектов;

- посредством повторной обработки используемых изображений обозначения из промежуточной библиотеки переносились в исходную в режиме дообучения.

Результаты исследования обучения в рамках описанной экспериментальной выборки характеризовались тем, что:

- доля распознанных слов (нижняя граница вероятности правильного прогноза) составила около 70% ;

- случаи неоднозначной идентификации отсутствовали (верхняя граница вероятности правильного прогноза - 100%);

- при попытке распознавания объектов типа печатных шрифтов не наблюдалось случаев ложной идентификации.

Следует подчеркнуть, что для промышленной реализации чтения рукописного текста представляется необходимым дополнить имеющуюся программную систему помимо программ дискретизации также программами выделения отдельных букв. В контексте развития предложенного в настоящей работе подхода к выделению видеообъектов охватить решение задач такого типа и других позволяет включение в структуру данных атрибутивного сим-

вольного представления объектов в виде множеств идентификаторов, пересечение которых сопоставляется объединению объектов. Дополнение признакового описания объектов символьными характеристиками связано с необходимостью обеспечения корректировки выделения объектов, осуществляемой посредством обучения.

Корректировка выделения объектов за счёт использования в качестве атрибутов составных обозначений в виде списков формальных символов предполагает организацию замкнутого цикла обработки: выделение объектов ---> распознавание ---> выделение объектов и достигается за счёт объединения распознанных смежных элементов объектов, пересечение обозначений которых не пусто. При этом алгебраическая организация данных (рис.25) приобретает точный физический смысл.

Дополнение структуры данных символьными атрибутами, преобразуемыми посредством логических операций, обеспечивает дальнейшее обобщение представления видеоданных. При этом к операциям со списками сводится описание соседства областей (п. 2.3.3 главы 2), а также анализ по локальным экстремумам (п. 2.6 главы 2), что представляет самостоятельный интерес.

В заключение следует отметить, что рамки промышленной разработки технических систем не допускают ошибочного проектирования организации данных программного комплекса. Поэтому научные результаты диссертационной работы, позволяющие на основе экспериментального исследования и теоретического обобщения разработать и обосновать комплексный вариант динамических структур данных обучаемой системы автоматического распознавания, представляют необходимый этап её эффективной реализации.

Выводы по главе 4.

На основе анализа особенностей распознавания изображений сформулированы принципы адекватной организации обучающей информации согласно концепции насыщения используемого объема оперативной памяти, характеризующиеся:

- симметричным рассмотрением всех признаков;
- непротиворечивостью логической схемы идентификации при произвольных обучающих данных;
- ограничением рабочего объема оперативной памяти линейной зависимостью от числа признаков, разброса признаковых значений и количества идентифицируемых видеообъектов;
- исключением необходимости проведения идентификации посредством перебора вариантов.

В качестве реализации перечисленных принципов предложена структура данных обучающей информации (библиотеки объектов), элементами которой являются составные обозначения в виде списков упорядоченных идентификаторов, ассоциированных со значениями признаков. При этом объединению объектов (элементов объектов) сопоставляется пересечение обозначений, а результирующее обозначение

объекта определяется совокупностью пересечений обозначений, отвечающих различным признакам всех составных частей.

На экспериментальном материале изучены особенности распределения идентификаторов в библиотеке объектов действующей системы автоматического распознавания. Для общего случая идентификации видеообъектов предложено методологическое решение проблемы обучения с требуемой вероятностью правильного прогноза.

Разработаны приёмы автоматического обучения системы на основе интерполяции знаний, состоящей в формировании для каждого объекта ограниченного числа диапазонов непрерывного распределения идентификатора по значениям признака.

Предложен способ автоматического формирования необходимого числа диапазонов посредством использования при интерполяции знаний критерия однозначности.

Предложена алгебраическая организация библиотеки объектов и предусмотрено обобщение структур данных для выделения-идентификации объектов за счёт дополнения атрибутивного представления объектов составными символьными обозначениями.

ЗАКЛЮЧЕНИЕ

Основным результатом диссертационной работы является создание прототипа обучаемой системы автоматического распознавания изображений, в которой управление выделением объектов желаемого типа устанавливается за счёт формирования в процессе обучения диапазонов разброса значений признаков и определяется пересечением указанных диапазонов. Виды (классы) идентифицируемых объектов некоторой предметной области определяются наборами поддиапазонов непрерывного распределения в библиотеке идентификаторов классов объектов.

В процессе создания прототипа разработаны и экспериментально исследованы динамические структуры данных для выделения и для идентификации объектов. Первая реализует адаптивное представление изображения в виде вычисляемой иерархической системы связанных областей. Вторая включает признаковое структурное (частичное) представление объектов в виде базовых векторных и скалярных значений признаков, а также библиотеку идентификаторов, ассоциированных со значениями некоторого подмножества признаков расширенной признаковой системы. При этом общая логическая схема организации видеоданных включает аппроксимацию объектов (элементов объектов) наборами связанных областей, а также интерпретацию слияния связанных областей в терминах эквивалентных преобразований используемых формальных структур.

В рамках предложенных структур данных рассмотрены и реализованы восходящая и нисходящая стратегия аппроксимации

объектов (элементов объектов) посредством связанных областей. Первая осуществляется в итеративном процессе слияния некоторых элементарных областей исходного изображения. Вторая реализуется в виде разделения областей на фрагменты. Обе стратегии развиты в предположении использования как яркостных, так и геометрических признаков и ориентированы на описание свойств видеообъектов независимо от изменений условий съёмки.

Особое место в работе занимает аналитическое геометрическое описание объектов в виде изотропных фигур, которое является основой классификации объектов по интегральным метрическим признакам, служит для обобщения методов формирования и выделения (поиска) объектов, обеспечивает многократное расширение признаковой системы, используется при дискретизации изображения и определяет возможности реализации центрального звена вычисления признаков системы автоматического распознавания.

Опыт исследования системы признаков посредством прототипа системы автоматического распознавания, реализованной на основе предложенных способов организации видеоданных, обобщён в новой математической модели распознавания изображений, включающей логическую схему распознавания, аналитическое описание объектов, интерпретацию используемых формальных структур и элементарных операций со структурами, а также приёмы и методы применения. При этом аппарат модели не выходит за рамки стандартной математики. В целом новизна состоит в оригинальном сопоставлении (интерпретации) реальных и формальных понятий.

На основе построенной математической модели в работе достигается комплексное решение актуальной научной проблемы автоматического распознавания изображений на трёх последовательных стадиях – выделения, структурного анализа и идентификации видеообъектов.

Основные научные результаты сводятся к следующему.

1) Предложен метод многоуровневой сегментации (разбиения) полутоновых изображений на связные области посредством итеративного слияния соседствующих областей близкой яркости, в котором все области рассматриваются равноправно.

2) На основе обобщения метода сегментации построена многоуровневая модель представления изображения, обеспечивающая аппроксимацию наблюдаемых объектов без искажения формы минимальным числом частей в виде связных областей со значениями признаков в установленных диапазонах.

3) Для оптимизации выделения простых и составных объектов по памяти и по времени разработана динамическая структура данных, в которой многоуровневая сегментация изображения задаётся в виде дерева с индексированными связями, а слияние областей описывается эквивалентными преобразованиями деревьев, таблицы связности и функций от аддитивных характеристик.

4) 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 без использования управляющих параметров.

5) 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.

6) 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.

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

8) Для идентификации объектов по повторяющимся признакам предложен способ организации данных в виде ассоциированных с признаковыми значениями списков идентификаторов, предусматривающий алгебраическое представление обучающей информации и автоматизацию обучения согласно критерию однозначности.

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

Полученные в работе научные и практические результаты, разработанные алгоритмы и программные системы внедрены и используются в 2 организациях и являются составной частью ряда текущих тематических работ, проводимых в СПИИ РАН.

ЛИТЕРАТУРА

1. Александров В.В. Самоподобные рекурсивные структуры как способ представления знаний в ЭВМ.- В кн.: Информационно-вычислительные проблемы автоматизации научных исследований. -М.:Наука, 1983, -С.65-67.
2. Александров В.В., Арсентьева А.В. Информация и развивающиеся структуры. -Л.: ЛНИВЦ АН СССР, 1984.
3. Александров В.В., Горский Н.Д. Представление изображений. Рекурсивный подход. -Л.: Наука, 1985. - 190 с.
4. Александров В.В., Горский Н.Д. Базы видеоданных: проблемы и перспективы. -Л.:ЛНИВЦ АН СССР, 1985. - 73 с.
5. Александров В.В., Горский Н.Д. ЭВМ видит мир. -Л.: Машиностроение, 1990. - 140 с.

6. Александров В.В., Горский Н.Д., Поляков А.О. Структурный метод упорядочивания и классификации экспериментальных данных. -Л.: Изд. ФТИ им. А.Ф. Иоффе АН СССР, 1978.
7. Александров В.В., Горский Н.Д., Мысько С.Н. Экспертные системы анализа изображений. -Л.:ЛИИАН, 1987. - 36 с.
8. Александров В.В., Поляков А.О. Структурный метод классификации // Вопросы классификации. Адаптация в системах со сложной структурой.-М.: Изд. Научного совета по проблеме "Кибернетика" АН СССР, 1977.
9. Александров В.В., Харинов М.В. Представление изображений ранговыми распределениями.-Препринт № 61-Л.:ЛИИАН,1988.-48 с.
10. Александров В.В., Харинов М.В. Представление полутонового изображения с помощью нерегулярного дерева однородных по яркости областей // Искусственный интеллект 90 (ВКИИ - 90) / Сб.докл. 3 Всесоюз. конф. - Минск, 1990.- т.1, - С.187-190.

11. Анисимов В.А., Горский Н.Д. Поиск произвольно ориентированного объекта на изображении на основе пирамидально-рекурсивных структур данных // Информационные проблемы распознавания. - Л.: ЛИИАН, 1988. - С.84-94.
12. Афанасьев А.Н., Гелис С.И. и др. Автоматизированный экспресс-анализ изображений при геологической разведке дна океана // Автоматизированные системы обработки изображений (АСОИЗ-Войчук докл. 2-й Всесоюз. конф. - Львов, - 1986. - С.181-182.
13. Бакут П.А., Колмогоров Г.С., Ворневицкий И.Э. Сегментация изображений: методы пороговой обработки // Зарубежная электроника, 1987. - № 10. - С.6-24.
14. Бакут П.А., Колмогоров Е.С. Сегментация изображений: методы выделения границ областей // Зарубежная электроника, 1987. - № 10.- С.25-47.
15. Васильев В.И., Ивахненко Г.А. Синтез пространств оптимальной кластеризации // Распознавание образов и анализ изображений (РОАИ-1-91) / Тез.докл.1 Всесоюз. конф. - Минск, 1991.-ч.2.-С.7-9.
16. Васильев В.И. Распознающие системы. - Киев: Наука, 1983. - 421 с.
17. Войлов Ю.Г., Энтелис Е.М. Решение задачи однозначности ориентации двумерных объектов в задаче распознавания изображений // Распознавание образов и анализ изображений (РОАИ-1-91) / Тез.докл. 1 Всесоюз. конф. -Минск, 1991. - С.37-40.
18. Гиммельфарб Г.М., Головцов А.Л.и др. Пакет прикладных программ КРОКИС-СМ (версия 2.1) для диалогового анализа межзональных цифровых изображений.// Программные продукты и системы. 1988. - № 1. - С.77-84.

19. Гимельфарб Г.Л., Залесный А.В. Модели марковских случайных полей в задачах генерации и сегментации текстурных изображений // Сб. науч.тр.: Средства интеллектуализации кибернетических систем./ Ин-т кибернетики им. В.М.Глушкова АН УССР .-Киев. 1989. - 115с., - С.27-35.
20. Глезер В.Д. Механизмы опознавания зрительных образов // Вестник АН СССР. М.:Наука, 1970.- № 7. - С.30-37.
21. Гренандер У. Лекции по теории распознавания образов: Пер. с англ. - М.:Мир, 1979-1983. В 3-х томах.
22. Денисов Д.А., Дудкин А.К., Меденников П.А. Алгоритм вычисления геометрических характеристик областей изображений, представленных квадродеревьями // Известия ЛЭТИ:Сб.науч.тр./ Л.:ЛЭТИ, 1986. - Вып.37. - С.70-77.
23. Денисов В.М., Матвеев Ю.Н., Очин Е.Ф. Принципы организации систем обработки изображений на базе клеточной логики // Зарубежная электроника. 1984. - № 1. С.3-25.
24. Дуда Р., Харт П. Распознавание образов и анализ сцен. - М.: Мир, 1976. - 511 с.
25. Ефимов Ю.Н. Метод оптимальных признаков и организация распознающей системы для многоальтернативного распознавания // Распознавание образов и анализ изображений: новые информационные технологии (РОАИ-1-91) / Тез. докл. 1 Всесоюз. конф. - Минск. 1991. - ч.2. - С.82-83.
26. Журавлёв Ю.И. Непараметрические задачи распознавания образов // Кибернетика. - 1976. - № 6 - С.93-103.
27. Журавлёв Ю.И. Об алгебраическом подходе к решению задач распознавания и классификации // Проблемы кибернетики. - М.:Наука, 1978. - Вып.33. - С.5-68.

28. Журавлёв Ю.И. Корректные алгебры над множествами некорректных (эвристических) алгоритмов. 111 // Кибернетика. 1978. - № 2 - С.35-43.
29. Журавлёв Ю.И. Экстремальные задачи, возникающие при обосновании эвристических процедур // Проблемы прикладной математики и механики.-М.:Наука, 1971. - С.67-75.
30. Журавлёв Ю.И., Никифоров В.В. Алгоритмы распознавания, основанные на вычислении оценок // Кибернетика. 1971. - № 3 - С.1-11
31. Зубарев Ю.Б., Глариозов Г.Л. Передача изображений. - М.:Радио и связь, 1989.
32. Исаев У.Т., Акаев А.А., Кутанов А.А. Распознавание образов с использованием голографических фильтров на фототермопластическом носителе // Распознавание образов и анализ изображений (РОАИ-1-91) / Тез. докл. 1 Всесоюз. конф. - Минск, 1991,- ч.3.- С.58-61.
33. Искандеров П.М., Слиозберг Н.С., Харинов М.В. Итеративно-рекурсивный метод предварительной обработки полутоновых изображений в задаче выделения конкреций. // Информационные проблемы распознавания. - Л.:ЛИИАН, 1988. - С.107-117.
34. Искандеров П.М., Харинов М.В. Сегментация полутоновых изображений методом итеративного осреднения. // Проблемы создания систем обработки, анализа и распознавания изображений. - Ташкент, 1989. - С.25-26.
35. Искусственный интеллект : Кн.2. Модели и методы: Справочник / Под ред. Д.А.Поспелова. - М.: Радио и связь. 1990. - 304 с.
36. Кантишев В.Н., Устройство для селекции изображений. А.с. СССР № 1638718, кл. G 06 K 9/00, Б.И. № 12, 1991.

37. Корчинский В.М. Групповые интегральные признаки геометрических классов проекционных изображений в задачах их инвариантного распознавания / Распознавание образов и анализ изображений: новые информационные технологии (РОАИ-1-91) // Тез.докл. Первой Всесоюз. конф. -Минск:изд. НСК АН СССР, 1991. - ч.2. - С.90.
- 38 Крамер Г. Математические методы статистики, - М.:Мир, 1975. - С.305.
39. Кузьмин В.А. Метод преобразования информации в задаче классификации объектов трёх классов / Распознавание образов и анализ изображений: новые информационные технологии (РОАИ-1-91) / Тез.докл. 1 Всесоюзной конференции. -ч.2. - Минск. 1991. - С.95.
40. Ледюков М.А. Автоматизированная система предварительной обработки и определения параметров полутонного изображения "Контур" // Структуры данных, алгоритмы и программы обработки сложных изображений. Препринт № 162 - С.-П.: СПИИРАН, 1992. - С.20-33.
41. Марр Л. Зрение. Информационный подход к изучению представления и обработки зрительных образов: Пер. с англ. М.:Радио и связь, 1987. - 400 с.
42. Митчелл О.Р., Делп Э.Дж. Усечённое блочное кодирование многоуровневой графической информации. // ТИИЭР, 1980. - т.68. -№ 7. С.129-137.
43. Мясников В.А., Игнатъев М.Б., Торгашев В.А. Рекурсивные вычислительные машины. Препринт ИТМ и ВТ № 12, М.,1977.-36с.
44. Нечёткие множества и теория возможностей / Под ред. Р.Ягера. - М.: Радио и связь, 1986.

45. Павлидис Т. Алгоритмы машинной графики и обработки изображений. - М.: Радио и связь, 1986. - 399 с.
46. Паук В.Н. Автоматическая классификация объектов на изображении. // Автоматизированные системы обработки изображений (АСОИЗ -89) / Тез. докл. 3 Всесоюз. конф., - Л., 1989. с. 75.
47. Подвигин Н.Ф., Макаров Ф.Н., Шелепин Ю.Е. Элементы структурно- функциональной организации зрительно - глазодвигательной системы. - Л.:Наука, 1986. - 251 с.
48. Прэтт У.К. Цифровая обработка изображений. М.:Мир,1982. - 790с.
49. Русын Б.П. Структурно-лингвистические методы распознавания изображений в реальном времени. - Киев.:Наукова думка, 1986. - 192с .
50. Садыков С.С., Кан В.Н., Самандаров И.Р. Методы выделения структурных признаков изображений. Ташкент:Фан, 1990.-103 с.
51. Тейлор Э.Ф., Уилер Дж.А. Физика пространства-времени, - М.:Мир, 1969. - С.76.
52. Трифанов В.Н. Методологические основы технологии управления научно-техническими программами // Автоматизация региональных технологий, - Л.: Наука. Ленинградское отделение, 1989. - С.32-47.
53. Трифанов В.Н., Концевой С.А. и др. Автоматизированный контроль научно- технических программ // Автоматизация региональных технологий, - Л.: Наука. Ленинградское отделение, 1989. - С.32-47.
54. Трифанов В.Н. Методические основы синтеза динамических сетей: алгебраическое равновесие и статистика. Препринт № 148, - Л.:ЛИИАН, 1991. - 61 с.

55. Трифанов В.Н. Концептуальные основы анализа диагностики и экспертизы качества промышленной продукции. Препринт № 154, -Л.:ЛИИАИ, 1991. - 64 с.
56. Ту Дж., Гонсалес Р. Принципы распознавания образов: Пер. с англ. - М.: Мир, 1978. - 411с.
57. Фор А. Восприятие и распознавание образов: Пер. с фр. - - М.: Машиностроение, 1989. - 272 с.
58. Фу К.С. Структурные методы в распознавании образов/ Пер. с англ. под ред. М.А. Айзермана - М.:Мир, 1977. - 320 с.
59. Харинов М.В. Автоклассификация объектов двухградационных изображений на примере вычисления формы // Информационные системы распознавания. - Л.: ЛИИАИ, 1988. С.94-107.
60. Харинов М.В. Быстрая сегментация полутонового изображения методом итеративного осреднения. // Автоматизированные системы обработки изображений (АСОИЗ-89), Тез. докл. 3 Всесоюз. конф. - Л., 1989. - С.85.
61. Харинов М.В. Представление растрового полутонового изображения с помощью однородных по яркости областей. // Представление знаний и экспертные системы. - Л.:ЛИИАИ, 1989., - С.142-154.
62. Харинов М.В. Сегментация полутонового изображения на основе бескоординатного представления // Проблемы создания систем обработки, анализа и распознавания изображений. - - Ташкент: ФАН, 1990. - С.31-42.
63. Харинов М.В. Выделение объектов на полутоновых изображениях // Оптическое изображение и регистрирующие среды / Тез. докл. 2 Всесоюз. конф., - Л. 1990., - т.2, - С.171-172.

64. Харинов М.В. Выявление объектов на основе многоуровневого представления полутонного изображения // Проблемы создания систем обработки, анализа и распознавания изображений. - Ташкент:ФАН, 1990. - С.56.
65. 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.
66. Харинов М.В. Разработка динамических структур данных системы автоматизированного распознавания изображений // Информатизация России / под руков. В.М.Пономарёва / Отчёт по НТП (проект 1.4.1)- С.-П.:СПИИРАН, 1992. - 186с.
67. Юсупов Р.М., Козлов Ю.М. Обучение в экспертных системах // Адаптивные и экспертные системы в управлении / Тез.докл. 5 Ленингр. симпоз. по теории адаптивных систем. - Л.:Знание, 1991. - С. 112-113.
68. Якубович В.А. Некоторые общие теоретические принципы построения обучаемых опознающих систем. 1. // Вычислительная техника и вопросы программирования. - Л.:ЛГУ, 1965.- С.3-72.
69. Янковская А.Е. Преобразование пространства признаков в пространство образов на базе логико-комбинаторных методов и свойств некоторых геометрических фигур // Распознавание образов и анализ изображений: новые информационные технологии (РОАИ-1-91)/Тез.докл.1 Всесоюз. конф.-Минск:изд.НСК АН СССР, 1991. - ч.2. - С.178.
70. Ярославский Л.П. Введение в цифровую обработку изображений. - М.: Сов. радио, 1979.

71. Adelson E.H., Burt P.J. Image Data Compression with the Laplassian Pyramid. //Proc.,IEEE Comp. Soc. Conf. Pattern Recogn. Image Proc., Dallas, 1981, - P.218-223.
72. Alexandrov V.V., Gorsky N.D., Mysko S.N. Recursive Pyramids and Their Use for Image Coding // Pattern Recogn. Lett., 1984, - v.2, - № 5, - P.301-310.
73. Alexandrov V.V., Gorsky N.D. Computer and Human Vision Systems // Machine GRAPHICS and VISION / Proceedings of 2nd GKPO'92 Conference, Naleczow, May 18-23, 1992. -v.1. - № 1/2. - P. 5-18
74. Alexandrov V.V., Gorsky N.D. From Human To Computers: Cognition throw Visual Perception, World Scientific, Singapore, 1991. - 203 p.
75. Alexandrov V.V., Gorsky N.D. Can a Computer Vision System Work Like a Human One // International Journal of Imaging Systems and Technologies, v. 3, 1991. - P.269-277.
76. Arsonault H.H., Kchallasinska-Moscuchov Y.N., Young Y. Rotation invariant pattern recognition // SPIE Application of digital image processing, 1982. - v.359.
77. Artificial intelligence: implication for CIM / Ed. by A.Kusiak. - Berlin a.o.] Springer, Kempston: IFS (publ), 1988. - № 14, - 527p.
78. Beckers A.L.D., Smeulders A.W.M. Optimization of Length Measurments for Isotopic Distance Transformation in Three Dimensions // Image Understanding -v.55-№ 3 1992.-P.296-305.
79. Cappelini V., Del Bimbo A., Mecocci A. New complete processing system for recognition and classification in multiple image frames // Electronics letters, 1984. - v.20, - № 11, - P.474-475.

80. Casasent D., Xia S.-F., Lee A.J. and J.S. Song, Real-Time Deformation Invariant Optical Pattern Recognition Using Coordinate Transformations // Appl. Opt., 1987. -v.26, - P.938.
81. Chien C.H., Aggarwal J.K. A normalised quadtree representation // Proc. CVPR'83: IEEE Computer Society Conf. on computer vision and pattern recognition, Washington, 1983. - P.121-126.
82. Chou S.L., Liu J.C., and Tsai W.H., Fold principal axis - A new tool for defining the orientation of rotationally symmetric shapes // Pat. Rec. Lett. 1991. -v.12,-P. 109-115.
83. Enomoto H. and Shibata. Orthogonal Transform Coding System for Television Signals // IEEE Trans. Electromagnetic Compatibility, EMC-13, 1971. - v.3, - P.11-17.
84. Faugeras O.D. Decomposition and Decentralization Techniques in Relaxation Labeling // Comput. Graphics and Image Processing. 1981. - v.16, - № 2. - P.341-355.
85. Hu M.-K., Visual pattern recognition by moment invariants // Trans.Inform.Theory, 1962. - P.179-187.
86. Huang K.S., et al. Binary image algebra and digital optical cellular image processors.//Appl.Opt.,1988. - № 1. - P.20-23.
87. Klinger A.,Dyer C.R. Experiments on Picture Representation Using Regular Decompositions // Comput. Graphics Image Process., 1976, -v.5, -№ 2, P.305-321.
88. Kugel P. Induction, pure and simple // Information and Control. -1977. - № 35.- P.276-337.
89. Li B.C. and Shen J., Fast computation of moment invariants // Patt. Rec., 1991. - v. 24.- P.801-806.

90. Maurike A., et. al. Registration and rectification of images using image algebra.//IEEE (PR and AI), 1987.-№ 6, - P.16-19.
91. Rosenfeld A., Hummel R.A., Zucker S.W. Scene Labeling by Relaxation operations // IEEE Trans.Systems Man Cybernet. 1976. - SMS-6,- № 6.- P.420-433.
92. Schwartz T.J. Fuzzy Systems Come to Life in Japan // IEEE, Expert. Feb. 1990. - P.77-78.
93. Tanimoto S.L. Template Matching in Pyramids // Comp. Graph. Image Proc., 1981. -v. 16, -№ 2, - P.356-369.
94. Thomas VLSI for Solid Modelling // Advances in computer graphics. - Berlin etc.:Springer. 1988. - P.24-27.
95. Tomita F., Shirai Y., Tsuji S. Description of Textures by a Structural Analysis // IEEE Trans. on Pattern Anal. and Mach. Intel. 1982. - vol.PAMI 4. - № 2. - P.183-191.

ПРИЛОЖЕНИЕ

РАБОЧИЕ ФОРМУЛЫ

Приводятся аналитические формулы геометрического описания объектов главы 3. Нумерация формул и названия разделов указываются согласно основному описанию. Полезные дополнительные тождества, не использованные в основном тексте, обозначены значком (+).

3.1.2 Центральные и средние значения координат. Условие координатной симметрии.

$$\sum_{i=1}^n x_i \cdot y_i = \frac{\left(\sum_{i=1}^n x_i\right) \cdot \left(\sum_{i=1}^n y_i\right)}{n}. \quad (1)$$

3.1.3 Запись в 2-мерном подпространстве n -мерного векторного пространства.

$$(X, Y) - (X, E)(Y, E) \equiv (X', Y') = 0, \quad (2)$$

где $\begin{cases} X' = X - (X, E)E \\ Y' = Y - (Y, E)E \end{cases}$ - компоненты n -ых векторов X и Y ,

ортогональные нормированному вектору $E = \frac{1}{\sqrt{n}} \sum_{i=1}^n e_i$, $(E, E) = 1$.

$$\begin{cases} u \equiv \sqrt{(u, u)}(k \cdot \cos \varphi_1 + m \cdot \sin \varphi_1) = X' \\ v \equiv \sqrt{(v, v)}(k \cdot \cos \varphi_2 + m \cdot \sin \varphi_2) = Y' \end{cases}, \quad (3)$$

$$\cos(\varphi_1 - \varphi_2) = \frac{(u, v)}{\sqrt{(u, u)(v, v)}}. \quad (4)$$

$$\sin(\varphi_1 - \varphi_2) \neq 0 \quad (5)$$

Условие координатной симметрии: $(u, v) = 0$

$$A = \begin{cases} \tilde{x}_i = \alpha_{11}x_i + \alpha_{12}y_i \\ \tilde{y}_i = \alpha_{12}x_i + \alpha_{22}y_i \end{cases} \Rightarrow B = \begin{cases} \tilde{u} = \alpha_{11}u + \alpha_{12}v \\ \tilde{v} = \alpha_{12}u + \alpha_{22}v \end{cases}$$

3.1.4 Поворот. Определение ориентации.

$$\begin{cases} W\{u\} = u \cdot \cos \varphi + v \cdot \sin \varphi \\ W\{v\} = -u \cdot \sin \varphi + v \cdot \cos \varphi \end{cases}$$

$$(W\{u\}, W\{v\}) = (u, v) \cdot \cos 2\varphi - \frac{1}{2}[(u, u) - (v, v)] \cdot \sin 2\varphi = 0 \quad (6)$$

$$d = \sqrt{\frac{1}{4} \left(\sqrt{\frac{(u, u)}{(v, v)}} - \sqrt{\frac{(v, v)}{(u, u)}} \right)^2 + \frac{(u, v)^2}{(u, u)(v, v)}} \quad (7)$$

$$\operatorname{sh} \gamma = -\frac{1}{2} \left(\sqrt{\frac{(u, u)}{(v, v)}} - \sqrt{\frac{(v, v)}{(u, u)}} \right) \quad (8)$$

$$\operatorname{ch} \gamma = \frac{1}{2} \left(\sqrt{\frac{(u, u)}{(v, v)}} + \sqrt{\frac{(v, v)}{(u, u)}} \right) \geq 0$$

$$d = \sqrt{\cos^2(\varphi_1 - \varphi_2) + \operatorname{sh}^2 \gamma} \quad (9)$$

$$\begin{aligned} d &= \sqrt{(\operatorname{ch} \gamma \cdot \cos(\varphi_1 - \varphi_2))^2 + (\operatorname{sh} \gamma \cdot \sin(\varphi_1 - \varphi_2))^2} \equiv \\ &\equiv \sqrt{\operatorname{ch}^2 \gamma - \sin^2(\varphi_1 - \varphi_2)} \end{aligned}$$

$$(W\{u\}, W\{v\}) = \sqrt{(u, u)(v, v)} \cdot d \cdot \sin 2(\varphi_0 - \varphi) = 0 \quad , \quad (10)$$

$$\begin{cases} \sin 2\varphi_0 = \sigma \frac{\cos(\varphi_1 - \varphi_2)}{d} \\ \sin 2\varphi_0 = -\sigma \frac{\text{sh } \gamma}{d} \end{cases} \quad (11)$$

3.1.5 Собственные преобразования Лоренца.

Приведение к собственным осям.

$$\begin{cases} L\{u\} = (\text{ch } \theta + \text{sh } \theta) \cdot u \\ L\{v\} = (\text{ch } \theta - \text{sh } \theta) \cdot v \end{cases} \cdot$$

$$WL: \begin{cases} WL\{u\} = e^\theta (u \cdot \cos \varphi + v \cdot \sin \varphi) \\ WL\{v\} = e^{-\theta} (-u \cdot \sin \varphi + v \cdot \cos \varphi) \end{cases} \quad (12)$$

$$(WL\{u\}, WL\{v\}) = (W\{u\}, W\{v\}) \quad (13)$$

$$\text{th } 2\theta_0 = -\sigma \frac{d}{\text{ch } \gamma} \quad (14)$$

$$\text{sh } 2\theta_0 = -\sigma \frac{d}{|\sin(\varphi_1 - \varphi_2)|}$$

$$\text{ch } 2\theta_0 = \frac{\text{ch } \gamma}{|\sin(\varphi_1 - \varphi_2)|}$$

$$(W \cdot L\{u\}, W \cdot L\{u\}) \equiv (W \cdot L\{v\}, W \cdot L\{v\}) = \sqrt{(u, u)(v, v)} \cdot |\sin(\varphi_1 - \varphi_2)| \quad (15)$$

$$\begin{cases} u = e^{-\theta} \cdot \cos \varphi \cdot WL\{u\} - e^{\theta} \cdot \sin \varphi \cdot WL\{v\} \\ v = e^{-\theta} \cdot \sin \varphi \cdot WL\{u\} + e^{\theta} \cdot \cos \varphi \cdot WL\{v\} \end{cases} \quad (16)$$

3.1.6 Интерпретация.

Размеры фигуры (определение):

$$\begin{aligned} \text{длина} &= \max \left(\sqrt{\frac{(W\{u\}, W\{u\})}{n}}, \sqrt{\frac{(W\{v\}, W\{v\})}{n}} \right) \\ \text{ширина} &= \min \left(\sqrt{\frac{(W\{u\}, W\{u\})}{n}}, \sqrt{\frac{(W\{v\}, W\{v\})}{n}} \right) \end{aligned}$$

Размеры фигуры (рабочие формулы):

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

$$e^{2\theta_0} = \sqrt{\frac{\text{ch } \gamma - \sigma \cdot d}{\text{ch } \gamma + \sigma \cdot d}}$$

$$\theta_0 = \frac{\sigma}{2} \cdot \ln(\text{длина}/\text{ширина})$$

$$|\theta_0| = \frac{1}{2} \cdot \ln(\text{длина}/\text{ширина}) \quad (+)$$

$$d = n \cdot \frac{(\text{длина})^2 - (\text{ширина})^2}{2 \cdot \sqrt{(u, u) \cdot (v, v)}}$$

$$\text{площадь} = \frac{\sqrt{(u, u) \cdot (v, v)}}{n} \cdot |\sin(\varphi_1 - \varphi_2)|$$

$$\text{плотность} = \frac{n^2}{\sqrt{(u, u) \cdot (v, v)} \cdot |\sin(\varphi_1 - \varphi_2)|}$$

3.1.7 Разложение $W \cdot L$ в произведение ортогонального и симметричного операторов. Отождествление плоскостей.

$$V^{-1} : \begin{cases} V^{-1} \cdot W \cdot L\{u\} = W \cdot L\{u\} \cdot \cos \varphi_0 - W \cdot L\{v\} \cdot \sin \varphi_0 \\ V^{-1} \cdot W \cdot L\{v\} = W \cdot L\{u\} \cdot \sin \varphi_0 + W \cdot L\{v\} \cdot \cos \varphi_0 \end{cases} \quad (17)$$

$$L\{w\} = \operatorname{ch} \theta_0 \cdot w + \operatorname{sh} \theta_0 \cdot [(w, k) \cdot m + (w, m) \cdot k] \quad (18)$$

при

$$\begin{cases} \sin(\varphi_1 + \varphi_2) = \sigma \frac{\operatorname{ch} \gamma}{d} \cos(\varphi_1 - \varphi_2) \\ \cos(\varphi_1 + \varphi_2) = -\sigma \frac{\operatorname{sh} \gamma}{d} \sin(\varphi_1 - \varphi_2) \end{cases} \quad (19)$$

достигается:

$$W \cdot L = V \cdot L \quad . \quad (20)$$

$$\begin{aligned} \operatorname{ch} 2\theta_0 \cdot \cos(\varphi_1 - \varphi_2) + \operatorname{sh} 2\theta_0 \cdot \sin(\varphi_1 + \varphi_2) &= 0 \\ \cos(\varphi_1 - \varphi_2) + \operatorname{sh} 2\theta_0 \cdot |\sin(\varphi_1 - \varphi_2)| \cdot \sin 2\varphi_0 &= 0 \end{aligned} \quad (22)$$

$$-\operatorname{sh} \gamma \cdot \operatorname{ch} 2\theta_0 \cdot \sin(\varphi_1 - \varphi_2) + \operatorname{ch} \gamma \cdot \operatorname{sh} 2\theta_0 \cdot \cos(\varphi_1 + \varphi_2) = 0 \quad (+)$$

$$\cos(\varphi_1 + \varphi_2) = \cos 2\varphi_0 \cdot \sin(\varphi_1 - \varphi_2) \quad (+)$$

$$\sin(\varphi_1 + \varphi_2) = \cos 2\varphi_0 \cdot \operatorname{ch} \gamma \quad (+)$$

$$\operatorname{ch} \gamma = \operatorname{ch} 2\theta_0 \cdot |\sin(\varphi_1 - \varphi_2)| \quad (+)$$

$$V(\varphi_0) \cdot L(\theta_0) \Big|_{xxy} \Rightarrow V(\varphi_0) \cdot L(\theta_0) \Big|_{uxy}$$

Нормировка $W \cdot L\{u\}$ и $W \cdot L\{v\}$ определяет ортонормированные e_a, e_b :

$$e_a = \frac{W \cdot L\{v\}}{\sqrt{\sqrt{(u,u)} \cdot (v,v) \cdot |\sin(\varphi_1 - \varphi_2)|}}$$

$$e_b = \frac{W \cdot L\{u\}}{\sqrt{\sqrt{(u,u)} \cdot (v,v) \cdot |\sin(\varphi_1 - \varphi_2)|}}$$

Из (16) :

$$\frac{(u, e_a)}{\sqrt{(u, u)}} = - \sqrt[4]{\frac{(v, v)}{(u, u)}} \cdot \sqrt{|\sin(\varphi_1 - \varphi_2)|} \cdot e^{\theta_0} \cdot \sin \varphi_0 = -e^{\frac{\gamma}{2}} \cdot \sqrt{\text{ch } \gamma - \sigma \cdot d} \cdot \sin \varphi_0$$

$$\frac{(u, e_b)}{\sqrt{(u, u)}} = - \sqrt[4]{\frac{(v, v)}{(u, u)}} \cdot \sqrt{|\sin(\varphi_1 - \varphi_2)|} \cdot e^{-\theta_0} \cdot \cos \varphi_0 = -e^{\frac{\gamma}{2}} \cdot \sqrt{\text{ch } \gamma + \sigma \cdot d} \cdot \cos \varphi_0$$

$$\frac{(v, e_a)}{\sqrt{(v, v)}} = \sqrt[4]{\frac{(u, u)}{(v, v)}} \cdot \sqrt{|\sin(\varphi_1 - \varphi_2)|} \cdot e^{\theta_0} \cdot \cos \varphi_0 = e^{\frac{\gamma}{2}} \cdot \sqrt{\text{ch } \gamma - \sigma \cdot d} \cdot \cos \varphi_0$$

$$\frac{(v, e_b)}{\sqrt{(v, v)}} = - \sqrt[4]{\frac{(u, u)}{(v, v)}} \cdot \sqrt{|\sin(\varphi_1 - \varphi_2)|} \cdot e^{-\theta_0} \cdot \sin \varphi_0 = -e^{\frac{\gamma}{2}} \cdot \sqrt{\text{ch } \gamma + \sigma \cdot d} \cdot \sin \varphi_0$$

Из (3), (19) и последних равенств следует :

$$k = \frac{-e_a + \varepsilon_s \cdot e_b}{\sqrt{2}}$$

$$m = \frac{e_a + \varepsilon_s \cdot e_b}{\sqrt{2}} \quad ,$$

где $\varepsilon_s = \frac{|\sin(\varphi_1 - \varphi_2)|}{\sin(\varphi_1 - \varphi_2)}$.

Тригонометрические коэффициенты разложения u, v в базисе k, m (3) выражаются в виде:

$$\begin{aligned}
 \cos \varphi_1 &= e^{\frac{\gamma}{2}} \cdot \frac{\sqrt{\operatorname{ch} \gamma + \sigma \cdot d} \cdot \sin \varphi_0 - \varepsilon_s \cdot \sqrt{\operatorname{ch} \gamma - \sigma \cdot d} \cdot \cos \varphi_0}{\sqrt{2}} \\
 \sin \varphi_1 &= -e^{\frac{\gamma}{2}} \cdot \frac{\sqrt{\operatorname{ch} \gamma + \sigma \cdot d} \cdot \sin \varphi_0 + \varepsilon_s \cdot \sqrt{\operatorname{ch} \gamma - \sigma \cdot d} \cdot \cos \varphi_0}{\sqrt{2}} \\
 \cos \varphi_2 &= -e^{\frac{\gamma}{2}} \cdot \frac{\sqrt{\operatorname{ch} \gamma + \sigma \cdot d} \cdot \cos \varphi_0 + \varepsilon_s \cdot \sqrt{\operatorname{ch} \gamma - \sigma \cdot d} \cdot \sin \varphi_0}{\sqrt{2}} \\
 \sin \varphi_2 &= e^{\frac{\gamma}{2}} \cdot \frac{\sqrt{\operatorname{ch} \gamma + \sigma \cdot d} \cdot \cos \varphi_0 - \varepsilon_s \cdot \sqrt{\operatorname{ch} \gamma - \sigma \cdot d} \cdot \sin \varphi_0}{\sqrt{2}}
 \end{aligned} \tag{+}$$

V и L задаются посредством e_a, e_b в виде:

$$V: \begin{cases} V\{e_a\} = e_a \cdot \cos \varphi_0 + e_b \cdot \sin \varphi_0 \\ V\{e_b\} = -e_a \cdot \sin \varphi_0 + e_b \cdot \cos \varphi_0 \end{cases} \tag{24}$$

$$L\{w\} = e^{-\theta_0} \cdot (w, e_a) \cdot e_a + e^{\theta_0} \cdot (w, e_b) \cdot e_b$$

Активная и пассивная интерпретации:

- активная интерпретация:

$$A\{w\} = \sum_j (w, A^+\{e_j\}) \cdot e_j = \sum_j (w, e_j) \cdot A\{e_j\} ;$$

- пассивная интерпретация:

$$w = \sum_j (w, A^{+1}\{e_j\}) \cdot A\{e_j\} = \sum_j (w, e_j) \cdot e_j .$$

Отождествление плоскости $x \times y$ и двумерного подпространства $u \times v$ сводится к тому, что в подпространстве $u \times v$ n -мерного пространства вводятся точки фигуры w_i :

$$w_i = x_i \cdot e_x + y_i \cdot e_y ,$$

где e_x, e_y - некоторый ортонормированный базис $u \times v$.

В случае активной интерпретации:

$$\begin{cases} e_x = e_b \\ e_y = e_a \end{cases} .$$

$$\begin{cases} x_i = (w_i, V^{-1} \cdot L_1 \{e_x\}) = e^{\theta_0} (x_i \cdot \cos \varphi_0 + y_i \cdot \sin \varphi_0) \\ y_i = (w_i, V^{-1} \cdot L_1 \{e_y\}) = e^{-\theta_0} (-x_i \cdot \sin \varphi_0 + y_i \cdot \cos \varphi_0) \end{cases} .$$

где $L_1 = V \cdot L_1 \cdot V^{-1}$:

$$L_1 \{w_i\} = e^{\theta_0} \cdot (w_i, V^{-1} \{e_x\}) \cdot V^{-1} \{e_x\} + e^{-\theta_0} \cdot (w_i, V^{-1} \{e_y\}) \cdot V^{-1} \{e_y\}$$

При пассивной интерпретации:

$$\begin{cases} e_x = e_a \\ e_y = e_b \end{cases} .$$

$$\begin{cases} x_i = (w_i, V \cdot L^{-1} \{e_x\}) = e^{\theta_0} (x_i \cdot \cos \varphi_0 + y_i \cdot \sin \varphi_0) \\ y_i = (w_i, V \cdot L^{-1} \{e_y\}) = e^{-\theta_0} (-x_i \cdot \sin \varphi_0 + y_i \cdot \cos \varphi_0) \end{cases} .$$

3.1.8 Характеристика фигуры посредством сдвига системы координат.

$$\begin{aligned} (\tilde{u}, \tilde{u}) &= (u, u) + n \cdot (\delta x)^2 \\ (\tilde{v}, \tilde{v}) &= (v, v) + n \cdot (\delta y)^2 \\ (\tilde{u}, \tilde{v}) &= (u, v) + n \cdot \delta x \cdot \delta y \end{aligned} ,$$

$$\begin{aligned}
\delta x &= \varepsilon_x \cdot \sqrt{\frac{\sqrt{(u, u) \cdot (v, v)}}{n} \cdot (d + \operatorname{sh} \gamma)} \\
\delta y &= \varepsilon_y \cdot \sqrt{\frac{\sqrt{(u, u) \cdot (v, v)}}{n} \cdot (d - \operatorname{sh} \gamma)} \quad , \\
\varepsilon_x \cdot \varepsilon_y &= -\operatorname{sign}\{\cos(\varphi_1 - \varphi_2)\}
\end{aligned} \tag{25}$$

где $\varepsilon_x, \varepsilon_y = \pm 1$.

Достигается:

$$\begin{aligned}
(\tilde{u}, \tilde{v}) &= 0 \\
(\tilde{u}, \tilde{u}) &= (\tilde{v}, \tilde{v}) = \sqrt{(u, u) \cdot (v, v)} \cdot (\operatorname{ch} \gamma + d) \\
(\delta x)^2 + (\delta y)^2 &= (\text{длина})^2 - (\text{ширина})^2 = 2 \cdot \frac{\sqrt{(u, u) \cdot (v, v)}}{n} \cdot d
\end{aligned}$$

- 174 -



А К Т

использования алгоритмов и программ
 сегментации изображений в работах ПМГРЭ ПГО
 "Севморгеология"

нград

16 августа 1989г.

Комиссией в составе:

председатель КАЦЕВ В.А.
 члены МЕЛАМУД М.Е.
 СЛИОЗБЕРГ Н.С.

установлено следующее:

1. В соответствии с планами опытно-методических работ ПЦОИ ПМГРЭ ПГО "Севморгеология" разрабатывается комплект программ обработки изображений морского дна. С целью повышения эффективности обработки в комплекс программ включены программы сегментации изображений, разработанные на языке Микро Ассемблер комплекта "PERICOLOR 2000".

2. Указанные программы являются результатом совместной научной работы Искандерова П.М., Слиозберга Н.С., Харина М.В. Итеративно-рекурсивный метод предварительной обработки изображений в задаче выделения конкреций изложен в сборнике: "Информационные проблемы автоматизации" Л.ЛИИАН, 1988, с.с.107-114. Использование итеративно-рекурсивного метода позволяет повысить качество сегментации глубоковод-

2.

ных снимков и создает возможность автоматизированной оценки залежей.

3. В настоящее время программы сегментации реализованы на комплексе обработки изображений PERICOLOR 2000E и используются при предварительной обработке фотоснимков морского дна.

Председатель комиссии

В.А.Кацев

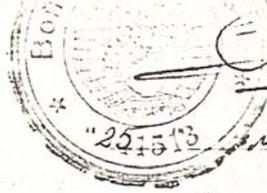
Члены комиссии:

М.Е.Меламуд

Н.С.Слиозберг

"УТВЕРЖДАЮ"

Заместитель командира войсковой части 41513
по научной работе
доктор технических наук профессор



В.Ф.ФАТЕЕВ

"22.03.1993" марта 1993г.

А К Т

реализации результатов диссертационных исследований
ХАРИНОВА М.В. в войсковой части 41513

Комиссия в составе председателя ктн снс ИЛЬИНА А.А. и
членов: ктн снс КОЧЕЛАЕВА Ю.С., ктн снс ТЕРЕХОВА В.К.,
ктн ПРИЩЕПА Ю.В. установила следующее:

Программное обеспечение, разработанное в рамках диссертационной работы Харинова М.В. и предназначенное для сегментации, многоуровневого представления и сжатия полутонных изображений, принято к эксплуатации в составе комплексного моделирующего стенда в/ч 41513 в качестве средств информационной поддержки и ведения специализированных баз данных видовой информации в условиях ограниченных ресурсов ЭВМ.

Председатель комиссии: ктн снс *Ильин* ИЛЬИН А.А.

Члены комиссии: ктн снс *Кочелав* КОЧЕЛАЕВ Ю.С.
ктн снс *Терехов* ТЕРЕХОВ В.К.
ктн *Прищеп* ПРИЩЕПА Ю.В.



"УТВЕРЖДАЮ"

Директор ТЦ ИНФОРМ

А.Д. Тонкин

" 9 " апреля 1993 г.

А К Т

передачи прототипа программной обуча-
емой системы автоматического распоз-
навания изображений

Настоящим актом подтверждается, что результаты диссертационной работы Харинова Михаила Вячеславовича "Разработка динамических структур данных системы автоматизированного распознавания изображений" реализованы в виде прототипа программной обучаемой системы многоцелевого распознавания и переданы ТЦ ИНФОРМ для опытной эксплуатации.

Прототип системы многоцелевого распознавания разработан для ЭВМ PC AT/286 на языке СИ в OS MS DOS 3.30, предназначен для обработки сегментированных полутоновых и двухградационных изображений и включает:

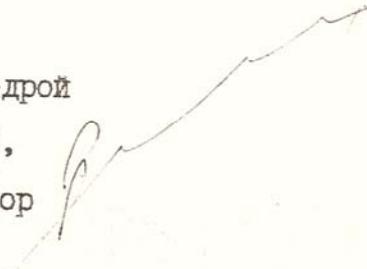
I. Программу RDR спецификации первых и вторых моментов связанных областей изображения, обеспечивающую подготовку данных для вычисления яркостных и геометрических характеристик наблюдаемых объектов в процессе однократного построчного сканирования.

2. Программу RCGR частичного анализа связанных областей, обеспечивающую формирование блока видеоданных в виде первых и вторых моментов элементов изображения с последующим автоматическим распознаванием объектов по относительным метрическим характеристикам соизмерных частей.

3. Программу CHILD (модификация RCGR), обеспечивающую формирование ассоциативной обучающей информации посредством интерактивного обучения на примерах.

Характерным отличием данного технического решения является реализация выбора необходимого набора признаков из предусмотренного множества. Система перечисленных программ обеспечивает новые возможности создания многоцелевых комплексов автоматизированного распознавания изображений, а также экспериментальное исследование и адаптацию признаковых систем к конкретным областям применения.

Заведующий кафедрой
автоматики ЛИВТ,
д.т.н., профессор



Францев Р.Э.