

В.А. ГОРОХОВАТСКИЙ

Т.В. ПОЛЯКОВА

**ПРИМЕНЕНИЕ
ПРОСТРАНСТВЕННЫХ СТРУКТУР ПРИЗНАКОВ
ДЛЯ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ
В КОМПЬЮТЕРНОМ ЗРЕНИИ**

Харьков - 2018

УДК 004.932.2:004.93'1

Гороховатский В.А., Полякова Т.В.

Применение пространственных структур признаков для классификации изображений в компьютерном зрении (монография)/ В.А. Гороховатский, Т.В. Полякова. – Харьков: ФОП Панов А.Н., 2018. – 120 с.

ISBN 978-617-7541-58-4

Монография содержит материалы исследований в области компьютерного зрения в плане создания структурных методов распознавания объектов на изображениях в условиях геометрических преобразований и влияния помех. Предложены современные методы формирования структурных описаний изображений, построения отношений и пространственных структур признаков, вычисления мер подобия описаний. Осуществлена трансформация описаний путем фильтрации и сжатия в целях улучшения показателей быстродействия и достоверности, изучены подходы к оптимизации параметров предложенных методов. Описаны результаты экспериментального исследования и применения разработанных методов при решении задач распознавания для прикладных баз изображений.

Для научных работников, преподавателей, аспирантов, студентов и всех, кто интересуется проблемами создания информационных технологий и интеллектуальных методов обработки данных и их применения в системах компьютерного зрения.

УДК 004.932.2:004.93'1

ISBN 978-617-7541-58-4

СОДЕРЖАНИЕ

Основные сокращения	4
Введение	5
1. Анализ методов и моделей классификации изображений на основе структурных описаний	9
1.1 Структурное представление в задаче анализа изображений	10
1.2 Построение инвариантных структурных описаний	23
1.3 Сравнительный анализ моделей установления эквивалентности структурных описаний	29
1.4 Постановка задач исследований	36
Выводы к разделу	37
2. Построение и сопоставление описаний визуальных объектов в виде множеств признаков пространственных структур	38
2.1 Особенности построения современных способов для формирования систем характерных признаков	38
2.2 Построение описаний на основе признаков пространственных структур	45
2.3 Метод сопоставления описаний в виде ППС при классификации	54
2.4 Конфигурации описания и модели голосования при сопоставлении ППС	58
Выводы к разделу	64
3. Повышение эффективности классификации путем группового анализа элементов описаний	66
3.1 Пространственная группировка элементов описания	66
3.2 Фильтрация множества ППС	70
3.3 Гибридные модели признаков пространств	74
Выводы к разделу	80
4. Оптимальное сопоставление описаний как множеств признаковых структур. Результаты экспериментальных исследований и внедрения методов	82
4.1 Постановка и решение задачи оптимизации при сопоставлении описаний	83
4.2 Результаты экспериментов для моделей классификации на основе разнотипных признаков	92
4.3 Результаты экспериментов по сопоставлению на основе моделей инвариантных признаков структур	97
Выводы к разделу	104
Выводы	106
Список использованных источников	108

ОСНОВНЫЕ СОКРАЩЕНИЯ

АБ	– аффинный базис
АИ	– аффинный инвариант
АП	– аффинное преобразование
БК	– барицентрические координаты
БП	– базовые признаки
ВМ	– Венгерский метод
ГИП	–геометрические инвариантные признаки
ГП	– геометрическое преобразование
ГХ	– геометрическое хеширование
КТ	– ключевые точки
МБЛП	–метод базовых локальных признаков
ППС	– признаковые пространственные структуры
ПХ	– преобразование Хафа
СП	– структурные признаки
СТЗ	– системы технического зрения
СЭ	– структурные элементы
ХП	– характерные признаки

ВВЕДЕНИЕ

Перспективным инновационным направлением в мире компьютерных технологий считается компьютерное зрение, а также сопряженные технологий программной и аппаратной обработки в системах распознавания и анализа изображений. Ключевой и практически важной задачей при создании систем искусственного интеллекта для современных высокотехнологических производств является автоматизация рутинного анализа визуальных данных, связанная с анализом и обработкой изображений с нескольких полей зрения, а также классификацией визуальных объектов для получения статистических данных.

Значительный вклад в развитие методов распознавания изображений в интеллектуальных СТЗ внесли исследования украинских и зарубежных ученых, таких как: С.Г. Антошук, А.М. Ахметшин, Р.А. Воробель, Ю.И. Журавлев, А.Г. Ивахненко, В.А. Ковалевский, В.Г. Лабунец, В.П. Машталир, Е.П. Путятин, Б.П. Русын, И.Б. Сироджа, Ю.А.Скобцов, Ю.П. Шабанов-Кушнаренко, М.И. Шлезингер, R.O. Duda, D.A. Forsyth, R.C. Gonzales, P.V.C. Hough, S. Kim, D.G. Lowe, W.K. Pratt, A. Rosenfeld, L.G. Shapiro, C. Schmid, а также ряд других [1–16]. Однако, сегодня проблема распознавания визуальных объектов для реальных применений еще далека от своего исчерпывающего решения.

Развитие технологий искусственного интеллекта в системах компьютерного зрения вызывает острую потребность в разработке новых эффективных подходов, связанных с растущей сложностью решаемых проблем. Решение задачи распознавания изображений значительно усложняется изменениями объекта под влиянием внешней среды в виде фона и помех, а также в условиях действия геометрических преобразований. Такое воздействие может привести к искажениям фрагментов анализируемого объекта, а также сформировать ложные объекты с близкими характеристиками [1–23].

Эффективность методов интеллектуального анализа видео-объектов в значительной мере зависит от характера представлений о строении описаний

множеств их элементов. В случаях, когда описание представляется как множество разрозненных элементов, трудно рассчитывать на создание высокоэффективных методов синтеза или классификации объектов. Рациональность методов анализа описаний повышается, если известна структура множества, устанавливающая определенные отношения между входящими в него элементами.

Структурные методы, основанные на анализе особенностей изображения в отдельных точках, позволяют извлекать из изображения осмысленную и структурированную информацию. Анализ свойств конечного числа характерных признаков (ХП) позволяет значительно упростить решение практических задач. Особую ценность для задач инвариантного распознавания видео-объектов имеет геометрическая информация, содержащаяся в координатах ХП. На основе координат ХП можно построить признаковые структуры более сложного плана, к которым относятся геометрические инвариантные признаки (ГИП), построенные на вычислении аффинных инвариантов (АИ) [4]. Вектор АИ отражает геометрическую форму объекта и обладает инвариантностью к аффинным преобразованиям на плоскости. ГИП можно считать частным случаем более общих представлений в виде признаков пространственных структур (ППС), где геометрические свойства интегрируются со свойствами детекторов ХП.

Существующие методы анализа объектов в реальных сценах не позволяют качественно решить сложные задачи распознавания, связанные с искажениями объектов под действием композиции из нескольких аффинных преобразований. Разработка новых моделей и методов на основе синтеза пространственных структур признаков позволит повысить эффективность решения прикладных задач.

Целью монографии есть теоретическое развитие и экспериментальное исследование методов интерпретации видео-объектов на основе структур пространственных инвариантных признаков для решения задачи распознавания при воздействии геометрических преобразований и помех. Применение

предложенных методов обеспечивает высокую достоверность классификации в условиях неопределенности при влиянии фоновых и локальных искажений.

Изучение принципов формирования и применения пространственных структур признаков осуществлено на основе положений теории регистрации, обработки и анализа изображений, элементов теории множеств и математической логики; при синтезе методов сопоставления структурных описаний применен аппарат распознавания образов, кластеризации, теории метрических пространств. Для проверки теоретических положений и оценки эффективности предложенных методов и моделей для баз реальных изображений применено имитационное моделирование.

В монографии в развитие исследования [12, 25] в общей постановке решается важная научно-прикладная задача классификации изображений визуальных объектов в условиях геометрических преобразований и помех внешней среды на основе применения разработанного метода с использованием моделей пространственных структур характерных инвариантных признаков.

Применение предложенных методов в интеллектуальных системах компьютерного зрения дает возможность осуществить создание универсальных или специализированных программно-аппаратных средств распознавания видеоинформации. Полученные результаты существенно расширяют спектр и уровень сложности решаемых задач компьютерного зрения, в частности, повышают надежность слежения и автоматического распознавания объектов в условиях организованных помех. Созданные компьютерные модели дают возможность оценить эффективность распознавания и проводить последующее усовершенствование методов обработки. Реализация результатов представлена в виде исследовательских программных комплексов и ряда отдельных специализированных программных систем, в частности, для решения прикладных задач определения архитектурного стиля, а также в технологиях поиска и обработки синтезированных изображений.

Авторы выражают благодарность и наилучшие пожелания любимому учителю и научному консультанту – Заслуженному деятелю науки и техники Украины, доктору технических наук, профессору

Путятину Евгению Петровичу

за всестороннюю поддержку, плодотворное сотрудничество в рамках научной школы кафедры информатики ХНУРЭ, обсуждение проблем, заданий, направлений, результатов и их применения при проведении многолетних совместных исследований в сфере задач компьютерного зрения.

1 АНАЛИЗ МЕТОДОВ И МОДЕЛЕЙ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ СТРУКТУРНЫХ ОПИСАНИЙ

Качественный переход на новые технологии стал причиной развития и достижений, в частности в области сетевых, микро-технологий, компьютерной графики и мультимедийных систем. В то же время возможности автоматизированного интеллектуального анализа изображений пока не обладают достаточной гибкостью при решении задач, связанных с углубленной обработкой и распознаванием. Одной из центральных и практически важных задач на сегодня при создании систем искусственного интеллекта является распознавание изображений [1–21]. Это обуславливает многообразие исследований в плане сопоставления огромного разнообразия визуальных объектов с целью разработки новых эффективных и устойчивых методов структурного анализа изображений в условиях геометрических преобразований (ГП) и помех [22–46].

Решать задачу распознавания объектов необходимо, в частности, при поиске соответствия неизвестного объекта эталону, хранящемуся в базе визуальных данных. Сложность сопоставления заключается в зависимости от вида сопоставляемых объектов и точки обзора. Для успешного поиска соответствий объектов необходима процедура совмещения точек обзора. Она включает в себя поиск в многомерном пространстве с перебором значений всех возможных параметров преобразования. Этот поиск является очень дорогостоящим с точки зрения вычислительной сложности, если может быть выполнен в принципе.

Проведем анализ современного состояния и направлений развития основных подходов к сопоставлению структурных описаний для решения задачи распознавания видео-объектов при влиянии ГП и помех. Рассмотрим существующие модели и методы, используемые в системах технического зрения (СТЗ) на основе структур данных визуальных объектов. Установлено,

что основную вычислительную сложность представляет процесс сопоставления описаний в признаковом пространстве, а также процедура определения эквивалентности объекта с множеством эталонов. Перспективу имеют исследования методов анализа и представления множества структурных элементов описаний, а также подходы к сокращению числа характерных признаков для повышения быстродействия и достоверности.

1.1 Структурное представление в задаче анализа изображений

Понятие структуры, в том числе и структуры данных, предполагает взаиморасположение и связь составных частей чего-либо. Под структурой объекта понимается способ объединения целого из отдельных частей – структурных элементов (СЭ) с их взаимосвязями. Эффективность применения структуры определяется количеством, значением, формой и содержанием ее СЭ, существующими между ними отношениями и местом, которое они занимают в целом [7].

В самом общем плане структурный подход к представлению объектов основывается на наличии внутренней и внешней структур связей, согласно которым в объекте выделяются и анализируются СЭ [17, 24]. Целостность объекта при воздействии изменений обеспечивается совокупностью устойчивых связей и отношений между его составляющими. Под СЭ в задачах анализа изображений подразумевается нечто большее, чем физические «детали» целого. В общем случае это – качества, признаки, описывающие объект в целом, его атрибуты. Распространение принципа деления объектов на уже выделенные компоненты целого позволяет строить многоуровневые иерархические представления. Части объекта могут интерпретироваться независимо друг от друга как элементы множества. Для задачи анализа изображений на множестве СЭ вводятся и интерпретируются различные семантические отношения (родовидовые, пространственные, временные и т.д.) между частями исходной сцены, выделенными объектами и их частями,

описывающие структуру данных. Принципы декомпозиции объектов и выделения системы отношений базируются, как правило, на интеллектуальных механизмах человеческого мышления. Термин «структурное распознавание образов» используется применительно ко всем методам, в которых объекты описываются в терминах частей и отношений [2, 8, 9, 10, 12].

Структурное представление данных, в основном, нацелено на описание, ориентированное на сокращение содержательной избыточности и уменьшение вычислительной сложности. При этом достигается сжатие исходного изображения в признаковом пространстве, и, как следствие, уменьшение требуемого объема памяти. Следовательно, уже само структурное представление видеоданных реализует некоторое сжатие.

В качестве основы структурного представления изображения используются такие элементы [78]:

- геометрические, в основе которых лежат точки, не объединенные в связанные контуры (точки пересечения, угловые, концевые точки и т.п.);
- выделенные элементы (контура), путем их сегментации и аппроксимации, обнаружения точек максимальной кривизны (направленные прямые линии, линии малой кривизны);
- характерные признаки (ХП) – точки, в которых значение некоторого локального признака достигает экстремального значения. В литературе также встречаются и другие названия ХП, такие как – точки интереса, ключевые точки (КТ) и, которые будем считать эквивалентными.

Существующие средства распознавания изображений используют различные способы сопоставления визуальных объектов с эталонами в базах изображений. Однако, основой для удачного сопоставления является формирование эффективного пространства структурных признаков, которое позволяет извлекать структурированную информацию об особых частях объектов или о нанесенных на них маркерах для их идентификации и распознавания. Процесс сопоставления, как правило, сводится к установлению соответствия между множествами ХП исходного изображения и эталонов.

По сути, ХП представляет собой наиболее простой геометрический элемент дискретного представления функции объекта распознавания. ХП должен иметь ряд параметров, существенно отличающих его от множества соседних с ним точек изображения объекта [12]. В частности, совокупность ХП может описывать некоторые характерные области на изображениях, например, вершины геометрических фигур, края плоскостей, углы, небольшие окружности и круги, перепады яркости или контрастности, резкие цветовые переходы и т.п.

Определение ХП основано на окрестности точки изображения. ХП будем называть точку u_i , окрестность которой $O'(u_i)$ можно отличить от окрестности $O'(u_j)$ любого другого ХП изображения u_j .

В настоящее время создана серия методов для определения ХП на изображениях с разными подходами к формированию пиксельной оценочной функции поиска и выделения признаков. К наиболее известным относятся:

- SIFT (Scale Invariant Feature Transform) [16]. Полученные точки в виде вектора-дескриптора инвариантны к масштабированию и поворотам изображения, устойчивы к изменениям освещения, шумам и изменениям позиции наблюдателя;

- SURF (Speeded Up Robust Features) [18]. Методика основана на поиске ХП и создании дескрипторов, инвариантных к масштабированию и вращению с помощью матрицы Гессе. При этом для каждой точки определяется градиент максимального изменения яркости и коэффициент масштабирования по матрице Гессе.

Основными недостатками методов SIFT и SURF являются такие, как нечеткое выделение объекта относительно фона и низкая вероятность правильного распознавания элементов изображений без ярко выраженной текстуры.

Дальнейшее развитие принципов обработки, положенных в основу методов SIFT, SURF, привело к созданию более эффективных по быстродействию методов ORB, BRISK [23].

В работе [37] предложен эффективный метод извлечения ХП ORB (Oriented FAST and Rotated BRIEF), основанный на сочетании алгоритма FAST (Features from Accelerated Test) [40] для поиска ХП и усовершенствованного подхода BRIEF (Binary Robust Independent Elementary Features) [42] для определения дескриптора ХП в виде бинарной строки.

Метод FAST является одним из наиболее распространенных в современной практике применения детекторов характерных точек, суть которого заключается в построении простых деревьев решений для классификации пикселей на ХП и остальные. На первом этапе поиска ХП для изображения строится масштабная пирамида Гаусса. Затем на каждом масштабном уровне определяются экстремумы яркости. Для этого применяется алгоритм FAST, согласно которому рассматривается окрестность в виде круга некоторого радиуса с центром в точке u с координатами (x, y) для каждой точки изображения и подсчитывается число смежных пикселей, яркость которых не принадлежит диапазону $[p(u) - t; p(u) + t]$, где $p(u)$ - яркость точки u , t - пороговая величина. Кандидатом на роль ХП объявляется точка u , окрестность которой содержит более 75% таких пикселей. В методе ORB предусмотрено регулирование количества ХП. Максимальное количество ХП по умолчанию не более 500, если их больше, то к ним применяется детектор углов Harris [93] для исключения наименее значимых детекторов.

На следующем этапе вычисляется параметр угловой ориентации θ найденных ХП для обеспечения устойчивости детектирования при вращении объекта. Для этого при текущем масштабе формируется область и вычисляются моменты первого порядка

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y),$$

где $I(x, y)$ – функция изображения, а $\forall p, q \in \{0, 1\}$ – параметры выбора x и y ориентации.

Затем определяется ее центр массы $c = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right)$ и параметр угловой ориентации $\theta = \arctan 2(m_{01}, m_{10})$. Ориентацию ХП задает вектор, направленный от координат ХП к центру массы. Введение параметра угловой ориентации позволяет добиться некоторой инвариантности детектирования при вращении объекта. Он основан на направлениях градиента яркости относительно центра точки, направление с наибольшей интенсивностью назначается ориентацией ХП.

Бинарный вектор дескриптора BRIEF длиной $n = 256$ вычисляется путем сравнения набора яркости пар областей внутри квадратного окна центрированного относительно координат ХП и согласованного с ее ориентацией [42]. В квадратном окне по заданному правилу выбирается набор пар областей x, y , значение яркости в которых сравниваются между собой

$$\tau(p; x, y) = \begin{cases} 1 & \text{if } p(x) \geq p(y) \\ 0 & \text{otherwise} \end{cases},$$

где p – средняя яркость выбранной области. Далее формируется дескриптор

$$f_n(p) = \sum_{i=1}^n 2^{i-1} \tau(p; x_i, y_i).$$

Структура алгоритма ORB показывает, что он менее требователен к вычислительным ресурсам по сравнению с другими алгоритмами. Выигрыш в скорости вычислений объясняется простой процедурой построения дескрипторов. ORB дает по результатам тестирования заметный выигрыш в быстродействии при сравнительной или лучшей точности, чем SIFT и SURF [37].

Еще одним методом извлечения ХП на изображении, не уступающим по популярности ORB, является BRISK (Binary Robust Invariant Scalable Keypoints), представленный в работе [39]. В этом методе ХП извлекаются с помощью FAST при линейном изменении масштаба начального изображения. Дескрипторы описываются в бинарном виде по модифицированному алгоритму BRIEF [42].

В BRISK представлены различные альтернативы форм маски для выявления ХП. Используется маска 9–16, которая с помощью FAST критерия сравнивает яркости 9 последовательных точек в 16-пиксельном кругу с центральной точкой. Для разнообразия масштабов применены и другие маски.

Учитывая расположение ХП и соответствующие значения масштаба, BRISK составляет дескриптор в виде бинарной строки. Идентифицируется угол направления градиента каждого ХП для получения ориентированно-нормализованных дескрипторов и обеспечения инвариантности к вращению.

Концепция дескриптора BRISK использует шаблон путем анализа точек, равномерно распределенных по концентрическим кругам с центром в ХП. Это обеспечивает интегрированный анализ и высокую скорость обработки или хранения. Создание BRIEF-дескриптора обеспечивает распознавание одинаковых участков изображения, снятых с разных точек зрения.

Основные преимущества инновационных методов ORB и BRISK заключаются в обеспечении существенно более высокой производительности за счет упрощения процесса обработки и построения и использования дескриптора бинарного типа.

Каждой точке u_i с координатами (x_i, y_i) из множества U , извлеченных на изображении ХП, ставится в соответствие q -мерный вектор-дескриптор $v_i = (v_{i1}, \dots, v_{iq})$ из множества V , описывающий структуру окрестности точки, выделяющей её из остального множества точек, где $i = \overline{1, n}$, n – число элементов в описании, v_{iq} – компоненты вектора ХП. При этом дескриптор рассчитывается так, чтобы его значение было инвариантным по отношению к допустимым геометрическим преобразованиям изображения. Множество векторов-дескрипторов вместе с координатами ХП определяет описание изображения, которое в дальнейшем используется для сопоставления изображений с эталонными образцами из базы данных.

В частности, если U – множество векторов координат (это множество с неповторяющимися элементами), V – множество векторов признаков.

Определим отображение $\varphi: U \rightarrow V$, которое ставит в соответствие элементу $u_i \in U$ элемент $v_i \in V$, тогда $\varphi(u_i) = v_i$. Кортеж $Z = \langle V, U, \varphi \rangle$ будем считать структурным описанием объекта [45, 51].

Признак выражает некоторую значимую характеристику объекта в численной форме и определяется как функция от значений пикселя или группы пикселей. Под характерными признаками (ХП) понимают [8, 9] те точки изображения, которые являются наиболее информативными для распознавания. Перечисленные термины являются названиями элементов из множества локальных признаков (другие названия – структурные признаки (СП), структурные элементы, примитивы и т.д.), представляющих не все изображение в целом, а только значимые части.



Рисунок 1.1 – Примеры характерных признаков объекта

Выбор способа структурного представления данных определяет суть и цели решаемой задачи. Изображение является функцией $B(x, y)$ двух переменных – пространственных координат x, y элемента, а значение $B(x, y)$ характеризует яркость элемента в точке x, y .

Пространство W изображений $B(x, y) \in W(D)$ (D – поле зрения) можно представить как

$$W(D) = \{B(x, y), (x, y) \in D \mid D = \{(x, y)\},$$

$$(x, y) \in \{1, \dots, X\} \times \{1, \dots, Y\}, 0 \leq B \leq B^m\},$$

B^m – максимальная яркость, где функция $B(x, y)$ принимает дискретные значения [1].

Поиск соответствий между СЭ производится с учетом их взаимного расположения, размеров, взаимной ориентации и др. Структурный подход обеспечивает решение задачи классификации в сложных условиях, к которым относятся следующие. Изображения предъявляются на сложном фоне. Изображения классифицированных объектов (эталонов) и входные изображения отличаются положением в поле зрения. Под эталоном подразумевается нечто, принимаемое за образец, идеальный прототип, модель, правило или пример, содержащий информацию для ее сравнения и оценки в других объектах. Входные изображения не совпадают с эталонными за счет воздействия случайных помех. Основное назначение эталона состоит в том, что он содержит информацию для ее оценки в других объектах [1, 20]. Прикладное значение имеет изучение математических моделей вида

$$B(x, y) = \alpha(x, y)B_0[T(x, y)] + \xi(x, y), \quad (1.1)$$

где функции $\alpha(x, y), \xi(x, y)$ описывают действие фона и помех,

$T(x, y)$ – оператор ГП.

Отличия входных и эталонных изображений возникают также при изменении освещенности, подсветки, локальных помех. Эталоны и изображения могут не совпадать из-за наличия ГП, включая такие сложные, как аффинные и проективные.

Общепринятое понятие классификации заключается в поиске соответствия на множестве признаков наблюдаемого изображения и базы ранее классифицированных объектов – эталонов [8, 12]. Суть задачи классификации состоит в отображении пространства признаков в пространство решений. Результатом классификации является принятие решения о принадлежности объекта тому или иному классу. Структурная классификация имеет ряд преимуществ, таких как сокращение признакового пространства, универсальность в плане возможности структурного анализа ситуаций частичного изменения объекта, учет взаимных перемещений и геометрических трансформаций элементов, устойчивость к влиянию фоновых искажений и ложных объектов [1, 4, 5, 8, 20, 63].

Рассмотрим задачу классификации в следующей постановке. Задано $\Omega = \{U\}$ – конечное множество структурных описаний объектов, J – множество классов. Существует отображение $\Omega \rightarrow J$, значение которого определяется множеством эталонных описаний $E = \{E_j\}$, $E \subset \Omega$, $j \in J$. Классификацию можно сформулировать как построение разбиения множества Ω на непересекающиеся классы $j \in J$, что $U(j) \cap U(l) = \emptyset$, при $j \neq l$, $U(j), U(l) \in \Omega$ [1, 14]. При этом разбиение осуществляется на основе некоторого критерия.

Классификация, где описание распознаваемого объекта имеет вид конечного множества $U = \{u_i\}_{i=1}^m$ из элементов u_i , получила название структурной классификации.

В литературе обсуждается несколько основных подходов к структурному распознаванию изображений при действии геометрических искажений. Главное их отличие состоит в выборе меры сравнения и пространства, в котором осуществляется сопоставление описаний изображений объектов.

Одним из видов структурных методов являются синтаксические методы, которые используют представление образов в виде предложений специального языка. При применении синтаксических методов формируются структурно-лингвистические признаки, когда изображение дробится на части – производные элементы. Вводятся правила соединения этих элементов, одинаковые для эталона и входного изображения. Правило классификации сводится к выводу грамматики, описывающей язык классифицируемых образов. Распознавание заключается в определении выводимости рассматриваемого предложения с помощью найденной грамматики [2, 28].

В настоящее время развиваются принципы локальной обработки, к которым относятся методы, основанные на преобразовании Хафа (ПХ) [8], вейвлет-анализ [109], частные корреляции [3, 21, 75]. Эти методы являются структурным вариантом известных спектральных и интегральных методов, но по сравнению с ними обладают важными избирательными по пространству

свойствами, позволяющими анализировать на предмет соответствия эталону отдельные части сигнала изображения.

Методы, основанные на преобразовании Хафа, позволяют выделять простые формы на изображении и осуществлять поиск объектов, принадлежащих определённому классу фигур с использованием процедуры голосования. Они широко используются для анализа визуальных объектов в задачах компьютерного зрения [8].

Описания объектов здесь представляются в виде множеств элементов простых форм, таких как прямые, окружности и другие элементарные подобъекты. Аналитически задается модель преобразования. Задается также некоторая аналитическая поверхность $\gamma(x,a)=0$, где $x=(x_1, \dots, x_n)$ – вектор из пространства сигналов, $a=(a_1, \dots, a_m)$ – вектор из пространства параметров преобразования. Поверхность γ , задающая связь сигналов и их преобразований, включает в себя непосредственно и вид преобразования. В результате применения $\gamma(x,a)$ в пространстве сигналов выделяется подпространство, которое описывается характеристической функцией принадлежности сигнала подпространству, где выполняется условие $\gamma(x,a)=0$. Область значений параметров преобразований дискретизируется с некоторым шагом по каждой компоненте a_i вектора a . Формируется массив-накопитель A , в который записываются значения отсчетов. Суть получения спектров состоит в том, что для каждого вектора x находятся значения функции $\gamma(x,a)$ при всех возможных значениях a . В массиве-накопителе A последовательно путем обработки каждого СЭ увеличиваются на 1 (инкрементируются) значения ячеек, для которых выполняется $\gamma(x,a)=0$. После просмотра всех точек содержимое ячейки накопителя соответствует величине спектрального отсчета ПХ. В соответствии со значением максимума в массиве A принимается решение о величине параметра и о соответствии объекта эталону. Таким образом, критерием принятия решения при использовании ПХ есть число голосов, отданных за величину преобразования или класс объекта.

Эффективность использования методов на основе ПХ как методов извлечения глобальной информации из совокупности локальных свойств объектов в большой степени обусловлена качеством входных данных. Для зашумленных изображений необходим этап предварительной обработки с целью подавления шума. Вычислительные затраты этого метода значительно возрастают с увеличением числа и множества значений пространства параметров. Успех в применении ПХ достигнут в практических задачах при распознавании изображений японских иероглифов [110], при контурном анализе форм облаков, объектов с формой окружностей различных радиусов [112], объектов типа самолетов [29, 94], а также в задачах слежения за движущимися малоразмерными и протяженными объектами [30, 110].

Преобразование Хафа традиционно считается весьма ресурсоёмким. Вторым недостатком ПХ считается неустойчивость при наличии шума и малом числе исходных данных. Проблема неустойчивости решается сглаживанием исходной выборки с использованием параметра сглаживания, который находится подбором. При такой схеме ПХ приходится делать неоднократно, для различных вариантов сглаживания, что приводит к дополнительным вычислительным затратам.

Вейвлет-анализ относят к методам спектрального анализа, основой в котором являются функции особого рода, называемые вейвлетами [109]. Вейвлет-преобразование – интегральное преобразование, которое представляет собой свертку вейвлет-функций с сигналом. Это способ преобразования функции (или сигнала) $f(t)$ в форму, которая позволяет сжимать исходное множество данных и делает некоторые величины исходной функции удобными для изучения. В основе вейвлет-преобразований лежит использование двух непрерывных, взаимозависимых и интегрируемых по независимой переменной функций: вейвлет-функции $\psi(t)$, с помощью которой выделяются детали сигнала и его локальные особенности и масштабирующей функции $\phi(t)$ для аппроксимации сигнала.

В пространстве $L^2(R)$ функций $f(t)$, определенных на всей действительной оси $R(-\infty, +\infty)$ и обладающих конечной энергией (нормой)

$$E_f = \int_{-\infty}^{\infty} |f(t)|^2 dt < \infty, \quad \text{непрерывным вейвлет-преобразованием функции}$$

$f(t) \in L^2(R)$ называют функцию двух переменных:

$$T(a, b) = \langle f(t), \psi(a, b, t) \rangle = \int_{-\infty}^{\infty} f(t) \psi(a, b, t) dt,$$

где $b \in R$ – параметр положения, соответствующий сдвигу,

$a \in R, a \neq 0$ – параметр растяжения, который задает масштабирование,

$\psi(a, b, t)$ – вейвлеты, масштабированные и сдвинутые копии порождающего вейвлета $\psi(t) \in L^2(R)$.

Совокупность $\psi(a, b, t)$ создает новый базис пространства $L^2(R)$. Вейвлетный базис функционального пространства образуется путем масштабных преобразований и сдвигов порождающего вейвлета

$$\psi(a, b, t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right).$$

За счет изменения масштабов вейвлеты способны выявить различия в характеристиках на разных шкалах, а путем сдвига проанализировать свойства сигнала в различных точках на всем изучаемом интервале.

В дискретном представлении вейвлеты характеризуются числовыми коэффициентами в функциональных уравнениях, содержащих изменения масштаба и сдвиг. Функции вейвлетных преобразований порождают структуры данных вида $\{c, S\}$, где c – вектор коэффициентов преобразования, а S – управляющая матрица для определения порядка следования коэффициентов в векторе c . Общеизвестным недостатком вейвлетов является их достаточно высокая вычислительная сложность и необходимость хранения значительных объемов данных [109].

Классификация изображений с использованием СЭ признается менее трудоемкой с вычислительной точки зрения, так как размерность данных в этом

подходе ограничена. Поскольку СЭ не используют напрямую значения интенсивностей, они гораздо более устойчивы к таким отличиям сопоставляемых изображений, как изменение освещения, изменение типа сенсора и другие. Однако само выделение подобных признаков на изображении является трудной задачей. Для многих классов изображений это является серьезным недостатком, поскольку точность регистрации может оказаться не лучше, чем точность, достигаемая при первоначальном выборе опорных точек. Число структурных элементов и точность положения соответствующих им опорных точек обычно ограничены [16, 22, 43].

Помехоустойчивыми при действии как случайных, так и локальных помех являются методы, основанные на подходе частных корреляций [3, 21]. Это структурный вариант одного из наиболее применяемых в приложениях корреляционного подхода. При этом частичные коэффициенты корреляции, полученные для отдельных фрагментов эталона в сигнальном пространстве, могут рассматриваться как признаки (в общем случае разношкальные). Обработка таких признаков, т.е. их свертка, зависит от типа изображений, помеховой обстановки (например, степени заслонения полезного изображения, наличия ложных изображений в поле зрения) и может быть осуществлена методами проверки статистических гипотез. Общеизвестным недостатком корреляционных подходов считается существенное увеличение размерности пространства признаков при распознавании в условиях действия многопараметрических групп геометрических преобразований.

Значительно более простые с точки зрения вычислительной сложности методы основаны на переходе в пространство признаков, которые характеризуются существенно меньшей размерностью по сравнению с пространством сигналов (изображений). В зависимости от поставленной цели (например, достижения заданной точности) выполняется корреляционная обработка признаков, полученных от эталона и входного изображения как с использованием порогов по величине сходства, так и без установления порога (когда ищется максимум сходства). При этом актуальной является задача

комплексирования разнотипных и разношкальных признаков (метрических, статистических, логических, текстурных, структурно-лингвистических и др.), полученных различными измерительными средствами с целью решения задачи распознавания или обнаружения. Подходы, основанные на параллельно-фрагментной обработке комплексированных данных корреляционным методом, изложены в [4].

Алгоритмы корреляционно-признаковых и признаковых методов достаточно близки, и первые можно рассматривать как частный случай вторых. Хотя признаковые методы и не обладают такой помехозащищенностью, как корреляционные или методы частных корреляций (в сигнальных пространствах), однако, благодаря меньшим трудозатратам, их применение часто полезно на этапе установления неэквивалентности описаний. При этом часть входных изображений, признаки которых не соответствуют ни одному из эталонов, отбраковываются.

1.2 Построение инвариантных структурных описаний

В большинстве задач идентификации объектов одним из важнейших признаков является форма объектов. Основной задачей сопоставления описаний геометрических объектов является разработка информативного представления формы, которое должно удовлетворять ряду требований: аффинная инвариантность, устойчивость, компактность, вычислительная простота извлечения и возможность оценить сходство [80]. Немаловажное значение имеет способность признака отражать как общее строение формы, так и локальные свойства объекта.

К группе геометрических признаков относят признаки, основанные на использовании геометрических характеристик видео-объектов. Например, геометрические размеры объекта по вертикали или горизонтали, расстояние между наиболее удаленными точками объекта, периметр, площадь и соотношение их на объекте, числовые характеристики описанных или

вписанных в изображение объекта фигур – окружностей, многоугольников и т.п. Особенно важными для результата структурной классификации видео-объектов есть признаки, отражающие геометрические соотношения характерных точек внутри объекта. К таким признакам можно отнести аффинные инварианты [1, 4, 5].

Достаточно специфические, но широко используемые геометрические признаки связаны с представлением геометрии контура объекта. Контур является одной из важных характеристик объекта в восприятии человеком. Поэтому описание контура – одна из распространенных задач в обработке изображений. К наиболее типичным методам описания контура относятся следующие.

Метод цепного кодирования для представления контура объекта заключается в том, что границу объекта, расположенного на дискретной сетке, представить в виде множества элементарных отрезков. Тогда полной характеристикой границы в каждой точке является направление требуемого отрезка. Достоинством представления границы объекта цепным кодом есть простота реализации алгоритма его описания и получения некоторых других геометрических характеристик (периметр, площадь, линейные размеры по вертикали и горизонтали) объекта на основе этого описания, возможность достижения инвариантности описания к преобразованиям подобия – масштабированию, переносу и повороту изображения. Основным недостатком является высокая неустойчивость получаемых описаний к искажениям в изображениях.

Ортогональные представления кривизны контура связаны с углом наклона к касательной контура в виде периодической функции $\kappa(s) = \kappa(x_1(s), x_2(s))$, где представление контура получается в виде ряда с коэффициентами

$$c_k = \frac{1}{L} \int_0^L \kappa(s) \varphi_k(s) ds, \quad (1.2)$$

где $\{\varphi_k(s)\}_{k=0}^{K-1}$ – множество ортонормированных или ортогональных функций на $[0, L]$,

$$\frac{1}{L} \int_0^L \varphi_k(s) \varphi_j(s) ds = \lambda_k \delta(k - j),$$

где L – периметр объекта.

В качестве функций $\varphi_k(s)$ чаще всего используют комплексную экспоненциальную функцию:

$$\varphi_k(s) = \exp\left(-i \frac{2\pi}{L} s \cdot k\right). \quad (1.3)$$

В этом случае говорят о представлении контура в виде ряда Фурье, а коэффициенты λ_k называют фурье-дискрипторами контура.

Описание (1.2) обладает свойствами инвариантности к преобразованиям подобия. Для его использования необходимо получить промежуточное представление контура в виде функции $\kappa(s)$, которое само является его характеристикой. Это может быть сделано различными способами, в частности с помощью ψ -кривых.

Представление контура с помощью ψ -кривой основано на аппроксимации прямолинейных участков границы объекта в виде отрезков ломаных, в области изменения направления границы – в виде дуг окружностей. В этом случае функция $\kappa(s)$ представлена прямолинейными горизонтальными участками и короткими кривыми, соответствующими областям изменения направления границы. При надлежащем построении ψ -кривых можно также добиться инвариантности к преобразованиям подобия.

Аппарат сплайнов достаточно широко используется для практических приложений, например для решения задачи описания границы [113]. Суть метода сплайн аппроксимации – представления границы объекта в виде кусочно-полиномиальной функции часто с выполнением ряда ограничений, накладываемых на ее гладкость. Наиболее популярными являются сплайны третьего порядка, поскольку обладают наименьшей степенью, при которой полиномиальная функция может менять знак кривизны. Это позволяет добиться

определенной гладкости в точках соединения сплайнов за счет равенства первых производных в этих точках.

Распознавание на основе контурных методов обладает высокой степенью устойчивости, но затруднительным оказывается поиск при влиянии пространственных преобразований сложного типа и наличии большого числа ложных контуров на объекте.

Аффинная геометрия на плоскости играет важную роль для анализа изображений, так как является самой общей геометрией, в которой еще имеет смысл понятие прямой. Преобразование называется аффинным, если любые три неколлинеарные точки оно переводит в три неколлинеарные точки. Аффинное преобразование (АП) сохраняет параллельность прямых, отображает скрещивающиеся прямые на скрещивающиеся прямые. Множество АП является группой, имеющей своей подгруппой группу преобразований подобия, любая композиция последовательно выполняемых АП также является АП [46].

Аффинное преобразование Π описывается парой (T, t) и определяется выражением:

$$\Pi(x) = Tx + t, \quad (1.4)$$

где x – произвольная точка изображения.

На рис. 1.2 приведен пример действия АП для точек объекта. АП (1.4) может быть представлено в координатной форме:

$$\begin{cases} y_1 = t_1 + t_{11}x_1 + t_{12}x_2, \\ y_2 = t_2 + t_{21}x_1 + t_{22}x_2, \end{cases} \quad (1.5)$$

где $T = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}$ – невырожденная матрица центроаффинного преобразования,

$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ – координатные столбцы образа и прообраза

соответственно,

$t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$ – столбец вектора переноса начала координат.

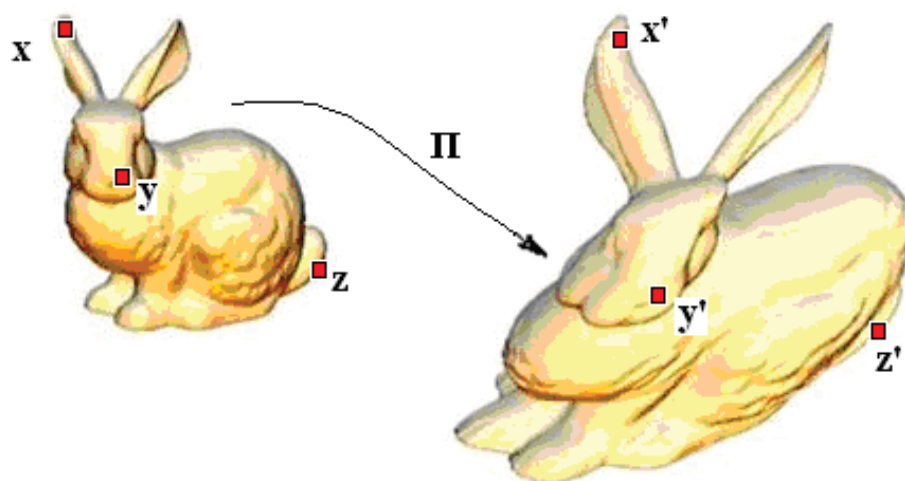


Рисунок 1.2 – Пример аффинного преобразования точек объектов

АП можно охарактеризовать таким свойством: прямые отображаются в прямые, параллельные прямые в параллельные прямые и сохраняются отношения с соответствующим знаком длин отрезков на параллельных прямых [5, 44]. Группа АП включает в себя такие преобразования, как перенос, поворот, масштабирование, сдвиг. Масштабирование, перенос и поворот часто называют преобразованиями подобия. Теория нормализации АП и распознавания изображений в условиях ГП развита в работах профессора Путятин Е.П. и его учеников [1, 3, 12, 20, 80].

В условиях действия группы ГП на анализируемое изображение понятие соответствия эталону преобразуется в математическое понятие эквивалентности множества изображений относительно действия заданной группы [1]. Изображения B_1, B_2 считаются эквивалентными относительно группы G , если существует преобразование $g \in G$, которое переводит B_1 в B_2 , $gB_1 = B_2$. Пусть $E_j \in W$ – эталонное изображение из класса j , а $E = \{E_j\}$, $E \subset W$ – множество всех эталонов. В таком случае все изображения gE_j , полученные путем действия элементов $g \in G$ на эталон E_j , являются эквивалентными и образуют класс эквивалентности $E^j = \{gE_j, g \in G\}$, причем различные классы E^j в силу свойств группы не пересекаются, т.е. выполняется $E^j \cap E^i = \emptyset$, $j \neq i$.

В соответствии с теорией нормализации множество изображений W может быть представлено как объединение классов эквивалентности [1]. Классификация сводится к установлению класса эквивалентности путем отображения из множества объектов в множество эталонов.

Преимущество использования инвариантных признаков заключается в существенном повышении быстродействия классификации. Геометрические инварианты в изложении [5] – это признаки формы объекта, которые остаются неизменными при воздействии ГП. Пусть V, U – множества признаков, отображение $\Theta: V \rightarrow U$ инвариантно относительно группы ГП G , если для него справедливо $\Theta(gB) = \Theta(B)$, $g \in G$, B – описание изображения. Другими словами, инвариантные признаки несут информацию, независимую от поля зрения объекта и изменяющуюся для различных объектов [5]. Основная идея применения инвариантов состоит в хранении в базе данных не описаний объектов, а множество производных признаков от них, которые не изменяются при воздействии преобразований. Признаками могут быть числа, кривые или другие величины. Каждый эталон в базе данных будет иметь соответствующее множество таких инвариантных признаков. На неизвестном изображении также извлекаются инвариантные признаки для последующего сравнения с множеством эталонных признаков.

Инвариантные признаки играют важную роль в распознавании объектов, так как позволяют получить описание, независимое от точки обзора и других геометрических преобразований. Переход к множеству инвариантов снижает сложность поиска и увеличивает на несколько порядков эффективность процесса классификации. Также увеличивается количество объектов, которые могут храниться в базе данных, потому что не нужно хранить одинаковые объекты, отличающиеся только ГП.

Инварианты ГП были впервые введены в компьютерном зрении Вайсом [78]. Особый интерес в практике компьютерного зрения с точки зрения универсальности вызывают проективные инварианты в виде двойного отношения [8]. Однако применение самых общих проективных инвариантов

эффективно и строго выполняется только в идеальном случае, в практической реализации координаты точек подвержены изменениям и на их основе сложно достичь точных инвариантных соотношений для сравниваемых изображений.

1.3 Сравнительный анализ моделей установления эквивалентности структурных описаний

Структурная классификация рассматривается как отображение СЭ эталона на элементы неизвестного объекта. Основная задача сопоставления связана с определением, содержит ли анализируемое изображение образ данного эталона. В случае успешного обнаружения могут быть вычислены параметры расположения объекта относительно камеры [5].

Проанализируем в сравнительном аспекте методы, которые применяются для двумерного сопоставления в пространстве ХП.

Метод базовых локальных признаков (local-feature-focus method, МБЛП) предполагает анализ локальных ХП объекта и их взаимного расположения. МБЛП состоит из двух основных этапов – предварительной обработки и сопоставления описаний. На этапе предварительной обработки формируется множество эталонных описаний, по одному эталону для каждого распознаваемого объекта. Для каждого эталона выделяется множество ЛП. В классическом описании МБЛП [43] в качестве признаков используется множество ограниченных областей, представленных последовательностью отрезков линий и площадями окружностей. Структура описания представляет собой множество кластеров вокруг базовых признаков (focus features, БП). Кластеры формируются как множество ЛП из ближайшей окрестности БП и однозначно определяют ориентацию и расположение БП на объекте для предопределения эффективной стратегии поиска для каждого БП. На практике структура каждого вторичного признака содержит тип признака, расстояние к базовому признаку и список значений возможных полных совпадений для признака. Пусть эталон E содержит множество БП $U = \{u_i\}$, $i = \overline{1, N}$,

N – число БП на эталоне, $S(u_i)$ – множество вторичных близкорасположенных признаков для каждого u_i .

Процедура сопоставления состоит из нескольких этапов [4, 101]. На исходном изображении V выделяется множество ЛП. Затем формируется множество БП $V = \{v_j\}$, $j = \overline{1, n}$, где n – число БП изображения. Выполняется поиск вначале БП, при обнаружении базового признака v_j , соответствующего БП u_i из эталона, производится попытка обнаружить в окрестности базового признака множество $S(v_j)$ близкорасположенных признаков. Отбор признаков осуществляется в соответствии с определенными заранее критериями. Следующими этапами процедуры сопоставления являются: преобразование множества БП эталона и изображения S_i, S_j соответственно; поиск максимальной клики – наибольшего числа соответствий кандидатов; формирование гипотезы о присутствии объекта на эталоне. Наибольшее число максимальных клик не гарантирует совпадения объекта с эталоном, поэтому выполняется процедура подтверждения гипотезы путем вычисления параметров АП, отображающего признаки из S_i на соответствующие признаки из S_j , и применения этих параметров к выбранному эталону (см. рис. 1.3).

Известными недостатками методов с использованием клик является их высокие вычислительные затраты.

Эффективность метода непосредственно зависит от надежности сформированного множества БП и их структурных отличий в объекте.

Метод кластеризации по расположению использует низкоуровневые ХП, от которых переходят в пространство значений геометрических признаков. Для совмещения эталона и объекта вычисляется АП для всех возможных пар управляющих точек. Затем на множестве параметров преобразований выполняется поиск кластера, содержащего параметры с близкими значениями. Кластер в пространстве параметров удастся обнаружить, если эталон и объект имеют большое количество сопоставленных признаков [4].

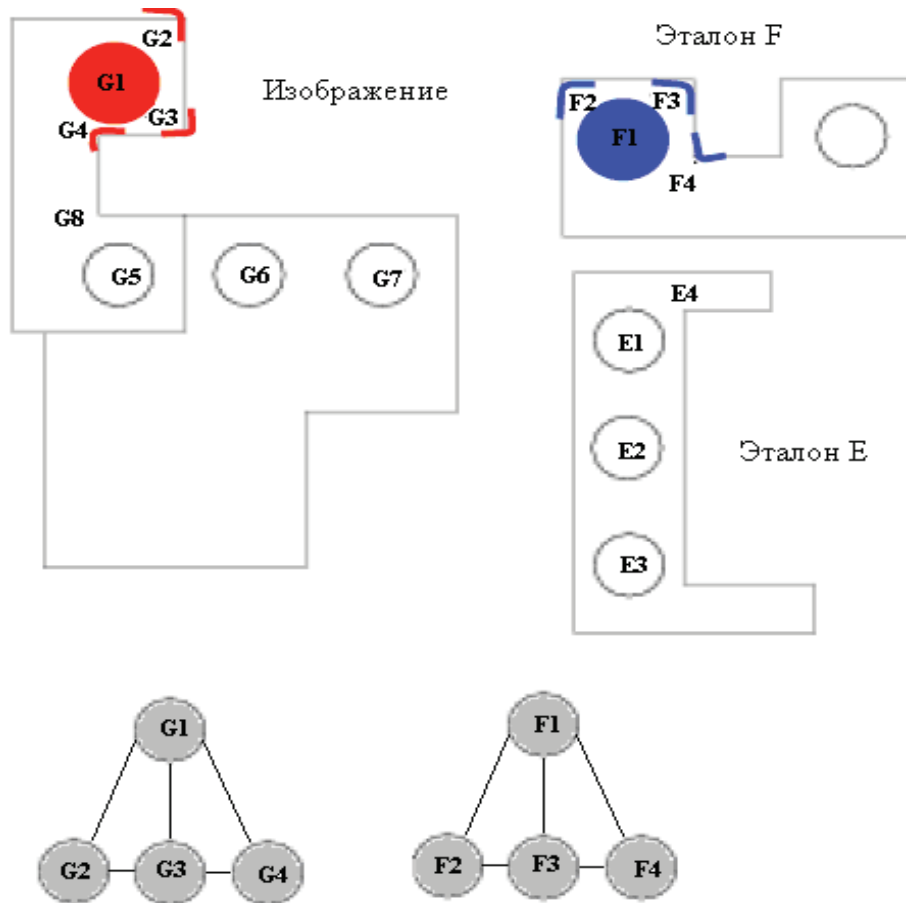


Рисунок 1.3 – Пояснение понятия «лучшее соответствие»

Множество пар соответствий точек объекта и эталона запишем в виде: $d = \{(p_i, P_i) | i \in S\}$, p_i – признаки эталона, P_i – признаки объекта, $S = \{1, \dots, m\}$ – индексы сопоставленных пар. Пусть f_i – параметры геометрического преобразования p_i в P_i , будем считать, что множество $f = \{f_1, \dots, f_m\}$ – это конфигурация “позы”. В случае отсутствия воздействия помех для идеального соответствия имеем m уравнений, которые преобразуют признаки эталона в признаки объекта, т.е. выполняется: $P_i = f_i(p_i)$, $i \in S$. Цель сопоставления заключается в нахождении оптимального расположения:

$$f^* = \arg \min_f U(f | d).$$

Предположим, что каждый f_i , принадлежащий определенному классу ℓ из преобразования позы, такой, что допустимое пространство поз есть $F = \ell^m$. Это налагает ограничения на параметры f_i . Необходимое количество

параметров преобразования зависит от класса преобразований и принятых представлений о позе объекта [88].

При фильтрации признаков по типу точность результатов повышается, а временные затраты сокращаются. Для каждого текущего набора параметров a определяется количество других наборов a_i , которые близки к текущему набору по некоторому критерию расстояния.

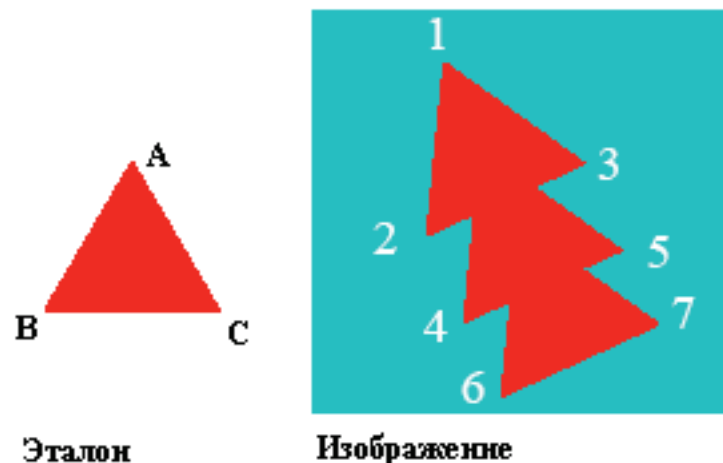


Рисунок 1.4 – Пример установления присутствия эталона на изображении

Альтернативный вариант организации вычислений предполагает применение дискретизации. При этом обеспечивается более быстрая обработка, но ее сложнее адаптировать к конкретной задаче. Каждый набор параметров учитывается в дискретном элементе параметрического пространства. После обработки всех наборов выполняется поиск существенных значений счетчиков. Если несколько похожих a_i попадают в различные соседние разряды, то кластер может быть потерян.

Метод базовых локальных признаков и кластеризации по расположению были разработаны для сопоставления одного эталона и объекта [4]. Для решения задачи классификации (нескольких эталонов) в этих подходах требуется сопоставлять все эталоны по очереди, что часто неприемлемо из-за значительного объема вычислений. Для работы с большими базами

видеоданных разработан метод геометрического хеширования (ГХ). Метод ГХ требует большой предварительной обработки, но благодаря этому обеспечивается быстрое распознавание. ГХ разработан в компьютерном зрении для сравнения геометрических характеристик в базе данных и находит применение в ряде других областей. В методе ГХ для голосования за классы используются геометрические инварианты [4, 5, 27, 69, 70] и поиск соответствия определенному классу осуществляется в условиях воздействия преобразований или частичного перекрытия объектов. ГХ по оценкам исследователей имеет высокую эффективность и низкую полиномиальную сложность [27].

Процедура сопоставления описаний на основе ГХ предполагает два основных шага. На первом шаге осуществляется предварительная обработка базы эталонов, формируется хеш-таблица описаний. Описания в хеш-таблице представлены в виде значений признаков, инвариантных к геометрическим признакам. Этот шаг требует больших вычислительных затрат. На втором шаге, непосредственно при сопоставлении, выполняется быстрый поиск соответствия объекта эталону за счет использования предварительно сформированной хеш-таблицы.

В качестве аналогии можно привести кластеризацию поз, где также используется голосование, но теперь голоса отдаются за геометрию, а не за позу. Основной принцип ГХ изначально был предложен для неоткалиброванных аффинных проекций двумерных изображений.

Формируется хеш-таблица, в качестве индексов которой используются значения $\alpha = (\zeta, \eta)$, инвариантные к аффинным преобразованиям [27]. Для каждого эталона из базы и для каждой группы из трех неколлинеарных точек – аффинного базиса (АБ) в этом эталоне вычисляются значения α для всех остальных точек множества ХП. Используя α как индексы, в таблицу вводится позиция, в которой записано номер эталона и АБ, на основе которого получены вычисленные значения. Величины α и значение АБ выступают как гипотетические значения, характеризующие j -й эталон.

После формирования хеш-таблицы для базы эталонов можно выполнять поиск. Для исследуемого объекта берется любая тройка неколлинеарных точек, и вычисляются значения α для всех остальных точек. Согласно индексам α восстанавливается содержимое таблицы. Если АБ эталона соответствует АБ объекта, то за комбинацию j -го эталона и его АБ должно быть получено много голосов. Данный подход может применяться и для других геометрических групп, отличных от точек. Подход может быть эффективным, если в базе эталонов существует мало групп характерных элементов с одинаковыми значениями инвариантов. При сопоставлении объектов с помощью ГХ участвуют все группы ХП эталона, поскольку неизвестно какая группа, принадлежащая объекту, исследуется.

Модели сопоставления описаний на основе голосования анализируют «близость» между элементами эталонных описаний и описанием объекта, который необходимо классифицировать. Наличие близости между элементами описаний служит частичным прецедентом (голосом) и оценивается по некоторому правилу посредством числовой оценки. По набору голосов за близость вырабатывается общая оценка классифицируемого объекта, которая и является значением функции принадлежности классу [22, 103]. Суть голосования сводится к простой схеме, в которой каждый признак, выделенный на изображении, голосует за ближайший к нему класс. Преимущество отдается классу с наибольшим количеством сформированных голосов и считается, что объект принадлежит данному классу. Учитывая, что голосование опирается исключительно на эталонные описания, оно может иметь лучшие характеристики достоверности, нежели метрическое сопоставление двух описаний в виде множеств.

Формально функцию голосования элемента v можно представить в виде:

$$\text{vote}(v, j) = \begin{cases} 1, & \rho(v, v_j) \leq \varepsilon, \\ 0, & \rho(v, v_j) > \varepsilon, \end{cases} \quad (1.6)$$

где v, v_j – ХП объекта и j -го эталона соответственно, $j \in J$,

$\rho(v, v_j)$ – расстояние между v, v_j ,

ε – порог значимости.

На основе (1.6) формируется целевая функция для кандидата с номером j :

$$\text{votes}(j) = \sum_{v \in V} \text{vote}(v, j), \quad (1.7)$$

где V – множество ХП объекта.

Голосование – достаточно эффективная практическая стратегия. Проблемы могут возникнуть в случае, когда признаки различных классов имеют близкие значения и принятие решения на основе максимума голосов может быть затруднено из-за наличия нескольких кандидатов с большим числом голосов (рис. 1.5). В процесс голосования могут быть включены как непосредственно ХП, так и их отношения, применение которых формирует многоуровневые признаки.

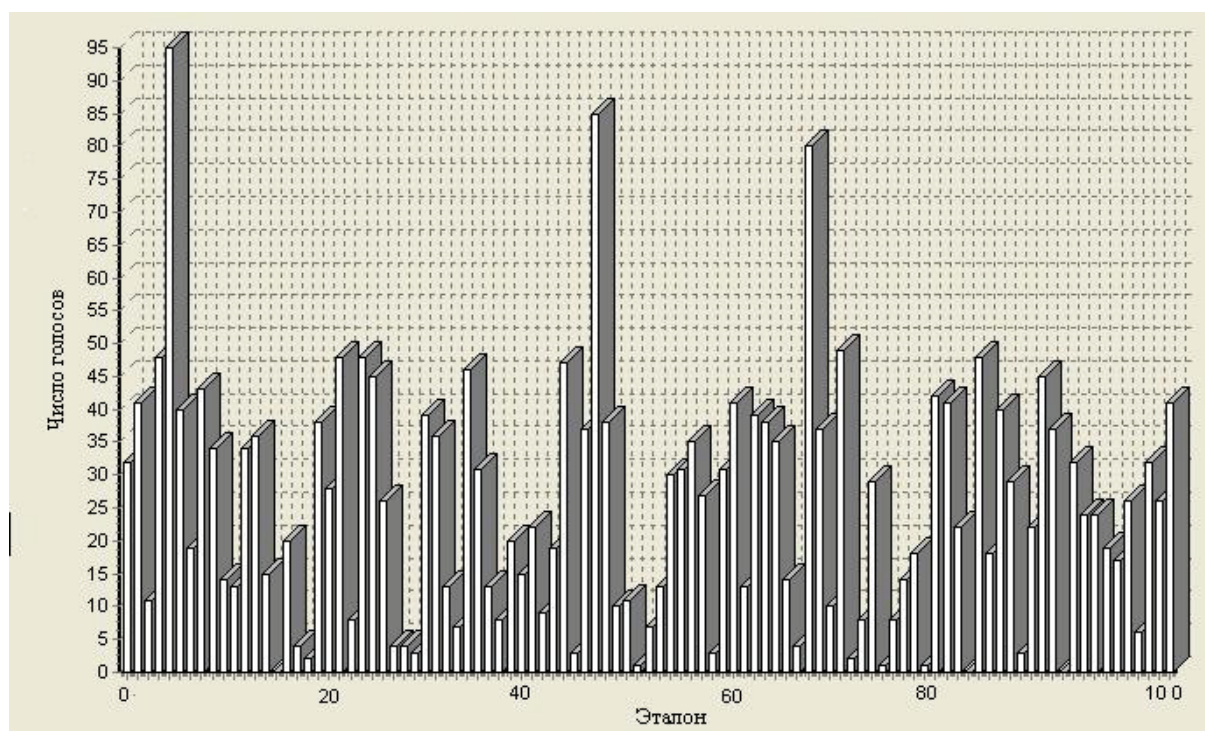


Рисунок 1.5 – Результат голосования как число голосов эталона

Основная идея, которая отличает разработанные в монографии методы, состоит в непосредственном использовании геометрических конструкций, построенных на координатных соотношениях ХП внутри структурного описания.

1.4 Постановка задач исследований

Характер представлений о строении множеств элементов описаний изображений в значительной мере влияет на эффективность методов интеллектуального анализа визуальных объектов. Повышение показателей надежности и быстродействия системы классификации может достигаться за счет использования новых структур на множестве признаков, устанавливающих определенные отношения между входящими в него элементами. Важным свойством для качественной классификации является инвариантность структурных признаков к геометрическим преобразованиям, так как они позволяют получить описание, не зависящее от точки обзора и других геометрических преобразований, за счет чего снижается сложность поиска и увеличивается эффективность процесса классификации.

Для эффективного решения важной научно-прикладной задачи классификации изображений в условиях геометрических преобразований и помех необходимо выполнить следующие задачи:

1) синтезировать метод сопоставления структурных описаний видео-объектов на основе построения моделей инвариантных признаков структур для повышения достоверности классификации за счет учета контекста и пространственных свойств элементов описания;

2) разработать модели построения описания видео-объектов путем применения пространственных структур и процедуры целенаправленного сжатия описаний для управления объемом описания, снижения вычислительной сложности и обеспечения надежности процесса классификации;

3) разработать гибридные модели классификации описаний на основе анализа разнотипных признаков с целью повышения достоверности классификации за счет многостороннего и углубленного анализа состава описаний;

4) изучить и применить модель установления максимального паросочетания между элементами описаний с учетом их взаимно-однозначного соответствия, основанную на принципах формирования оптимальных соответствий, которая обеспечивает минимум значения критерия суммарного расстояния между элементами описаний для эффективного вычисления подобия;

5) провести экспериментальные исследования, направленные на проверку эффективности разработанных методов и моделей с целью определения рациональных путей их реализации и установления особенностей практического использования.

Выводы

1. Анализ современного состояния проблемы классификации изображений на основе структурных описаний объектов в условиях геометрических преобразований и помех показали необходимость разработки новых эффективных методов структурного анализа, а также практическую необходимость снижения вычислительной сложности процедур сопоставления в многомерном пространстве параметров.

2. Прикладной интерес при сопоставлении структурных описаний представляют геометрические признаки, которые обладают способностью отражать как общее строение формы, так и локальные свойства объекта. Перспективным направлением является развитие теории использования инвариантных пространственных признаков, что в целом обуславливает снижение сложности классификации.

3. Наиболее эффективными в настоящее время моделями установления эквивалентности описаний являются метод базовых локальных признаков, кластеризация по расположению, геометрическое хеширование и модели на основе голосования, которые используются для работы с прикладными базами изображений. Модели голосования целесообразно развивать в плане создания эффективных процедур формирования гипотезы о принадлежности объекта с использованием пространственных геометрических конструкций, построенных на координатных соотношениях ХП внутри описания.

4. Сформулированы задачи исследования, связанные с развитием новых эффективных подходов к формированию признаковых пространств на основе построения и применения пространственных признаковых структур.

2 ПОСТРОЕНИЕ И СОПОСТАВЛЕНИЕ ОПИСАНИЙ ВИЗУАЛЬНЫХ ОБЪЕКТОВ В ВИДЕ МНОЖЕСТВ ПРИЗНАКОВЫХ ПРОСТРАНСТВЕННЫХ СТРУКТУР

Основным преимуществом рассмотренных методов построения, анализа и сопоставления признаков пространственных структур в сравнении с традиционными подходами является тот факт, что использование пространственной информации напрямую увеличивает достоверность распознавания изображений объектов. Например, если элементы описания перемешать произвольным образом, изменив их координаты на плоскости, то голосование традиционным методом не изменит величину подобия. В то же время построение признаков пространственных структур для сравниваемых описаний в этой ситуации приведет к существенно иному значению меры подобия [53]. Настоящий раздел работы содержит обоснования и формализацию построения описаний в виде пространственных структур.

2.1 Особенности построения современных способов для формирования систем характерных признаков

В процессе применения структурных методов анализа объектов на изображениях осуществляется решение двух основных задач [1, 4, 5, 8]. Для получения структурного представления производится обнаружение структурных элементов в виде ХП, которые отражают степень изменчивости (кривизну) функции яркости и уровень отличия фрагмента с особым признаком от других фрагментов. На изображении данные изменения двумерного сигнала соответствуют контрастным и изолированным точкам, углам, перепадам, концам линий или кривых.

Рассмотрим суть построения и свойства структурных описаний, полученных современными технологиями [16, 18, 62].

Технология SIFT (масштабно-инвариантное преобразование признаков) [16] основана на принципах локальной обработки, включающих фильтрацию, формирование значимых признаков, анализ дискретного пространства преобразований, аппроксимацию координат признаков и др. Основные применения методы SIFT нашли в задачах распознавания двумерных объектов и реконструкции трехмерных сцен [16].

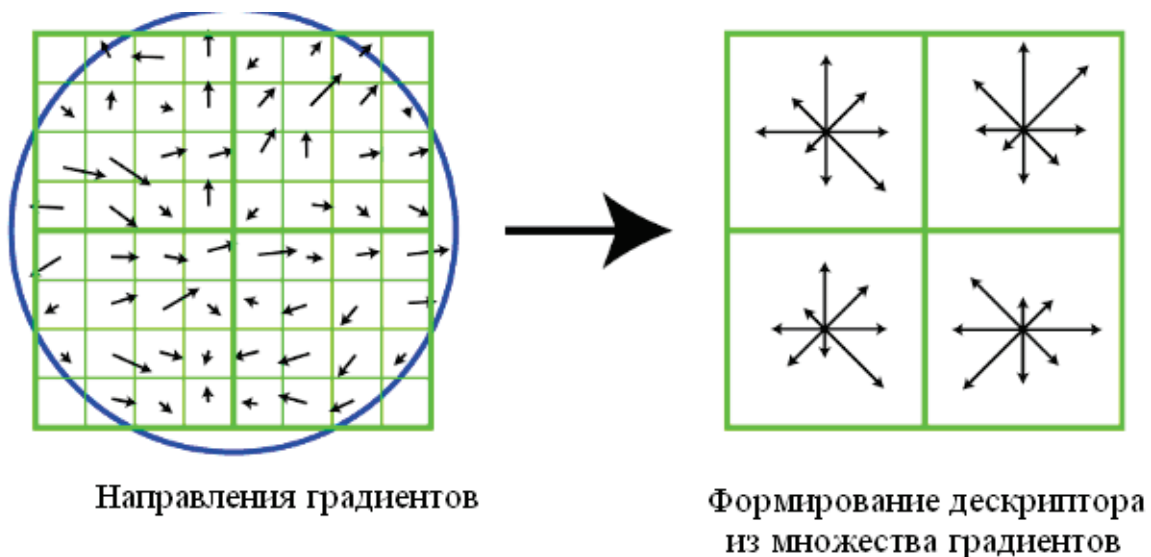


Рисунок 2.1 – Фрагмент изображения (слева) и полученный на ее основе дескриптор признаков (справа)

Главный акцент в построении SIFT сделан на устойчивом определении ключевых точек изображения в плане формирования их признаков путем отделения от других структурных элементов. В процессе многоэтапной обработки путем усиления различий в признаковом пространстве достигается необходимая величина расстояния между структурными элементами объектов, что в целом приводит к высоким характеристикам распознавания. По этим причинам представляется перспективным применение SIFT для распознавания изображений частично видимых объектов.

Построение масштабируемого пространства сводится к тому, что исходное изображение $B(x, y)$ преобразуется к виду множества функций $L(x, y, k\sigma) = B(x, y) \otimes G(x, y, k\sigma)$, каждый элемент которого получается

путем свертки $\otimes$ изображения с маской гауссиана $G(x, y, \sigma)$ (параметр σ). Величина σ изменяется дискретно в пределах отрезка $[\sigma_0, \sigma_1]$.

Далее на базе системы функций L формируется множество DoG-изображений $D(x, y, \sigma)$ вида $D_i(x, y, \sigma) = L(x, y, k_{i+1}\sigma) - L(x, y, k_i\sigma)$. Изображение DoG представляет собой разность двух гауссово-сглаженных изображений L с масштабами $k_{i+1}\sigma$ и $k_i\sigma$, где параметры k_i, k_{i+1} выбираются, смежными внутри октавы. Разность двух изображений, полученных путем сглаживания гауссианами с разными величинами $k_i\sigma$, трактуется как аппроксимация действия оператора лапласиана.

По набору DoG-изображений формируется первоначальное множество ХП. Точка $z^* = (x^*, y^*, \sigma)$ считается кандидатом в список ХП, если она находится в центре и является максимумом или минимумом в своей окрестности $3 \times 3 \times 3$, включающей 8 точек при конкретном масштабе σ и 2×9 точек ближайших «слоев» сверху и снизу, задаваемых соседними масштабами.

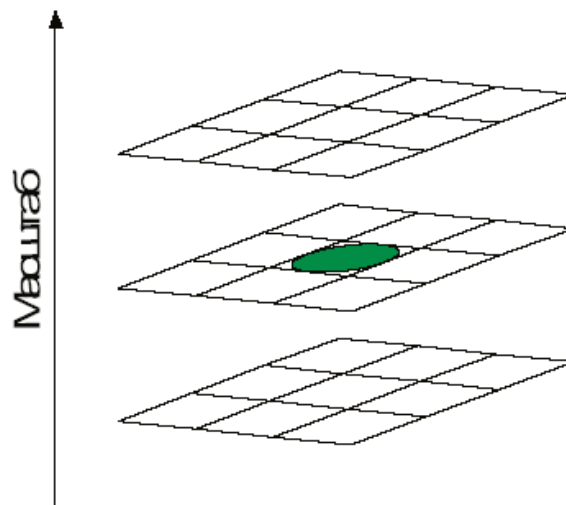


Рисунок 2.2 – Схема поиска точек экстремума

Следующий этап включает отбор устойчивых точек путем детального анализа свойств близлежащих по координатам данных в построенном пространстве функций $D(x, y, \sigma)$ на основе интерполяции. Для определения точного положения ХП осуществляют разложение $D(x, y, \sigma)$ в ряд Тейлора в точках с координатами ХП

$$D(c) = D + \frac{\partial D^T}{\partial c} c + \frac{1}{2} c^T \frac{\partial^2 D}{\partial c^2} c, \quad (2.1)$$

где $c = (x, y)$ – смещение от выбранной ХП z^* в пределах фиксированного масштаба σ . Местоположение $\hat{c}$ экстремума функции $D(c)$ по смещению определяется путем приравнивания нулю производной от $D(c)$ относительно c . Вычисления выполняются в числовом виде. Если хотя бы одна из компонент смещения $\hat{c}$ превышает значение 0,5, то координаты новой ХП вычисляются с учетом поправки $c^* = c^* + \hat{c}$, и для нее процедура интерполяции вновь повторяется. Далее в целях устранения точек с низкой контрастностью отбрасываются те из них, для которых член второго порядка в разложении (2.1) меньше величины 0,03.

Устранение откликов нестабильных краевых точек осуществляется путем анализа соотношения между главными значениями кривизны в направлениях x, y для каждой выделенной ХП. Определение параметров кривизны сводится к нахождению собственных значений матрицы гессиана

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}, \text{ где } D_{uv} \text{ представляет собой вторую производную функции } D$$

по переменным u, v . Эта производная вычисляется путём сканирования окном фиксированных размеров в окрестности точки (x, y) [16].

Определение направлений и амплитуд градиента осуществляется путем назначения каждому ХП из списка одной или несколько ориентаций (величин углов) на основе локальных значений градиента изображения. Для изображения $L(x, y, \sigma)$ при фиксированном σ вычисляются амплитуда $m(x, y)$ и ориентация $\theta(x, y)$ градиента для ХП с координатами (x, y)

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2},$$

$$\theta(x, y) = \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}. \quad (2.2)$$

На основе множества откликов значений и направлений градиента в окрестности размером 3σ (с центром в ХП) строится гистограмма из 36 столбцов. Каждый столбец охватывает диапазон ориентаций в 10 градусов. Значения гистограммы – суммы градиентов $v = \sum m_q$, соответствующие фиксированному направлению с номером q . Для устранения выбросов гистограмма дополнительно сглаживается по трем точкам и нормируется путем деления на максимум с целью обеспечения устойчивости к изменениям яркости. Пики гистограммы соответствуют доминирующим ориентациям градиента в окрестности ХП.

Признаковая информация содержится в массиве направлений $T = \{t_i\}$, $t_i = (x, y, m, \theta, \sigma)$. В окрестности 16×16 каждой точки при фиксированном масштабе анализируется множество из 16 непересекающихся фрагментов 4×4 , образующих разбиение. Для каждого из фрагментов, в свою очередь, строится восьмистолбцовая гистограмма направлений по переменной θ , охватывающая 360 градусов. Одно направление соответствует диапазону в 40 градусов. Значениями дескриптора являются суммы $v_i = \sum_{q=i} m_q$. Для обеспечения инвариантности к поворотам направления нормализуются. В результате для каждого ХП имеем вектор v (дескриптор), содержащий $16 \times 8 = 128$ значений направлений.

Метод SIFT за счет представления в «развернутом» по множеству значений параметров (масштабы и повороты) пространстве обеспечивает устойчивое вычисление ХП. В SIFT применен корреляционный принцип анализа на основе перебора дискретных значений преобразований и представлении их в виде единого вектора. На рис. 2.3 приведен пример объекта с выделенным множеством ХП.

Технология SURF (Speeded Up Robust Features – ускоренное обнаружение устойчивых признаков) решает аналогичные задачи – поиск ХП изображения и создание дескрипторов – векторов, компоненты которых обладают инвариантностью к преобразованиям смещения, масштаба и вращения

объектов [18]. SURF можно считать некоторой более современной и эффективной реализацией SIFT. В целом методы SURF и SIFT являются близкими и решают одни и те же задачи.

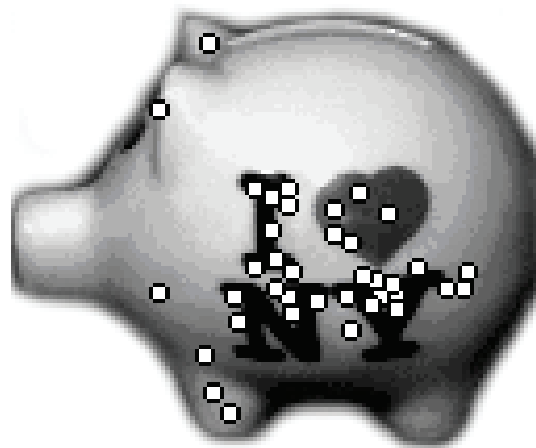


Рисунок 2.3 – Объект из базы COIL-20 с выделенными ХП

Метод SURF считается более быстрым, чем SIFT. Различаются диапазоны значений дескрипторов двух данных методов – у метода SURF значения могут быть как положительные, так и отрицательные, тогда как значения дескрипторов SIFT только положительные. Дескриптор SURF представляет собой вектор, содержащий 64 (или 128 в расширенной версии) числовых компоненты. Метод SURF в настоящее время получил большую популярность, чем SIFT, его реализации есть во многих математических библиотеках, таких как OpenCV, SURFmex, libmv SURF, JOpenSURF, boofCV, Dlib C++ Library [18].

Методы SIFT, SURF используются для поиска объектов на изображении, однако они работают исключительно с сигналом изображения и никак не выделяют объекты из фона. При этом особенности могут быть как внутри объекта, так и на фоне, а также на точках границы. В связи с этим методы недостаточно качественно работают для объектов простой формы и без ярко выраженной текстуры [16, 18]. Внутри таких объектов метод скорее всего не найдет особых точек. Точки будут найдены либо на границе объекта с фоном,

либо вообще только на фоне. А это может привести к тому, что объект иногда не сможет быть распознан в другом изображении на другом фоне.

ASIFT [62] обеспечивает большее количество совпадений на паре изображений, находящихся в разных системах координат – декартовых и сферических (координаты объекта на двумерном изображении и координаты определяемые ориентацией камеры), где SIFT и SURF не всегда дают приемлемые для приложений результаты. Основное отличие ASIFT от технологии SIFT – это объединение процедур симуляции и нормализации окрестности ХП. SIFT нормализует вращения и перемещения, и симулирует все масштабы (уменьшенные фрагменты области ХП). Поэтому масштабирование не влияет только на особые точки. ASIFT симулирует три параметра: масштаб и сферические координаты (longitude angle and the latitude angle), которые эквивалентны наклону, и нормализует другие три (смещения и вращение).

На первом этапе алгоритма ASIFT каждое изображение преобразуется путем имитации возможных аффинных искажений, вызванных изменением оптической ориентации оси камеры относительно фронтальной позиции. Эти искажения зависят от параметров долготы φ и широты θ . Изображения поворачиваются на угол φ , после чего следуют наклоны с параметром

$t = \frac{1}{|\cos \theta|}$ (наклон t в направлении x является операцией вида

$u(x, y) \rightarrow U(tx, y)$. Для цифровых изображений наклон осуществляется созданием t -подвыборки, что обеспечивает нахождение смоделированных изображений вблизи любого другого возможного вида, порожденного другими значениями φ и θ . Далее все смоделированные изображения сравниваются алгоритмом SIFT.

На рис. 2.4 изображения сопоставляемых символов А и В соответствуют объекту и эталону. В процессе применения ASIFT моделирует все искажения, вызванные изменением направления оси оптической камеры, что отражает представление в виде параллелограмма.

Метод ASIFT содержит в качестве главного компонента технологию SIFT, но обеспечивает инвариантность не только к изменению масштаба, поворота и смещения (как классический SIFT), но также и к изменению угла, определяющего оси ориентации камеры, снимающей данный объект. Этот метод дает большое количество совпадений на паре изображений в различных аффинных системах координат, где SIFT не дает приемлемых результатов. К недостаткам ASIFT можно отнести невысокую устойчивость дескрипторов на больших изображениях, а также недостаточную производительность, связанную с увеличением объема вычислений за счет процедуры симуляции всех возможных аффинных искажений изображения.

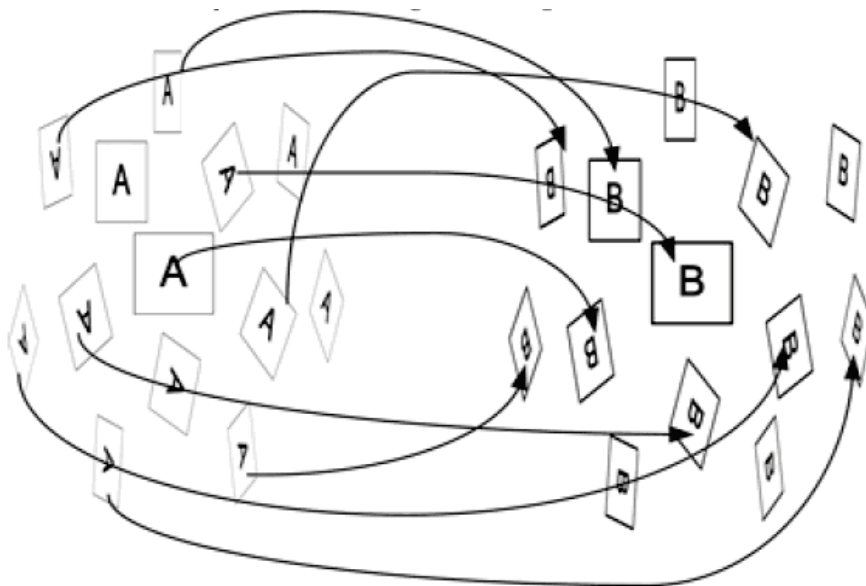


Рисунок 2.4 – Схема применения метода ASIFT

2.2 Построение описаний на основе признаков пространственных структур

Для формирования исходного множества структурных признаков в монографии выбран метод SURF.

Под признаковой пространственной структурой (ППС) в монографии понимается объединение разрозненных структурных элементов описания объекта в виде пространственных конгломератов признаков на основе

формирования между структурными элементами отношений, отражающих геометрические свойства распознаваемого объекта [25, 52].

Описание визуального объекта после выделения ХП формально представлено в виде конечного множества $U = \{u_i\}_{i=1}^n$ элементов u_i – ХП со своими значениями координат x_i, y_i , n – число ХП в описании.

Вначале сосредоточимся на геометрических свойствах двумерных объектов, отложив использование значений векторов ХП для последующего анализа. Отметим, что человеческое зрение редко прибегает к анализу функции яркости визуальных объектов, часто опираясь при решении практических задач исключительно на их геометрические характеристики, которые можно получить непосредственно из множества координат ХП [4, 69, 33].

Рассмотрим в качестве исходных данных множество координат. Известно, что с помощью трех точек на плоскости $u_0, u_1, u_2 \in U$ можно сформировать систему координат и охарактеризовать произвольную точку $u_i \in U$ в этой системе парой чисел $\zeta, \eta \in \mathbb{R}$ в виде:

$$u_i = u_0 + \zeta(u_1 - u_0) + \eta(u_2 - u_0). \quad (2.3)$$

Применив аффинное преобразование $\Pi = \{T, t\}$ (T – центроаффинное преобразование, t – вектор смещений) к описанию U (см. рис. 2.5), получим новое описание $U' = \{u'_i\}_{i=1}^n$ с элементами

$$u'_i = \Pi(u_i) = T u_i + t. \quad (2.4)$$

Связь элементов $u_i \in U$ и $u'_i \in U'$ в соответствующих аффинных системах координат можно представить как

$$\begin{aligned} u'_i &= T u_i + t = T(u_0 + \zeta(u_1 - u_0) + \eta(u_2 - u_0)) + t = \\ &= (1 - \zeta - \eta)(T u_0 + t) + \zeta(T u_1 + t) + \eta(T u_2 + t) = \\ &= (1 - \zeta - \eta)u'_0 + \zeta u'_1 + \eta u'_2 = u'_0 + \zeta(u'_1 - u'_0) + \eta(u'_2 - u'_0). \end{aligned}$$

Это показывает, что значения ζ, η не зависят от воздействия АП и они могут быть использованы как инвариантные признаки для сопоставления описаний в условиях действия АП. Назовем пару значений ζ, η геометрическими инвариантными признаками (ГИП). Множество значений

ГИП и их взаимосвязи в пространстве ГИП являются основой для формирования разновидностей ППС.

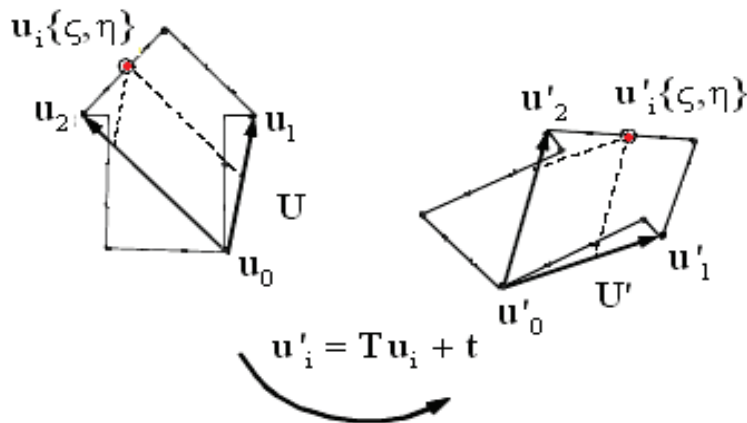


Рисунок 2.5 – Схема применения АП для описаний из множеств ХП

В общем случае при перестановке порядка элементов базиса u_0, u_1, u_2 для одного и того же ХП значения ГИП (ζ, η) изменяются, число таких пар для одного и того же ХП равно 6.

Если дополнительно рассмотреть параметр γ такой, что

$$\gamma + \zeta + \eta = 1, \quad (2.5)$$

то представление u_i можно записать в виде [4, 5, 25, 33, 72]:

$$u_i = \gamma u_0 + \zeta u_1 + \eta u_2. \quad (2.6)$$

Значения γ, ζ, η однозначно определяют расположение ХП u_i относительно базиса u_0, u_1, u_2 и называются барицентрическими координатами (БК) (рис. 2.6) [44].

Значение каждой из величин γ, ζ, η является отношением расстояния точки до одной из сторон к высоте, опущенной на эту сторону из противоположной вершины [26]. Из этого утверждения следует то, что БК принимают постоянные значения на линиях, параллельных соответствующей стороне, например, ζ принимает постоянные значения от 0 до 1 на линиях, выделенных пунктиром на рис 2.7. Это свойство БК можно использовать для построения структур на множестве ГИП или в процедурах сжатия описания [47, 53].

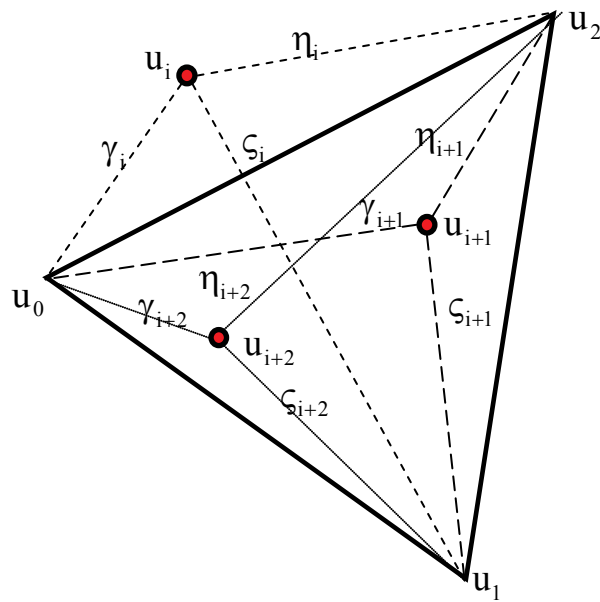


Рисунок 2.6 – Геометрическая интерпретация БК
для множества точек относительно АБ (u_0, u_1, u_2)

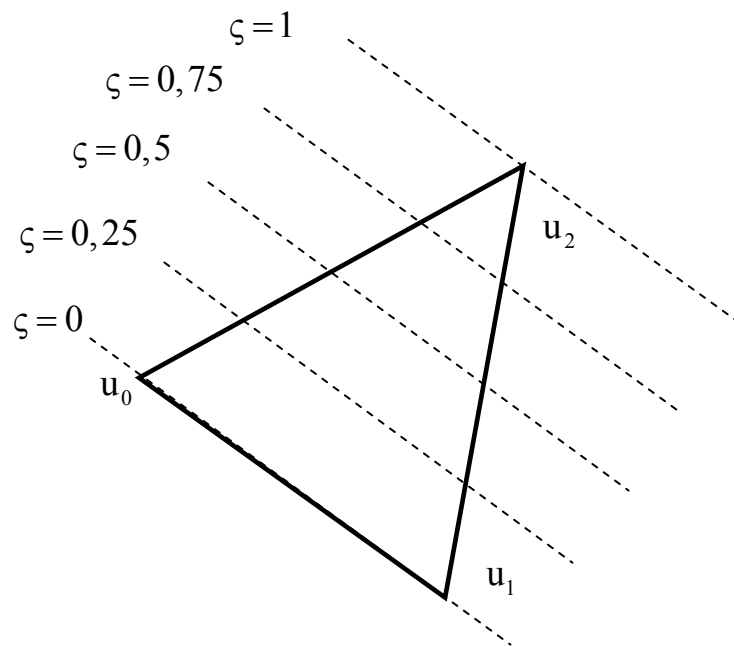


Рисунок 2.7 – Линии константных значений координаты ζ
относительно базиса (u_0, u_1, u_2)

Известны следующие свойства БК [116]. Если точка P находится внутри координатного треугольника-базиса, то все три её барицентрические координаты имеют один знак (их можно считать положительными). Если точка

P – на какой-либо стороне координатного треугольника или на её продолжении, то хотя бы одна барицентрическая координата этой точки равна нулю. В остальных случаях две координаты точки P – одного знака, а третья имеет противоположный знак.

В физическом смысле точка u_i представляет собой центр тяжести системы из трёх материальных точек с массами γ, ς, η , расположенных соответственно в пространственных точках u_0, u_1, u_2 . Коэффициенты γ, ς, η могут быть как положительными, так и отрицательными. БК нашли применение в химии при решении проблем аналитического описания растворов, сплавов, различных химических соединений. Их сводят к задачам о нахождении центров масс для систем материальных точек, а БК используются как степень концентрации различных трехкомпонентных смесей. Известно также применение БК в колориметрии в качестве координат цветности точки в системе RGB. В популяционной генетике БК используют, например, для определения вероятности формы генов признаков случайно взятой особи из популяции [44].

Для двумерного случая координаты точек ХП и базиса удовлетворяют уравнению в матричной форме:

$$\begin{bmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \gamma \\ \varsigma \\ \eta \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}. \quad (2.7)$$

Используя, например, правило Крамера для решения системы линейных уравнений, легко вычислить значения аффинных координат для ХП u_i при условии неколлинеарности точек АБ (u_0, u_1, u_2):

$$\varsigma = \frac{\begin{vmatrix} x_0 & x_i & x_2 \\ y_0 & y_i & y_2 \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ 1 & 1 & 1 \end{vmatrix}}; \quad \eta = \frac{\begin{vmatrix} x_0 & x_1 & x_i \\ y_0 & y_1 & y_i \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ 1 & 1 & 1 \end{vmatrix}}; \quad \gamma = \frac{\begin{vmatrix} x_i & x_1 & x_2 \\ y_i & y_1 & y_2 \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ 1 & 1 & 1 \end{vmatrix}} = 1 - \varsigma - \eta, \quad (2.8)$$

где $\| \|$ – символ определителя.

Отметим важное свойство, что координаты γ, ζ, η отображают геометрические особенности объекта, т.к. получены на основе геометрической информации (координат), и являются вторичными геометрическими инвариантными признаками объекта. Графическое отображение значений γ, ζ, η , полученных для множества ХП в системе координат АБ u_0, u_1, u_2 , позволяет представить компактное и понятное описание структуры взаимосвязи значений признаков из двух разных пространств и судить о взаимном пространственном расположении множества ГИП. Анализ соответствий множества ХП и их проекции (значений γ, ζ, η) в системе АБ показал, что отношения взаимной удаленности ХП друг от друга сохраняются (рис. 2.8).

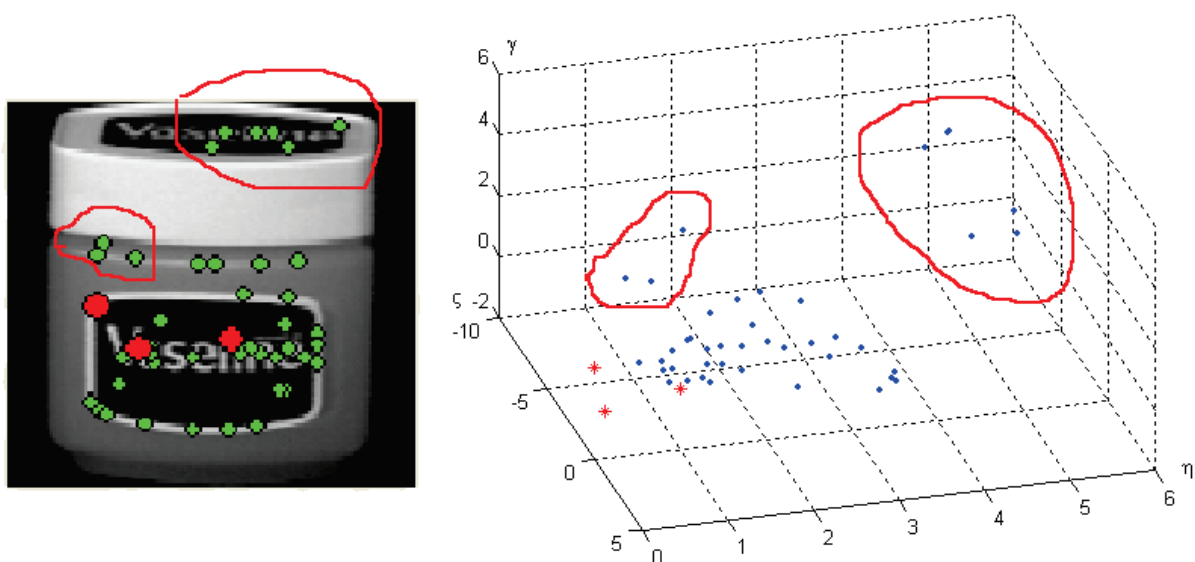


Рисунок 2.8– Множества ХП и их представления

Чтобы однозначно определить ХП $u_i \in U$ относительно АБ, достаточно двух параметров ζ, η , значение γ получаем из уже известных ζ, η . Необходимость хранения и обработки всех трех параметров γ, ζ, η возникает при решении задачи сопоставления описаний на основе ГИП в случае, когда порядок следования ХП объекта неизвестен или часть множества ХП потеряна при воздействии помех [25,53]. Для задачи классификации в пространстве ГИП представление в виде троек γ, ζ, η лучше в вычислительном плане. Выигрыш

объясняется тем, что сокращается объем исходных данных для сопоставления и тем самым повышается быстродействие сопоставления [47].

В результате построения ГИП структура элемента описания приобретает вид:

$$z_i = \langle x_i, y_i, \{\alpha_{qi}\}_{q=1}^{m_i} \rangle, \quad (2.9)$$

где $\alpha_{qi} = (\zeta_{qi}, \eta_{qi}, \gamma_{qi})$ – тройка ГИП для q -го АБ, $i = \overline{1, n}$, m_i – число ГИП для z_i -го ХП. Тогда описание объекта в виде совокупности элементов представлено, как $Z = \{z_i\}_{i=1}^n$.

Предложенную систему признаков $\{\alpha_{qi} = (\zeta_{qi}, \eta_{qi}, \gamma_{qi})\}$ можно рассматривать также с точки зрения иерархического представления [107]. ППС – это признаки второго уровня, построенные на исходном множестве ХП. Основу построения иерархии здесь определяет применение пространственных и атрибутивных отношений на множестве ХП.

В общем виде иерархию $H(Z)$ применительно к описанию Z можно представить как конечное множество $H(Z) = (h_1, \dots, h_m)$, где m – число уровней, $h_k = \varphi(h_{k-1})$, $k > 1$, φ – функция перехода от уровня к уровню, реализующая процесс «группирования» или отбора ХП нижнего уровня [122]. Случай $k = 1$ соответствует исходному описанию. Одним из вариантов построения $H(Z)$ для множества ХП есть формирование уровней h_k исходя непосредственно из множества Z . В таком случае $h_k = \varphi_k(Z)$, и иерархия определяется функциями φ_k . Как правило, процесс сопоставления описаний осуществляется на равноценных уровнях. В результате мера подобия $\vartheta[H(Z_1), H(Z_2)]$ иерархических описаний представляется композицией f результатов сравнения для систем признаков отдельных уровней

$$\vartheta[H(Z_1), H(Z_2)] = f(\vartheta_1[h_1(Z_1), h_1(Z_2)], \dots, \vartheta_m[h_m(Z_1), h_m(Z_2)]). \quad (2.10)$$

Для случая ППС выражение (2.10) вычисляется как подобие между множествами ХП второго уровня, т.е. между $\{\alpha_{qi} = (\zeta_{qi}, \eta_{qi}, \gamma_{qi})\}$, представляющими объект и эталон.

Построение иерархий для случая ППС можно распространить для более высоких уровней. Компонентное представление путем иерархического/стратифицированного описания обладает одновременно свойствами целостности и детализации, и позволяет анализировать объект на соответствующем уровне абстрагирования. Для каждой страты существуют характерные особенности, законы и принципы, отражающие ее свойства при распознавании независимо от других страт. В то же время в рамках иерархии описание каждого последующего уровня основано на данных предыдущего уровня. Анализ и сопоставление описаний может осуществляться как независимо по уровням, так и с использованием обобщенной меры подобия, включающей множество описаний разных уровней. В результате распознаваемый объект представляется как многоэлементная совокупность компонент, для которых вычисляется значение вектора признаков. Выбор способов вычисления дескрипторов и формирования компонент в значительной мере определяет свойства классов объектов в плане их разделения.

Иерархическая модель меры подобия соответствует сопоставлению иерархий $H(Z_1), H(Z_2)$, построенных на множествах Z_1, Z_2 ,

$$\vartheta = \Psi_2[\Psi_1[H(Z_1), H(Z_2)]] \quad (2.11)$$

Варианты иерархии в (2.11) предполагают пространственную (в плане координат) и/или дескрипторную (в плане значений ХП) группировку элементов множеств [5]. Одним из путей есть построение системы глобальных признаков по описаниям Z_1, Z_2 . При этом иерархия может быть построена как на предварительном этапе (например, в виде кластеризации элементов Z_1, Z_2), так и непосредственно в процессе сопоставления путем отбора групп схожих элементов множеств. Применение иерархии в целом усиливает интегральные свойства сопоставления в плане влияния аддитивных помех, уменьшая в то же время возможности по селекции пространственных помех типа фона и наложений.

Элементы α_{q_i} формируются путем перебора всевозможных троек ХП, координаты которых не лежат на одной прямой [58]. Для каждого базиса (тройки) вычисляются $(n - 3)$ значений ГИП. Максимально возможное число базисов из трех неколлинеарных точек равно числу комбинаций C_n^3 . Количество ГИП для конечного числа ХП, равного n , достигает величины

$$m = nC_{n-1}^3 = (n - 3)C_n^3. \quad (2.12)$$

По формуле (2.12) вычислим m_i для отдельных ХП, а общее число m значений ГИП для всего описания запишем как $m = \sum_i m_i$.

На рис.2.9 приведен график зависимости числа m аффинных инвариантов от величины n при условии, что все тройки координат удовлетворяют требованиям к базису. Анализ графика показывает, что число ГИП при $n=35$ и выше резко возрастает с ростом количества точек в структурном описании. Так, при числе ХП, равном 35, количество ГИП достигает значения приблизительно двухсот тысяч [47].

Нелинейный характер роста числа ГИП в зависимости от значения числа n ХП в описании требует применения специальных методов фильтрации множества ХП в целях сокращения описания. Для баз видеоизображений, рассмотренных в работе, необходимое для практической классификации число ХП достигает $n=45$, и, соответственно число ГИП для одного ХП может достигнуть значения $m_i = 8\,000$. Подробно методы сжатия множества ХП и множества ГИП в целом будут рассмотрены и исследованы в последующих разделах.

Методы на базе ХП позволяют анализировать описание объекта на предмет соответствия эталону отдельных частей изображения. Множество ГИП и его подмножества опираются на систему связей между отдельными ХП, что дает возможность детального изучения целостных свойств образа. Представление объекта в виде множества пространственных ГИП позволяет построить и согласованно идентифицировать априорно заданную структуру анализируемого объекта.

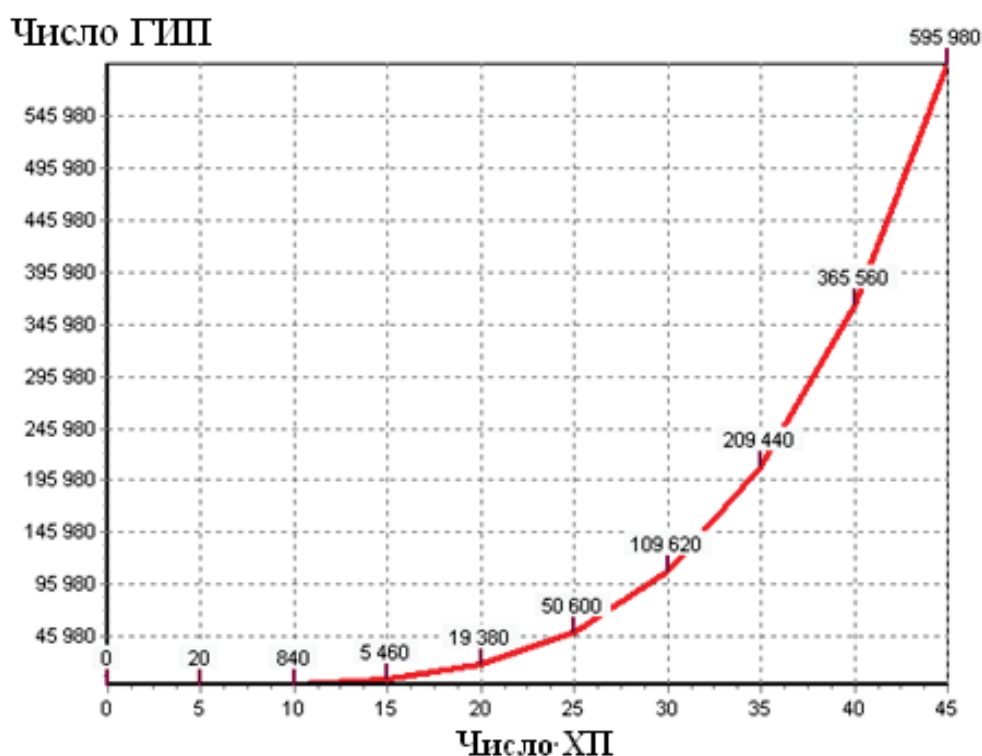


Рисунок 2.9 – Зависимость числа ГИП от базового количества ХП

Преимуществами построения описаний на основе ГИП как частного случая более общего представления в виде ППС есть такие ключевые факторы, как простота процедур формирования множества элементов описания в виде пространственных структур, возможность предварительной фильтрации множества признаков на этапе построения описания, возможность непосредственного управления гибкими геометрическими конструкциями, построенными на координатных соотношениях ХП внутри синтезированного описания для осуществления эффективной классификации.

2.3 Метод сопоставления описаний в виде ППС при классификации

В задаче сопоставления описаний не всегда исходные данные являются достаточно корректными, поэтому возникает необходимость выбора улучшенного в некотором смысле множества признаков для сопоставления описаний. В подобных случаях применимы методы оценки параметров, учитывающих присутствие выбросов в исходных данных, например, схемы

голосования, а также семейство методов на основе случайных выборок (RANSAC) [81]. Схема работы алгоритма RANSAC заключается в циклическом повторении поиска соответствия четырех ХП на сопоставляемом объекте и эталоне. Лучшим соответствием считается то, в котором достигнут минимум суммы отклонений всех ХП объекта и эталона при некотором преобразовании за заданное количество циклов (≤ 2000). В результате используются только те точки, отклонение которых не превосходит заданного порога.

Предлагаемый метод сопоставления структурных описаний видео-объектов на основе построения и анализа моделей ППС состоит в реализации следующих основных этапов:

- 1) выделение на изображении множества ХП и формирование массивов U_i со значениями координат X_i, Y_i ;
- 2) построение множества ППС z_i в виде структуры (2.9) на основе U_i ;
- 3) целенаправленная обработка множества ППС в целях уменьшения избыточности и сокращения объема вычислений;
- 4) вычисление меры подобия между описаниями ППС путем голосования компонентов описаний;
- 5) принятие решения о классе видео-объекта на основе значения вычисленного критерия подобия.

Результирующая мера подобия используется для оптимизации на множестве описаний эталонов (см. рис. 2.10).

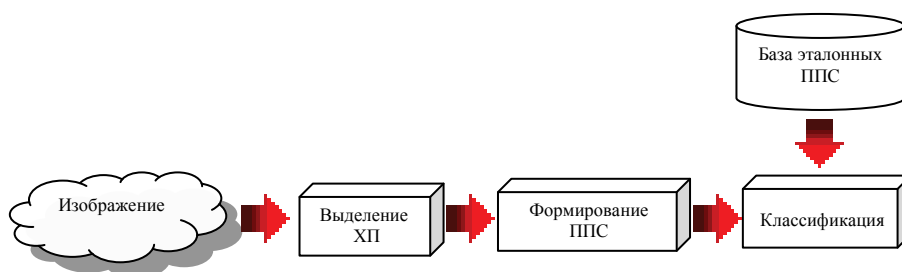


Рисунок 2.10 – Преобразование данных в процессе классификации

Основной предмет исследований сосредоточен на этапах 2–5 предложенного метода. При этом раздел 3 посвящен усовершенствованиям, связанным с

групповым пространственным анализом элементов описания, а также с применением целенаправленной фильтрации и совместного использования разнотипных признаков. В разделе 4 осуществлено усовершенствование меры сопоставления множеств ППС на основе оптимального венгерского метода.

Каждый из рассмотренных этапов, в свою очередь, имеет внутреннюю структуру. В целом применение разработанного метода направлено на повышение достоверности классификации за счет учета контекста и пространственных групповых свойств элементов описания.

Будем рассматривать классификацию как отображение структурных компонентов объекта в виде описания U на компоненты множества эталонов в виде конечного множества эталонных описаний $\{U(j)\}_{j=1}^J$. Реализуем это отображение путем решения задачи оптимизации вида

$$j^* = \arg \max_{j \in J} \vartheta(U, U(j)), \quad (2.13)$$

где J – множество номеров классов ($j^*, j \in J$),

$U, U(j)$ – соответственно описания в виде множеств ППС объекта и j -го эталона,

ϑ – мера подобия для множеств.

Мера подобия является безразмерным показателем для количественного определения степени сходства объектов. Максимальное значение величины $\vartheta(U, U(j))$ означает, что объект наилучшим образом соответствует j -му эталону.

В принципе, в качестве меры ϑ может быть использована и метрика для множеств. Такими метриками могут быть метрики Хаусдорфа, Танимото, и другие [115]. Наилучшее соответствие характеризует значение метрики, равное нулю. С другой стороны известно, что вычисление значений метрики для множеств требует большего числа вычислительных операций, чем вычисление подобия на основе числа голосов элементов. Поэтому в прикладных задачах в большинстве ситуаций для описаний используют меры подобия, а на уровне отдельных элементов описаний – метрики.

Традиционно значение критерия ϑ вычисляется путем анализа хеш-таблицы, построенной на предварительном этапе классификации на основе анализа множества эталонных описаний [4, 47, 48]. Структура элемента ППС имеет вид (2.9). Вычисление значения $\vartheta(U, U(j))$ сводится к подсчету числа голосов, отданных элементами множества базисов в описании объекта за эталон с номером j , представленный в виде ППС. Возможное число голосов с учетом нормировки зависит от применяемого метода голосования.

Достоверность классификации с применением конкретного метода может быть оценена значением критерия θ , вычисляемого как отношение ближайшего максимума меры подобия среди эталонов базы к значению меры подобия анализируемого объекта с истинным эталоном [51–53]. Значение критерия $\theta \in [0, 1]$ показывает, насколько уверенно осуществляется принятие решения на основе вычисленного максимума голосов, отданных за конкретный эталон. Чем меньше значение θ (ближе к 0), тем в большей степени глобальное оптимальное решение значимо по отношению к локальному оптимуму, соответствующему наиболее близкому из остальных (возможно, ложных) классов. Наши эксперименты показали (см. раздел 4), что для нескольких эталонов базы Coil-20 при применении традиционного метода параметр θ достигает значения $\theta = 0,8$, в то время как для предложенного метода на основе ППС значение $\theta = 0,02$, что значительно лучше в плане достоверности при различении объектов. Такое низкое значение критерия θ для разработанного метода подтверждает его высокий показатель правильной классификации для анализируемой базы видео-информации.

Предложенный метод сопоставления, основанный на анализе описаний в виде структур пространственных признаков с применением голосования, за счет введения и использования в процессе вычисления подобия новых информационных связей в виде соответствия структур в целом приводит к усовершенствованию процедуры классификации и обеспечивает решение задач с более высоким уровнем достоверности. Необходимого качества классификации визуальных объектов с точки зрения уровня устойчивости к

аддитивным помехам для конкретных баз видеоданных можно добиться путем выбора мер подобия и критерия принятия решений о классе с учетом параметрического определения.

2.4 Конфигурации описания и модели голосования при сопоставлении ППС

Рассмотрим два наиболее распространенных подхода к построению моделей голосования применительно к ППС [49, 57].

Однозначное голосование. Путем вычисления минимального расстояния от тройки α_q как признака объекта до элементов множества $\{\alpha_p\}_j$ j -го эталона определяется

$$\rho(\alpha_q, \{\alpha_p\}_j) = \min_p \rho(\alpha_q, \alpha_{pj}), \quad (2.14)$$

где $\rho(\alpha_q, \alpha_{pj})$ – некоторая метрика на множестве ГИП.

В качестве метрики при сопоставлении двух векторов ГИП (троек) можно использовать, например, манхэттенское расстояние

$$\rho(\alpha_q, \alpha_{pj}) = |\varsigma_q - \varsigma_{pj}| + |\eta_q - \eta_{pj}| + |\gamma_q - \gamma_{pj}| \quad (2.15)$$

или евклидову метрику

$$\rho(\alpha_q, \alpha_{pj}) = \sqrt{(\varsigma_q - \varsigma_{pj})^2 + (\eta_q - \eta_{pj})^2 + (\gamma_q - \gamma_{pj})^2}. \quad (2.16)$$

Евклидово расстояние – наиболее распространенный вид меры релевантности в пространстве геометрических признаков. Например, при применении манхэттенского расстояния, в сравнении с евклидовой метрикой, влияние отдельных больших разностей (выбросов) значений признаков уменьшается, так как разности не возводятся в квадрат.

Далее необходимо оценить значимость величины $\rho(\alpha_q, \{\alpha_p\}_j)$, полученной при определении значения (2.14) путем сравнения с некоторым порогом ε для метрики. Если выполняется условие $\rho(\alpha_q, \{\alpha_p\}_j) < \varepsilon$, то

значение числа голосов j -го эталона увеличиваем на единицу, считая, что признак α_q нашел соответствие в описании эталона j . Здесь каждая тройка α_q инвариантов эталона может сформировать всего один голос (или не одного). Максимальное число голосов в такой схеме равно числу признаков объекта.

Формально функция вычисления голоса имеет вид:

$$h(\alpha_q, j) = \begin{cases} 1, & \min_p \rho(\alpha_q, \alpha_{pj}) \leq \varepsilon, \\ 0, & \min_p \rho(\alpha_q, \alpha_{pj}) > \varepsilon. \end{cases} \quad (2.17)$$

Все известные подходы построения меры подобия между описаниями Z_1 , Z_2 используют в качестве базовой информации матрицу расстояний

$$P = \begin{bmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1n_1} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2n_1} \\ \dots & \dots & \dots & \dots \\ \rho_{n_2 1} & \rho_{n_2 2} & \dots & \rho_{n_2 n_1} \end{bmatrix}$$

между всевозможными парами ХП $z_i \in Z_1$, $z_j \in Z_2$, а n_1, n_2 – мощности конечных множеств Z_1 , Z_2 .

Реализация принципа однозначного голосования элементов z_i связана с анализом отдельных строк матрицы P на предмет формирования голоса, равного 0 или 1. По полученному набору соответствий (голосов) можно вычислить значение критерия подобия двух описаний в виде суммы

$$L(j) = \sum_{q=1}^{n_1} h(\alpha_q, j), \quad (2.18)$$

где n_1 – число признаков объекта.

Более практичным является сравнение описаний, начиная с признаков эталона. В таком случае осуществляется поиск эталонной информации на неизвестном объекте. Максимальное число голосов в этой ситуации равно числу признаков эталона.

Множественное голосование. Голос за эталон j отдается каждый раз в случае, если $\rho(\alpha_q, \alpha_{pj}) \leq \varepsilon$. Максимальное число голосов в этом методе равно произведению количества признаков объекта и эталона.

Формально функцию получения голоса можно представить в виде:

$$h(\alpha_q, j) = \begin{cases} 1, & \rho(\alpha_q, \alpha_{pj}) \leq \varepsilon, \\ 0, & \rho(\alpha_q, \alpha_{pj}) > \varepsilon, \end{cases} \quad (2.19)$$

где α_q, α_{pj} – признаки объекта и эталона j -го класса соответственно,

$\rho(\alpha_q, \alpha_{pj})$ – расстояние между, $j \in J$,

ε – порог значимости.

Итоговая функция для класса-кандидата с номером j в случае множественного и однозначного голосования совпадает и определена выражением (2.18). Полученное значение критерия подобия делится на максимально возможное число голосов, в результате имеем нормированное значение $L(j) \in [0, 1]$, позволяющее принять решение в соответствии с максимальной долей голосов, отданных за класс j .

В конечном плане результат применения методов сопоставления непосредственно зависит от порога ε . Кроме того, принятие решения при множественном голосовании всегда опирается на большее число голосов, чем однозначное, что требует дополнительного нормирования.

Использование структурных пространственных отношений элементов описания путем перехода в пространство ГИП, построенное на основе координат ХП, дает возможность усовершенствования структурной классификации путем введения новых ресурсов, обладающих информационными свойствами высокоуровневой системы признаков.

Экспериментальные исследования эффективности предложенных моделей голосования для базы изображений Coil-20 подробно описаны в разделе 4. Пространственная помеха моделировалась путем перехода координат точек описания в пределах от ± 1 до ± 3 с некоторой вероятностью p (уровень

помехи). Эксперименты показали, что применение модели однозначного голосования более помехоустойчиво, чем множественное голосование. Например, при значении $p = 0,35$ вероятность правильной классификации в окрестности искажения ± 3 для однозначного голосования составила 0,9, а для многозначного – 0,83.

Применение комбинаций координат ХП в виде отношений и построение решений, основанных на групповых свойствах подмножеств множества признаков, заложено в природе структурного представления как способ грануляции информации для получения новых знаний о распознаваемом объекте.

В традиционных методах с использованием инвариантных признаков сопоставление описаний происходит на основе голосования в пространстве признаков, сформированных в одной системе координат (для одного АБ), которая впоследствии подтверждается верификацией, например, с помощью метода наименьших квадратов [69]. Если верификация не прошла успешно, происходит возврат к голосованию признаков, принадлежащих следующему АБ. Считают, что объект будет классифицирован при положительном результате подтверждения. Предложенный метод исключает процедуру верификации, тем самым неоднократно увеличивая быстродействие.

Основное отличие рассмотренного метода состоит в непосредственном использовании при классификации пространственных и других взаимоотношений структурных элементов, за счет чего значение меры подобия будет зависеть уже не только от значений отдельных элементов, а также от их взаимного расположения в описании – структуры описания. Каждая применяемая на практике структура данных обладает особыми свойствами, эффективными для решения конкретных задач, в то же время применение ее в других задачах с точки зрения эффективности часто бывает не совсем целесообразным.

Конфигурация, которая описывает взаимное расположение элементов объекта, является одним из ключевых факторов, влияющих на качество и быстродействие моделей классификации. Ряд возможных конфигураций представлен на рис. 2.11.

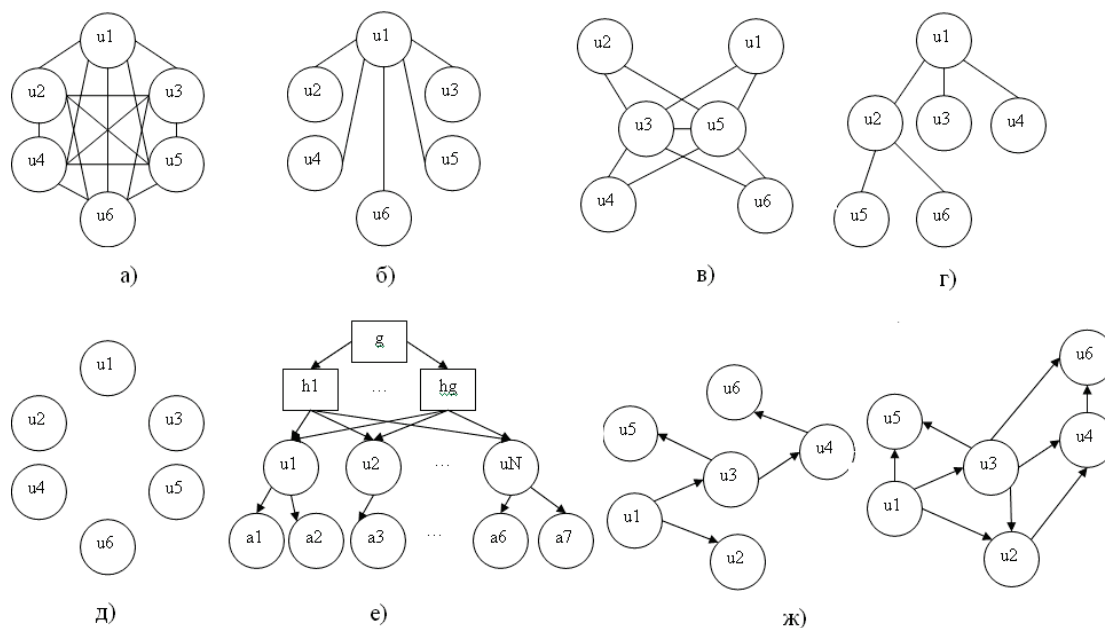


Рисунок 2.11 – Конфигурации структуры связей между элементами объекта:
а) Constellation («созвездие»); б) Star («звезда»); в) k-fan (модифицированная структура типа «звезда», где k – число элементов в клике графа,);
г) Tree («дерево»); д) Bag-of-features (множество ключевых точек); е) Hierarchy (иерархическая структура, описывающая зависимость между частями объекта и их объединениями); ж) Sparse flexible models (направленный ациклический граф, описывающий отношение между частями объекта).

Для выбора практических схем взаимосвязи элементов структурного описания при решении прикладных задач можно воспользоваться теорией построения сетевых конструкций [117]. Построение сети специальных конфигураций отношений между СЭ описания позволяет представить преобразование индивидуальных независимых элементов в подгруппы с центральным элементом и определить отношения признаков при переходе от одного уровня анализа к другому. Внутри одного и того же множества элементов можно построить различные типы структур. Поскольку количество связей между отдельными СЭ достаточно велико, возникает проблема определения значимых подгрупп в сформированном признаковом пространстве. В целом, сетевые структуры могут анализироваться на уровне центральных элементов, уровне подгрупп с повышенной сетевой

плотностью (клик) и общесетевом уровне. Уровень центральных элементов предполагает их сравнение в рамках сети. Клики характеризуются более интенсивными внутренним обменом. Общесетевой системный уровень описывается доминированием и связностью. Доминирование — это отклонение от равномерного распределения связей между «центрами» и «кликами». В системе, характеризующейся высокой степенью доминирования, большинство связей будут соединять «центры» и «клики» [117]. Комбинация доминирования и коммуникативности позволяет установить четыре типа сетевых структур (рис. 2.12). Высокая доминантность и коммуникативность присущи так называемой спутниковой структуре, где ресурсы перемещаются от центра к периферии.

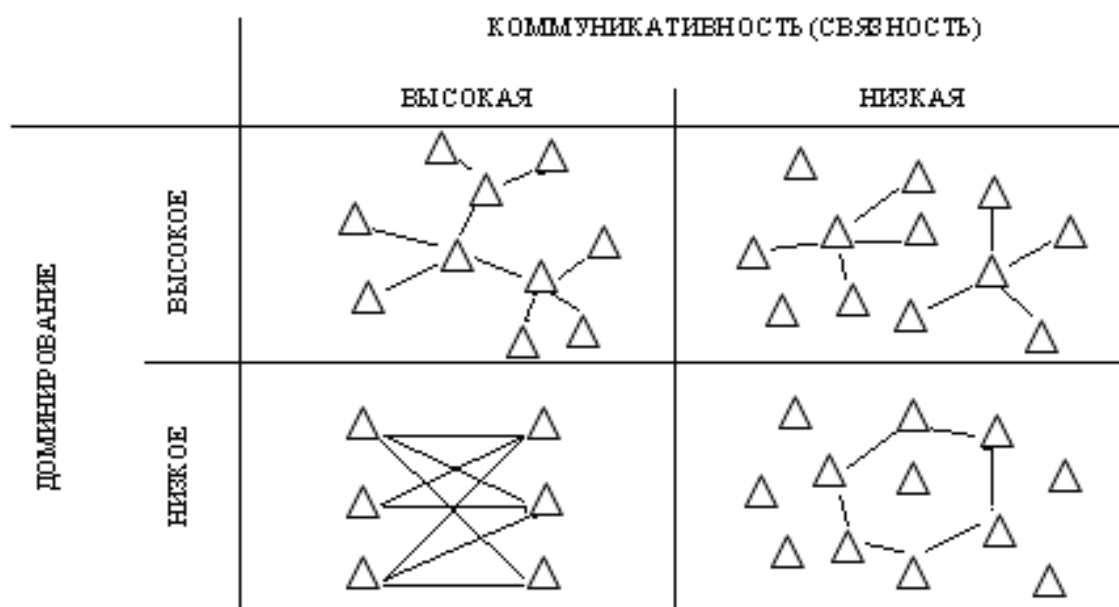


Рисунок 2.12 – Типы сетевых структур

Одно из возможных прикладных применений рассматриваемых в работе методов может быть связано с триангуляционным представлением плоских объектов, например, представляющих собой географические элементы местности в целях идентификации местоположения интересующих объектов или опознания элементов местности на карте [119, рис. 2.13].

Триангуляционная сеть является важным средством описания отношений между ХП, отображающая структуру пространственных отношений для

сопоставления описаний. За счет использования здесь принципа соседства сокращается избыточность представления в среде ППС.

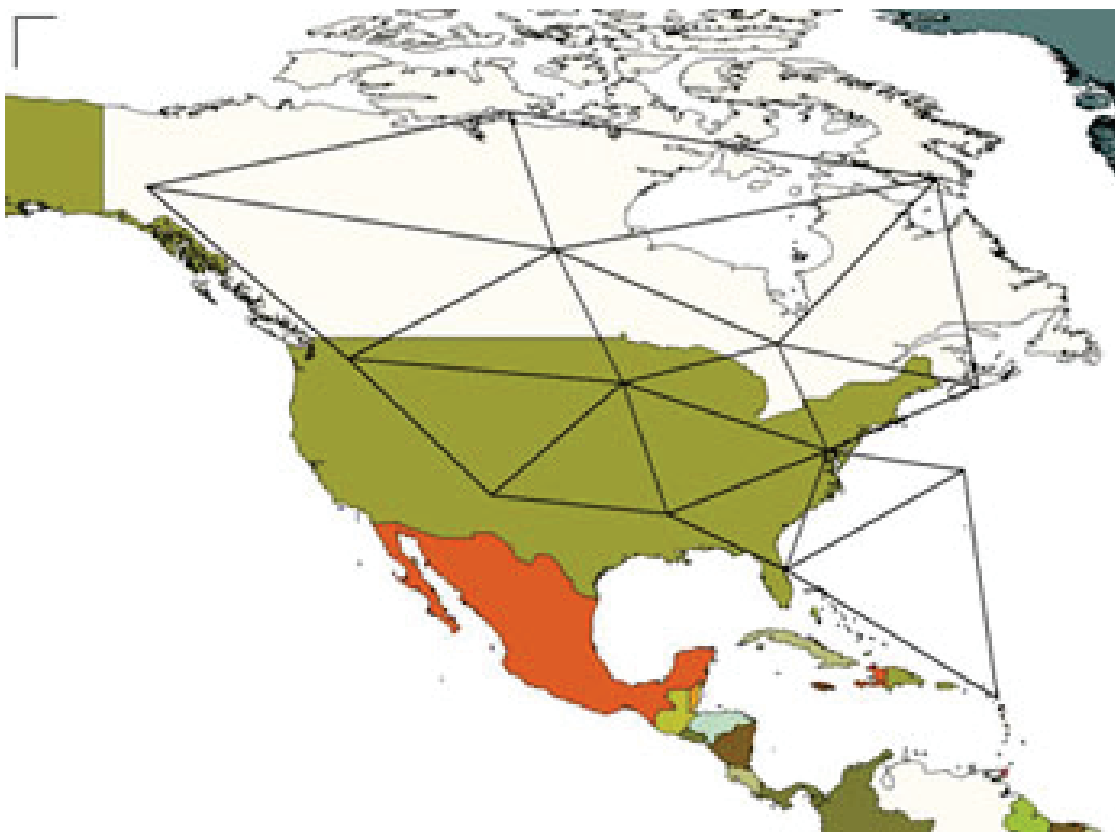


Рисунок 2.13 – Триангуляционная сетка для построения ППС

Таким образом, использование рассмотренных моделей позволяет построить процедуры вычисления подобия для модифицированных признаков описаний, представленных множествами структур.

Выводы

1. Важным фактором для построения системы классификации на основе ХП есть оптимальное соотношение объема хранимых данных, временных затрат на обработку и уровня достоверности классификации. Современные средства формирования ХП (SIFT, SURF, ASIFT, ORB, BRISK) основаны на принципах локальной обработки, дают надежные результаты при решении прикладных задач и могут быть с успехом использованы как основа при построении новых подходов к синтезу пространственных структур признаков.

2. Проведено обоснование выбора формирования описаний объектов при решении задач распознавания зрительных образов в виде ППС, построение которых опирается на систему сформированных связей между отдельными ХП, что дает возможность интегрированного изучения целостных свойств образа. Преимуществами построения описаний на основе ГИП как частного случая ППС являются простота процедур формирования описания, возможность предварительной фильтрации множества признаков на этапе построения описания, возможность непосредственного управления гибкими геометрическими конструкциями, построенными на координатных соотношениях ХП внутри синтезированного описания для осуществления эффективной классификации.

3. Предложен метод сопоставления структурных описаний видео-объектов на основе анализа структур пространственных признаков с применением голосования, который за счет введения и использования в процессе вычисления подобия новых информационных связей в виде соответствия структур объектов усовершенствует процедуру классификации и обеспечивает более высокий уровень достоверности. Эксперименты показали, что при применении традиционного метода показатель достоверности достигает значения 0,8, в то время как для предложенного метода на основе ППС он равен 0,02, что значительно лучше. Такое значение показателя достоверности для разработанного метода подтверждает его высокий уровень правильной классификации.

4. Исследование различных вариантов конфигураций элементов в описании и моделей голосования (однозначное и множественное) показало, что применение однозначного голосования более помехоустойчиво, чем множественного. Например, при уровне искажения $p = 0,35$ вероятность правильной классификации для однозначного голосования составила 0,9, а для множественного – 0,83. Использование моделей голосования позволяет построить процедуры вычисления подобия для модифицированных признаков описаний, представленных множествами структур.

3 ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ КЛАССИФИКАЦИИ ПУТЕМ ГРУППОВОГО АНАЛИЗА ЭЛЕМЕНТОВ ОПИСАНИЙ

Основное предназначение структурного представления визуальных объектов – использование при классификации пространственных и других взаимоотношений структурных элементов, составляющих описание визуального объекта. Формирование описания в виде структуры позволяет осуществить переход к компактному представлению данных при возможно более полном сохранении существенных аспектов информации, содержащихся в них. На сегодня еще не существует универсального подхода к формированию множеств структурных элементов, и каждая применяемая на практике структура данных обладает своими специфическими свойствами, эффективными для решения конкретных задач, в то же время применение ее в других задачах часто бывает неэффективным. Новые перспективные возможности для анализа описаний представляет одновременное изучение двух или более видов структур для одного и того же множества признаков или совместное использование различных типов данных для представления описания. Приведенные в монографии материалы исследований показывают, что, используя различные варианты формирования отношений и анализа множества ГИП как нового структурного пространства с точки зрения тематической группировки вокруг отдельных ХП или вокруг отдельных АБ, можно в результате достичь повышения характеристик качества распознавания [52, 59].

3.1 Пространственная группировка элементов описания

Группировка множества ГИП производится с целью получения более сжатого описания огромных массивов информации в виде групп данных, обладающих свойствами, нужными для осуществления классификации.

Значительное количество индивидуальных признаков заменяется меньшим числом групп признаков, что позволяет лучше уяснить структуру всего множества на предварительных этапах обработки данных для последующего использования результата группировки на конечном этапе классификации.

Рассмотрим два основных варианта представлений, анализа и обработки структурного описания в виде множества A , полученного путем вычисления ГИП на основе исходного множества $U = \{u_i\}_{i=1}^n$ ХП [52, 59].

Исследуем представление в виде объединения

$$A = \bigcup_k \{\alpha\}_k,$$

где $k = \overline{1, C_n^3}$,

$\{\alpha\}_k \in A$ – подмножество значений ГИП для k -го базиса, построенного для каждой из трех неколлинеарных точек исходного описания в виде множества U .

Подмножество $\{\alpha\}_k$ представляет собой набор векторов – троек ГИП для всех ХП в системе координат одного аффинного базиса. Модель множества $\{\alpha\}_k$ имеет вид:

$$\{\alpha\}_k = \langle e_k, \{\alpha_q^k\}_{q=1}^{n-3} \rangle, \quad (3.1)$$

где e_k – базис, вокруг которого осуществляется группировка (рис. 3.1).

Сопоставление двух описаний E и B в результате сведено к вычислению меры подобия, отражающей число эквивалентных в некотором плане элементов множеств [4]. Считая E эталоном с числом ГИП m_E , а B – анализируемым объектом с числом ГИП m_B , процедура определения их подобия в данном представлении сводится к установлению эквивалентности отдельных базисов, описанных сформированными подмножествами ГИП. Для этого нужно ввести некоторую процедуру или критерий для оценки идентичности базисов. Будем считать два базиса эквивалентными, если для конкретного базиса из E найдено более половины эквивалентных элементов из текущего базиса описания B . Если такая эквивалентность установлена, проверка остальных базисов из состава B уже может не осуществляться. В результате мера подобия может

определяться, например, как доля эквивалентных базисов к общему числу базисов эталона E :

$$\vartheta(E, B) = \frac{m_e}{m_E}, \quad (3.2)$$

где m_e – число эквивалентных базисов из E .

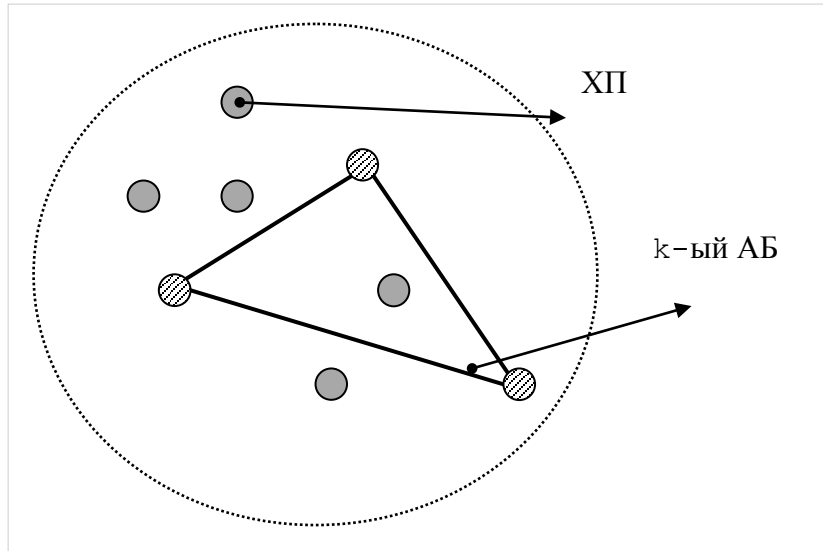


Рисунок 3.1 – Иллюстрация формирования $\{\alpha\}_k$

Другой вариант представления и сопоставления сводится к установлению соответствия между сжатыми множествами базисов эталона и объекта, что направлено на сокращение времени классификации [52].

Рассмотрим теперь второе представление описания в формате

$$A = \bigcup_q \{\alpha\}_q, \quad q = \overline{1, n},$$

где $\{\alpha\}_q$ – подмножество ГИП для ХП с номером q .

Подмножество $\{\alpha\}_q$ содержит совокупность тех α , которые связывают один ХП во множестве всех имеющихся АБ. Структура элемента описания схематично показана на рис. 3.2 как облако ГИП, которые концентрируются вокруг отдельного ХП, и принимает вид кортежа:

$$\{\alpha\}_q = \langle x_q, y_q, \{\alpha_k\}_{k=1}^n \rangle. \quad (3.3)$$

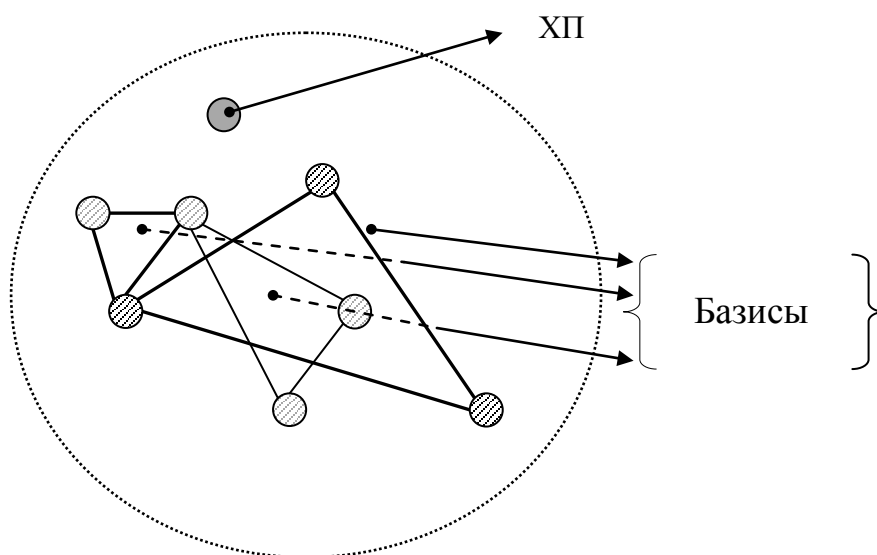


Рисунок 3.2 – Иллюстрация формирования $\{\alpha_k\}_q$

Так как множество ГИП содержит информацию о форме объекта и формируется на основе всевозможных комбинаций базисов для $p - 1$ ХП (всех, кроме фиксированного), то всегда можно выбрать некоторое подмножество $p_v < p$ элементов, от которых в определяющей мере зависят геометрические характеристики объекта. В этом случае подобие устанавливается на основе эквивалентности лишь p_v элементов. За счет сокращения количества точек в представлении эталона снижается и число АБ, что в результате уменьшает и вычислительную сложность. Учитывая в представлении (3.3) некоторое дублирование структурных данных для разных ХП, целесообразно осуществить предварительную фильтрацию множества ХП, что одновременно с сокращением данных часто обеспечивает достаточно высокий уровень достоверности при классификации. Конкретным примером фильтрации является проверка точек описания на коллинеарность, что фильтрует и сокращает описание [52]. В общем случае может быть сформулирована и решена задача о минимизации признакового пространства с точки зрения некоторого критерия для качества работы метода или множества методов на этом пространстве [31].

Применение разработанных моделей группировки элементов ППС во множестве ГИП демонстрирует получение описания в виде групп признаков со своей структурой, обладающей свойствами, нужными для осуществления эффективной классификации. Тем самым обеспечивается повышение показателей достоверности за счет доминирования при принятии решений группы связанных признаков над традиционными решениями, которые принимаются по отдельным кандидатам на множестве независимых признаков.

3.2 Фильтрация множества ППС

В связи с обсуждаемыми вариантами представлений структурного описания рассмотрим теперь пути и конкретные процедуры сжатия описания с целью усовершенствования процесса сопоставления с точки зрения критерия быстродействия [52]. В целях сокращения объема структурного описания при первом варианте группировки (3.1) вокруг отдельных АБ введем процедуру $F_1 : A \rightarrow A_1$ фильтрации на множестве базисов. В результате получим и изучим сформированное сжатое описание $A_1 = \bigcup_{k=1, \dots, s} \{\alpha\}_k$, где $k = \overline{1, s}$, s – число элементов в сжатом описании. Коэффициент сжатия равен C_n^3 / s , что приводит к соответствующему сокращению времени распознавания.

Одним из примеров F_1 может быть процедурное формирование A_1 на принципах участия каждого эталонного ХП хотя бы в одном из базисов, используемых при сопоставлении в целях классификации.

Рассмотрим в качестве практического примера один из принципов формирования связей. Его отличие от сетевых конструкций, описанных в подразделе 2.3, состоит в получении связей уже в процессе осуществления классификации. Суть принципа состоит в создании сети АБ в соответствии со случайной нумерацией ХП. Например, первый базис e_1 содержит ХП с номерами 1, 2, 3. Вторым АБ e_2 сформируем на основе 2, 3 и 4-го ХП. Процедура F_1 продолжается до тех пор, пока закончится возможность

построения АБ из троек ХП. В общем случае k -тый базис e_k содержит ХП с номерами $\{u_i, u_{i+1}, u_{i+2}\}$, а $e_{k+1} = \{u_{i+1}, u_{i+2}, u_{i+3}\}$, где $\{u_i\}_{i=1}^n \in U$. Число базисов k в сжатом описании достигает значения $n - 2$, что на порядок меньше, чем в полном описании.

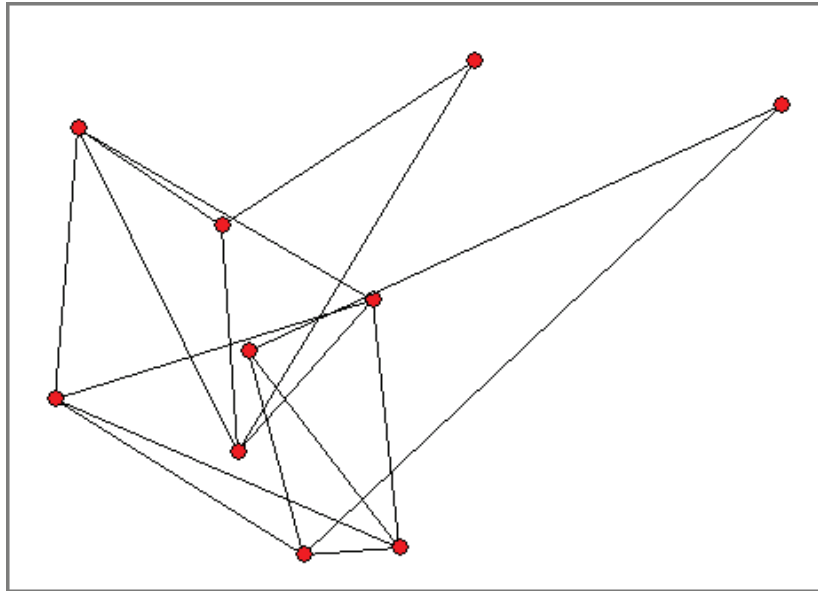


Рисунок 3.3 – Пример формирования случайной конфигурации АБ на основе предложенной процедуры F_1

В конкретном примере (рис. 3.3) для 10-ти ХП число АБ в полном описании достигает 120, а с применением процедуры F_1 – 8. Это существенно сокращает вычислительные затраты и сохраняет в этих условиях надежность принятия решения, т.к. в формировании базисов участвуют все эталонные ХП, т.е. эталонная информация не теряется.

Вариантом процедуры F_2 сжатия и фильтрации, связанным со вторым способом группировки (вокруг ХП), является проверка точек описания U на выполнение условия коллинеарности $F_2 : U \rightarrow U_2$ [51]. Известно, что необходимым условием для формирования значений ППС α есть неколлинеарность точек АБ, где они вычисляются. В дискретном представлении проверку условия неколлинеарности будем выполнять в соответствии с выражением для координат ХП

$$|(x_3 - x_1)(y_2 - y_1) - (x_2 - x_1)(y_3 - y_1)| < \varepsilon_k, \quad (3.4)$$

где $e = (x_r, y_r)_{r=1}^3$ – координаты АБ, ε_k – порог.

Если для координат (x_3, y_3) условие (3.4) выполняется, то точка (x_3, y_3) исключается из формируемого списка, т.к. лежит на одной прямой с двумя другими точками. Свяжем процедуру отбора точек с порогом точности ε_k . С увеличением порога ε_k путем применения обработки (3.4) число ХП, а с ним и число анализируемых АБ сокращается. Как результат, сокращается и множество элементов описания в виде ППС $A_2 = \bigcup_{k=1, \dots, s} \{\alpha\}_k$.

В проведенном нами эксперименте [59] для изображения рис. 3.4,а количество признаков в начальном описании было равно 38. После фильтрации неколлинеарных точек получено: для порога $\varepsilon_k = 2$ – 26, для порога $\varepsilon_k = 5$ – 13, для $\varepsilon_k = 10$ – 11 точек. При пороге в интервале $5 < \varepsilon_k \leq 10$, как показало компьютерное моделирование, число ХП в представлении этого конкретного объекта из базы soil-20 снижается незначительно на 5%.

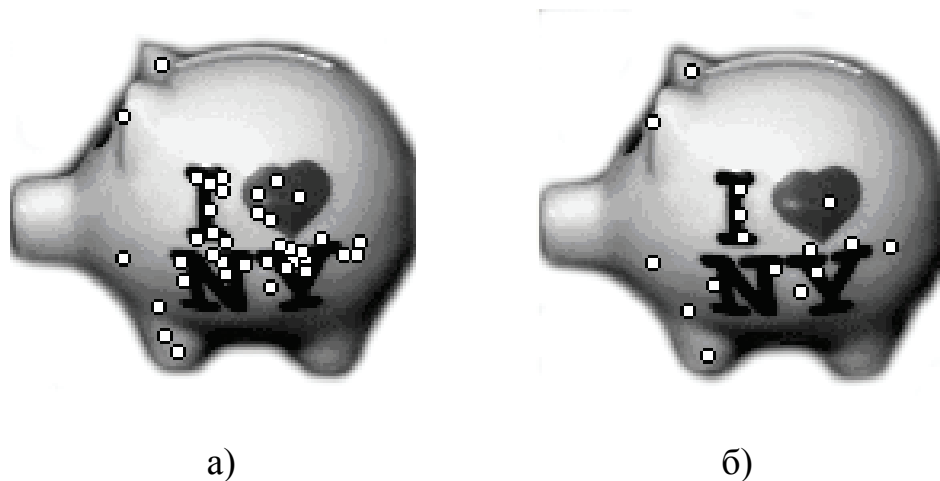


Рисунок 3.4 – Изменение числа ХП при фильтрации базисов:

а) исходное описание (38 ХП), б) сокращенное при $\varepsilon_k = 5$ (13 ХП)

Такая эффективная фильтрация (описание сокращается более, чем в три раза) имеющихся информативных точек изображения дает значительное преимущество в быстродействии модифицированного метода, основанного на пороговой обработке (3.4), по сравнению с классическим

подходом, использующим полное описание. Отметим, однако, что при уменьшении объема структурного описания, очевидно, несколько снижается достоверность и помехозащищенность классификации, поэтому необходимо оценить и контролировать эти характеристики при вновь сформированных сокращенных описаниях.

Еще одним вариантом способа сжатия, связанным со вторым путем группировки (вокруг ХП), является применение процедурного построения F_3 . Суть F_3 состоит в применении ансамбля (логической конъюнкции) пар соседних точек описания U для группового принятия решения о классе анализируемого объекта $F_3 : U \rightarrow U_3$. Сокращение объема вычислений здесь достигается не за счет полного перебора пар точек, а ограниченного, в котором участвуют только соседние точки. Смысл усечения состоит в формировании последовательности (цепочки) пар ХП, которая формируется в соответствии со случайной нумерацией. Например, если ХП с номером 1 и 2 объекта принимают решение о присутствии на эталоне, то далее выбирается следующая пара ХП с номером 3 и 4. Процедура продолжается до тех пор, пока закончится возможность построения пар ХП. Достоинством использования F_3 при классификации является нечувствительность к искажениям отдельных значений признаков в результате более надежного принятия решения и способность эффективно обрабатывать данные с большим числом признаков.

Результат применения обсуждаемых здесь методов F_1, F_2, F_3 сжатия описаний зависит от базы эталонов, в которой осуществляется классификация, а также от ряда других факторов: метода выделения ХП, порога на параметры коллинеарности, требуемого уровня достоверности, процедур выбора базисов и ХП в сжатом описании. Введенные процедуры построения эффективных описаний обладают таким достоинством, как возможность управления объемом анализируемых данных при классификации. Понятно, что факторы достоверности и вычислительных затрат являются взаимно исключаящими, что дает возможность выбора решения на основе оптимального соотношения

их взаимосвязи или предпочтения одного из этих факторов в зависимости от решаемой проблемы. Оба варианта обработки близки друг другу и используют одно и то же множество ГИП, построенное на множестве имеющихся ХП.

Как видим, применение структурной пространственной группировки совместно с предложенными процедурами фильтрации приводит к значительному сокращению объема описания и, как следствие, времени распознавания, оставляя вероятность классификации на необходимом для прикладных задач уровне.

3.3 Гибридные модели признаковых пространств

Применение в интеллектуальном анализе многомерных данных однотипных подходов, соответствующих одной научной парадигме, для решения сложных проблем далеко не всегда приводит к успеху. В гибридной архитектуре, объединяющей несколько парадигм, эффективность одного подхода может компенсировать слабость другого. Комбинируя различные методы при решении конкретной практической проблемы, можно обойти недостатки, присущие каждому из них в отдельности. Суть гибридных систем состоит в формировании некоторой совокупности признаков из различных компонентов, объединенных в интересах достижения поставленных целей. Основной задачей гибридизации является преодоление недостатков отдельных алгоритмов, и стремление не потерять преимуществ каждого из отдельно взятых алгоритмов.

Одним из путей развития моделей голосования в плане повышения их эффективности в задачах классификации есть комплексирование разнотипных и разно-шкальных признаков (атрибутивных, статистических, пространственных, текстурных, логических и др.), полученных различными способами [14, 83]. Многоаспектный анализ позволяет с большей точностью осуществить классификацию описаний за счет использования сочетаний признаков из разных пространств.

Один из способов комплексирования для нашей задачи классификации – совместное использование при голосовании атрибутивных (например, значения дескрипторов ХП в соответствии с детектором SIFT) и пространственных (значения ГИП) характеристик ХП. При этом ГИП из структурного описания объекта используются при поиске соответствующих им ГИП эталона независимо от атрибутов ХП, а при формировании голоса в целях улучшения достоверности дополнительно проверяется величина подобия соответствующих детекторов ХП.

Структуру описания визуального объекта $S = \{s_i\}_{i=1}^n$ (2.10) дополним k -мерным числовым вектором атрибутивных признаков $d_i \in R^k$, инвариантным к геометрическим преобразованиям из группы подобия, который отражает свойства изображения в локальной окрестности ХП (x_i, y_i) и полученный, например, в соответствии с методом SIFT или SURF [16, 18]:

$$s_i = \langle x_i, y_i, \{\alpha_q\}_{q=1}^{m_i}, d_i \rangle. \quad (3.5)$$

В общем случае $\{d_i\}$, $\{\alpha_q\}$ являются конечными мультимножествами. Структура элемента s_i проиллюстрирована на рис. 3.5.

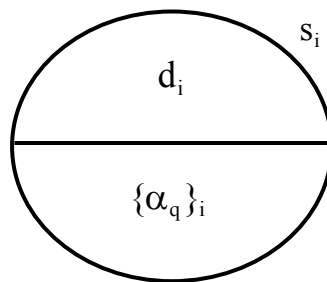


Рисунок 3.5 – ХП s_i в гибридном представлении

Как правило, перед применением основного метода классификации производится предварительная нормировка признаков на единичный гиперкуб.

Рассмотрим возможные подходы к гибридизации путем анализа их в виде следующих основных классов [48].

1. Один из способов используется для поиска перспективных областей в пространстве решений, другой осуществляет среди них локальный поиск для

дальнейшего нахождения глобального оптимума или, другими словами, выполняется иерархическая стратегия оптимизации, при которой локальный компонент поиска уточняет «грубые» решения, полученные в результате глобального поиска партнерским алгоритмом.

Эффективным средством устранения влияния ложных или изменившихся под влиянием помех элементов структурного описания есть логическая обработка значений меры подобия, например, в виде предиката

$$L_V(\rho_V, \varepsilon_V, a, b) = \begin{cases} 1, & \rho_V(a, b) \leq \varepsilon_V, \\ 0, & \rho_V(a, b) > \varepsilon_V, \end{cases} \quad (3.6)$$

где $a, b \in V$ – признаки или множества признаков из некоторого пространства V ,

ε_V – порог для значения метрики, который задает эквивалентность a и b [85].

Применяя бинарную обработку $L_V(.)$ в соответствующих пространствах признаков совместно с логической операцией Θ , можно построить меры подобия в виде функции:

$$\vartheta(s_i, s_j) = \varphi[\rho_\alpha, \rho_d, \Theta[L_\alpha(.), L_d(.)]], \quad (3.7)$$

где Θ задает логическую взаимосвязь для признаков из разных пространств,

ρ_α, ρ_d – метрики,

$\varepsilon_\alpha, \varepsilon_d$ – пороги в предикатах L_α, L_d для соответствующих типов признаков.

Если, например, в (3.7) выбрать в качестве Θ конъюнкцию, то имеем меру на базе проверки истинности обоих предикатов. Если Θ – дизъюнкция, то имеем меру на основе истинности одного из предикатов. Логическая модель (3.6) больше подходит для установления соответствия описаний в виде множеств. Один из вариантов: если минимум по $\{d\}$ и минимум связанного с ним $\{\alpha\}$ определяются внутри одного и того же эталона, то засчитывается голос за этот эталон. Более практичным есть применение конъюнкции, другие варианты приводят к последовательной модели.

Правило определения голоса $h(s_i, j)$ элемента s_i при вычислении (3.7) на основе Θ в виде конъюнкции можно получить в виде

$$h(s_i, j) = \begin{cases} 1, & [\theta(\alpha_i, j)) \neq \emptyset] \& [\theta(d_i, j) \neq \emptyset], \\ 0, & \text{otherwise,} \end{cases} \quad (3.8)$$

где $\theta(d, j) = \{d \in \Lambda(j) \mid L_d\}$, $\theta_j(\alpha) = \{\alpha \in \{\alpha\}_j \mid L_\alpha\}$ – множества соответствий атрибута d и ГИП α в компоненте j -го эталона, а символ $\&$ означает одновременное выполнение условий. Голос (3.7) засчитывается, если множества соответствий каждого из типов признаков в эталонном множестве не пусты.

Если в (3.7) считать ϕ способом комплексирования, то это выражение приобретает общий характер с учетом произвольного типа совместной обработки пространств признаков. Выражение (3.7) и его конкретизацию можно охарактеризовать как меры на пространственно-атрибутивных отношениях элементов α , d , т.к. именно отношения элементов выступают аппаратом при построении таких мер.

Применение мер вида (3.7) в задаче классификации сводится к подсчету и оптимизации числа голосов $h(s_i, j)$, отданных за эталон класса j . Результатом голосования есть значение функции $L(j) \in [0, 1]$, определенной (2.16).

2. Один алгоритм используется как «вторичный» метод оптимизации параметров другого. Вычислим в пространстве геометрических признаков $\{\alpha\}$ расстояние

$$\rho(\alpha_{iA}, \{\alpha_p\}_B) = \min_p \rho_\alpha(\alpha_{iA}, \alpha_{pB}) \quad (3.9)$$

от ХП с номером i эталона A до множества $\{\alpha_p\}_B$ ХП объекта B . Зафиксируем номер l элемента из B , для которого достигается минимум (3.9). Далее проверяем условие на значимость подобия (3.8) посредством предиката $L_\alpha(\cdot)$. При выполнении $L_A(\cdot) = 1$ голос $h(s_i, j)$ элемента α_{iA} засчитываем ($h(s_i, j) = 1$), если дополнительно выполняется также истинность предиката $L_d(\cdot)$ для атрибутивных признаков. В противном случае принимаем $h(s_i, j) = 0$.

Аналогично (3.9) можно вначале найти расстояние $\rho(d_{iA}, \{d\}_B) = \min_q \rho_d(d_{iA}, d_{qB})$ в пространстве атрибутов от точки d_{iA} (компонента кортежа $s_{iA} \in A$) до множества $\{d\}_B$ атрибутов объекта B , а затем

при условии истинности предиката $L_d(.)=1$ определить голос $h(s_i, j)$ путем проверки истинности $L_\alpha(.)$. Если действовать аналогично предыдущему способу, то подтверждение соответствия может быть выполнено в нескольких вариантах, так как каждый из атрибутов d_{iA}, d_{qB} здесь рассматривается вместе со своим множеством ГИП. Можно рассмотреть такие условия: расстояние между соответствующими множествами ГИП меньше некоторого порога, расстояние (3.9) от одного из ГИП до уже зафиксированного эталона меньше порога, расстояние от выбранного подмножества ГИП до множества ГИП зафиксированного эталона меньше порога и др. По объему вычислений этот способ последовательного анализа признаков может оказаться эффективнее предыдущего, т.к. объем данных для элемента пространства $\{d\}$ значительно меньше, чем для $\{\alpha\}$.

Оба способа в предложенном варианте обработки отличаются тем, что один из типов признаков (атрибуты или геометрия) получает преимущество перед другим, который, в свою очередь, играет роль контроля или подтверждения.

Результирующее значение меры вычисляется в виде функции $L(j) \in [0,1]$, связанной с нормировкой на их максимально возможное число.

3. Ассоциативная модель. Совместное использование в одной мере пространственно-атрибутивных признаков можно реализовать в виде ассоциативной меры для $s_{iA} \in A, s_{qB} \in B$:

$$\rho(s_{iA}, s_{qB}) = \frac{1}{2} [\rho_\alpha(\alpha_{iA}, \alpha_{qB}) + \rho_d(d_{iA}, d_{qB})], \quad (3.10)$$

причем (3.10) является метрикой как линейная комбинация метрик [14, 83]. Отметим, что для всех ГИП, относящихся к одному ХП, используется одно и тоже значение дескриптора ХП. Минимум (3.10) определяется на всем множестве эталонных описаний и затем проверяется на превышение порога ε_u для ассоциативной метрики. Величины метрик в (3.10) обычно нормируются к отрезку $[0,1]$ для исключения преобладающего влияния одной из них. Как

показали наши эксперименты, диапазоны изменения признаков α и d достаточно близки. Ассоциативная мера (3.10) учитывает значения признаков из разных пространств равноценным образом, здесь возможно применение также взвешенных характеристик [14].

В общем случае меры (3.7), (3.9), (3.10) применимы и для варианта множественного соответствия ХП с порогами эквивалентности ε_α , ε_d , когда одному ХП объекта может соответствовать несколько ХП из эталонного описания. Множественные соответствия имеют более высокую помехозащищенность, чем однозначные [3].

Сходство пары объектов обуславливается наличием у каждого из них ХП, причем величина сходства зависит от наличия как совпадающих, так и не совпадающих признаков.

Существуют различные способы пространственной группировки ГИП для сопоставления описаний. В работе [48] для отдельного ХП формируется подмножество ГИП путем рассмотрения всевозможных базисов (разные системы аффинных координат). Другим способом есть формирование подмножества ГИП для всех ХП в одной и той же системе координат, т.е. для отдельного базиса. Такой способ представления имеет вид значительного числа подмножеств (количество базисов) с небольшим количеством элементов. Данное представление удобно для применения ВМ из-за небольших по размеру матриц P . Кроме того, оно дает возможность управлять количеством подмножеств (базисов) и принимать решение по неполному описанию.

Материалы экспериментального исследования предложенных гибридных моделей мер подобия подробно описаны в разделе 4. Результаты проведенных экспериментов показали, что при воздействии пространственных помех, искажающих значения ГИП, которое осуществлялось путем перехода ХП в любую из точек его окрестности $\pm 1, \pm 2, \pm 3$, вероятность распознавания в базе coil-20 для традиционного метода на основе α_q достаточно резко снижается до 0,67 с увеличением окрестности влияния помехи, в то время как комплексированные методы сохраняют достаточно высокую вероятность на

уровне 0,95 при значительных искажениях, в частности в окрестности ± 3 . Применение ассоциативной меры показало более устойчивые результаты сопоставления в условиях пространственных искажений по сравнению с комплексированием (3.7) с применением логической операции Θ в виде конъюнкции. Время распознавания для рассмотренных методов практически совпадает.

В материалах проведенных исследований первое показано, что объединение пространств представления структурных признаков в одной мере при сопоставлении описаний улучшает гибкость и универсальность анализа и повышает достоверность классификации за счет более точного и многостороннего анализа свойств описаний. Совместное использование в гибридных моделях голосования признаков из геометрических и атрибутивных пространств, которые свойственны системам компьютерного зрения, показывает предпочтительные характеристики классификации в условиях аддитивных и пространственных помех по сравнению с известными методами, что говорит о целесообразности их применения в прикладных задачах.

Выводы

1. Применение моделей группировки элементов ППС обеспечивает повышение показателей достоверности за счет доминирования при принятии решений группой связанных признаков над стандартными решениями, которые принимаются по отдельным кандидатам на множестве независимых признаков. В результате классификации с применением группировки элементов ППС степень подобия для неправильно классифицированных эталонов не превышает 2%, в то время как для традиционных методов она достигает 80%.

2. Совместное использование структурной группировки и фильтрации приводит к возможности значительного сокращения объема описания и снижения вычислительной сложности сопоставления при необходимом для

прикладных задач уровне вероятности классификации. В рассмотренных практических примерах объем описания уменьшился в 1,4 раза.

3. Предложенные гибридные модели построения мер подобия на основе объединения структурных признаков из разных пространств в одной мере при сопоставлении описаний повышают достоверность классификации за счет более гибкого, точного и многостороннего анализа свойств описаний. Так вероятность классификации составила не ниже 0,95 при воздействии пространственных помех. Формирование и совместный анализ разнотипных признаков позволяют с большей надежностью вычислить сходство объектов.

4 ОПТИМАЛЬНОЕ СОПОСТАВЛЕНИЕ ОПИСАНИЙ КАК МНОЖЕСТВ ПРИЗНАКОВЫХ СТРУКТУР. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ И ВНЕДРЕНИЯ МЕТОДОВ

Особый интерес при сопоставлении описаний распознаваемых объектов в задачах искусственного интеллекта представляет применение подходов, в основе которых лежит оптимизация значения некоторого критерия, связанного с величиной подобия описаний. Как правило, это осуществляется путем комбинаторного перебора на конечном множестве вариантов [65, 68]. Одним из эффективных способов структурного анализа данных, связанных с вычислением оптимального сходства конечных множеств, является венгерский метод, названный комбинаторной оптимизацией ограниченных множеств. Его можно рассматривать как решение задачи о назначениях для наиболее общего случая поиска максимального паросочетания в двудольном взвешенном графе [68]. Принцип оптимального сопоставления структурных описаний по сравнению с голосованием ХП позволяет реализовать схему построения однозначных соответствий ХП из сравниваемых описаний, что в целом благоприятствует улучшению достоверности распознавания.

Исследования данного раздела состоят в применении оптимальных методов для сопоставления структурных описаний в форме ППС, сравнительной оценке сложности вычислительных затрат с другими методами, экспериментальном анализе оптимальных методов для практических систем признаков и баз видеоинформации. Обсуждаются вопросы развития модели вычисления подобия на основе принципа формирования оптимальных групповых соответствий. Приведены наиболее важные результаты экспериментальных исследований предложенных метода и моделей, что в целом подтверждает целесообразность и эффективность их практического использования. Моделирование осуществлено в системах программирования

Matlab и Borland C++ Builder. Приведено описание используемых в экспериментах баз видеоданных.

4.1 Постановка и решение задачи оптимизации при сопоставлении описаний

Распространенная в прикладных исследованиях классическая транспортная задача состоит в минимизации при некоторых ограничениях линейной функции вида $C = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min$, означающей суммарные затраты при выполнении условий полного удовлетворения спроса и равенства вывозимого количества продукта запасам его в пунктах отправления [68]. К задачам транспортного типа сводятся многие другие задачи линейного программирования – о назначениях, сетевые, календарного планирования.

Решение транспортной задачи традиционным симплекс-методом часто не учитывает всей специфики условий и поэтому оказывается слишком громоздким. Структура конкретных прикладных ограничений учитывается в ряде специальных вычислительных методов ее решения. Одним из наиболее распространенных по практическим соображениям вычислительных подходов считается Венгерский метод (ВМ) [50, 60, 65, 68].

Суть ВМ состоит в последовательном формировании оптимального решения за конечное число шагов. Процедура на основе ВМ для решения наиболее общей транспортной задачи состоит из предварительного этапа и конечного числа однотипных итераций. На подготовительном этапе строится матрица $X_0 = (x_{ij}[0])_{m,n}$, элементы которой неотрицательны и удовлетворяют

неравенствам: $\sum_{j=1}^n x_{ij}[0] \leq a_i, i = \overline{1, m}, \sum_{i=1}^m x_{ij}[0] \leq b_j, j = \overline{1, n}$, m, n – число пунктов

отправления и назначения, a_i, b_j – запасы и потребности продукта [68].

Если эти условия являются равенствами, то матрица X_0 считается решением

задачи. Если среди этих условий есть неравенства, то осуществляется переход к первой итерации. На k -ой итерации строится матрица $X_k = (x_{ij}[k])_{m,n}$.

Близость этой матрицы к решению задачи характеризует число Δ_k –

$$\text{суммарная невязка матрицы } X_k : \Delta_k = \sum_{i=1}^m a_i + \sum_{j=1}^n b_j - 2 \sum_{i=1}^m \sum_{j=1}^n x_{ij}[k].$$

В результате осуществления первой итерации строится матрица X_1 , состоящая из неотрицательных элементов. При этом выполнено условие $\Delta_1 < \Delta_0$, где Δ_0 – суммарная невязка матрицы X_0 на подготовительном этапе. Если $\Delta_1 = 0$, то X_1 считается оптимальным решением. Если $\Delta_1 > 0$, то переходят к следующей итерации. Итерации проводятся до тех пор, пока Δ_k при некотором значении k не станет равным 0. Соответствующая матрица X_k и считается решением транспортной задачи.

ВМ наиболее эффективен при решении транспортных задач с целочисленными объемами производства и потребления. В этом случае оцениваемое число итераций не превышает величины $\Delta_0 / 2$.

Достоинством ВМ есть возможность в процессе его применения оценивать близость результата каждой из итераций к оптимальному плану. Это позволяет контролировать процесс вычислений и при необходимости (в целях сокращения временных затрат) прекращать его при достижении заданной точности. Такое свойство особенно важно для задач большой размерности, к которым относятся практически все задачи обработки изображений, например, число векторов ХП в структурных описаниях объектов достигает двух–трех сотен.

Важным с прикладной точки зрения частным случаем транспортной задачи есть задача о назначениях, которая формулируется следующим образом (рис. 4.1).

Имеется m работ и n кандидатов для их выполнения. Затраты i -го кандидата на выполнение j -ой работы равны c_{ij} , $i = \overline{1, m}$, $j = \overline{1, n}$. Каждый кандидат

может быть назначен только на одну работу, и каждая работа может быть выполнена только одним кандидатом. Признак $x_{ij}=1$ если i -й кандидат выполняет j -ю работу, $x_{ij}=0$ в противном случае. Требуется найти назначение кандидатов на работы, при котором суммарные затраты на выполнение работ минимальны. Возьмем в качестве модели вычисления оптимального подобия задачу о назначениях, где значение x_{ij} будем использовать как наличие или отсутствие соответствия между элементами сопоставляемых описаний с номерами i, j . Математическая модель задачи о назначениях представлена в виде:

$$C = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min, \quad (4.1)$$

$$\begin{cases} \sum_{j=1}^n x_{ij} = 1, \forall i = \overline{1, m}, \\ \sum_{i=1}^m x_{ij} = 1, \forall j = \overline{1, n}, \\ x_{ij} \in \{0, 1\}, i = \overline{1, m}, j = \overline{1, n}. \end{cases} \quad (4.2)$$

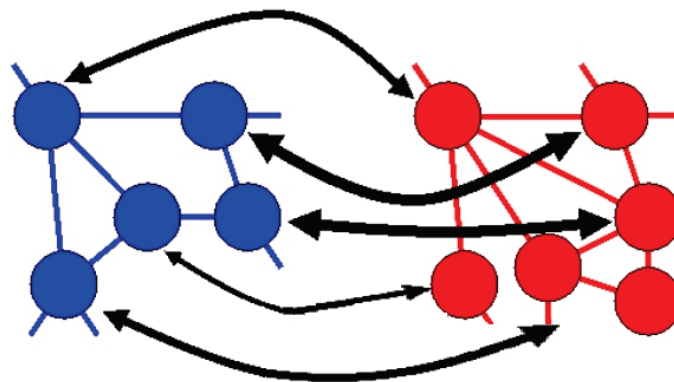


Рисунок 4.1 – Демонстрация задачи о назначениях

Задача (4.1) – (4.2) является задачей линейного программирования, в которой правые части ограничений равны 1, а переменные x_{ij} могут принимать только два значения, и поэтому она может быть эффективно решена с помощью ВМ. Классический вариант задачи о назначениях предполагает

минимизацию целевой функции C для квадратной матрицы при $m = n$:

$$C = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min. \text{ Переход к поиску максимума } C = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \max$$

осуществляется путем замены в (4.1) – (4.2):

$$c_{ij} \leftarrow c^* - c_{ij}, \quad c^* = \max_{i,j} \{c_{ij}\}, \quad i, j = \overline{1, n}. \quad (4.3)$$

Практический интерес для сопоставления структурных описаний произвольной размерности представляют задачи с прямоугольной матрицей, они могут рассматриваться как задачи при $m = n$ после формального дополнения матрицы $\{c_{ij}\}$ строками или столбцами с нулевыми элементами. Сканирование нулевых строк или столбцов будет бесполезным, но необходимым этапом реализации ВМ. Задача о назначениях решает проблему установления величины подобия структурных описаний объектов в условиях несоответствия порядка их следования или нарушения целостности описания из-за помех.

Представим описание Z визуального объекта в виде конечного множества $Z = \{z_i\}_{i=1}^m$ элементов z_i , где m – количество ХП в описании, z_i – ХП, который может представлять как дескриптор (значение признака), так и геометрическую информацию (координаты ХП, значения ППС на основе этих координат и т.д.) [50, 60]. Критерием при сопоставлении элементов является значение некоторой метрики $\rho(z_i, z_j)$, оценивающей различия z_i, z_j , взятых из разных описаний $z_i \in Z_1, z_j \in Z_2$. Методы с использованием голосования определяют сумму голосов элементов, исходя из величин $\rho(z_i, z_j)$ путем оптимизации на конечном множестве или через оценку с помощью некоторого порога. Оптимальные методы на основе $\rho(z_i, z_j)$ формируют решение, например, которое минимизирует сумму расстояний между описаниями Z_1, Z_2 с учетом однозначности соответствий элементов.

Применим ВМ для оптимального установления величины соответствия между двумя структурными описаниями с учетом возможного действия

пространственных помех, приводящих к искажениям отдельных элементов. Результатом выполнения ВМ является формирование максимального паросочетания для элементов двух множеств с минимизацией общей стоимости (весов), которую можно оценить в виде суммы расстояний между парами ХП из сравниваемых описаний. Сопоставление Z_1, Z_2 на основе ВМ формально сведем к решению оптимизационной задачи:

$$R(x) = \sum_{i=1}^m \sum_{j=1}^n \rho(z_i, z_j) x_{ij} \rightarrow \min, \quad (4.4)$$

где $z_i \in Z_1, z_j \in Z_2$,

m и n – количества точек в описаниях Z_1, Z_2 ,

x_{ij} – бинарный признак, $x_{ij} \in \{0,1\}$, отражающий соответствие i -го и j -го элементов из Z_1, Z_2 . Решение задачи (4.4) при ограничении на однозначность соответствия признаков из сопоставляемых множеств

$$\begin{cases} \sum_{j=1}^n x_{ij} = 1 \quad \forall i = \overline{1, m}, \\ \sum_{i=1}^m x_{ij} = 1 \quad \forall j = \overline{1, n}, \end{cases} \quad (4.5)$$

минимизирует расстояние между Z_1, Z_2 .

Как оптимальный, так и традиционный (основанный на голосовании) подходы используют в качестве базовой информации матрицу расстояний

$$P = \begin{bmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2n} \\ \dots & \dots & \dots & \dots \\ \rho_{m1} & \rho_{m2} & \dots & \rho_{mn} \end{bmatrix}$$

между всевозможными парами ХП $z_i \in Z_1, z_j \in Z_2$.

Реализация принципа однозначного голосования элементов z_i связана с анализом отдельных строк матрицы P на предмет формирования голоса, равного 0 или 1. По полученному набору соответствий (голосов) можно вычислить значение критерия подобия двух описаний в виде суммы

$S_1 = \sum_i L(i)$, где $L(i)$ – предикат, равный 1, если голос формируется (соответствие установлено) и нулю – в противном случае.

Реализация ВМ предполагает целенаправленный выбор такой последовательности из $\{\rho_{ij}\}$, сумма значений которой $S_2 = \sum \rho_{ij}$ будет минимальна на множестве возможных соответствий элементов описаний, и при этом из каждой строки и столбца матрицы P будет выбран только один элемент, что в результате обеспечивает выбор оптимального по критерию S_2 однозначного соответствия множеств элементов описаний. Реализация ВМ осуществляется путем целенаправленного эквивалентного преобразования матрицы P с последовательным анализом строк и столбцов для получения матрицы с неотрицательными элементами и системой m независимых нулей, из которых никакие два не принадлежат одной и той же строке или одному и тому же столбцу. При достижении ситуации с m независимыми нулями проблема выбора считается решенной, оптимальный вариант назначений определяется позициями независимых нулей в преобразованной матрице [65]. Пример паросочетания показан на рис. 4.2

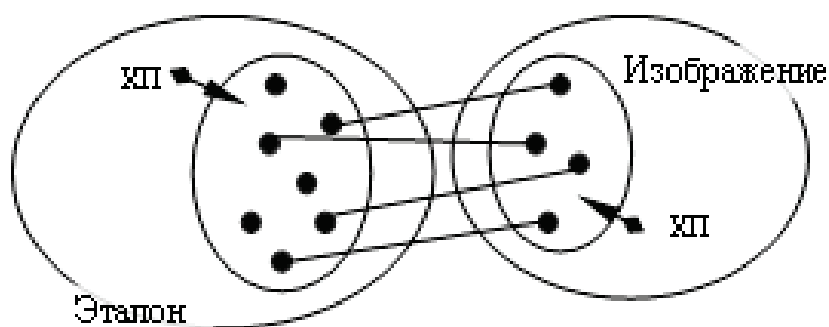


Рисунок 4.2 – Паросочетания для множеств ХП

Известно, что задача поиска паросочетаний с использованием ВМ решается за время, пропорциональное величине d^3 , где d – это размер сравниваемых описаний. Видим, что требуемые затраты вычислений существенно возрастают с увеличением d . В тоже время для метода с голосованием затраты прямо пропорциональны d , что не так критично к росту объема описаний.

При применении ВМ для сопоставления описаний Z_1, Z_2 их размерности m и n в общем случае могут отличаться. В этом случае в традиционном ВМ размер матрицы P увеличивают до максимального среди m и n , а возникшие элементы заполняют нулями [50, 60, 65]. Если исходная матрица не является квадратной, то дополнительно вводят необходимое количество строк (или столбцов), а их элементам присваивают значения, определяемые условиями задачи. Для задачи определения подобия описаний полученным за счет расширения элементам необходимо присвоить некоторое максимальное значение среди возможных значений $\rho(z_i, z_j)$.

Заметим, что для задачи сопоставления не совсем целесообразно усекать матрицу весов до квадратной традиционным способом в виде $\min(m, n)$, т.к. при этом может быть потеряна ценная информация об видео-объекте. С другой стороны, дальнейшее усложнение обработки за счет применения ВМ требует минимизации размеров матрицы на этапе предварительной обработки. Это достигается посредством применения специальных методов сжатия структурных описаний [47, 51]. Другим способом снижения затрат есть отбор в некотором смысле «значимых» строк и столбцов, например, содержащих большие по значению расстояния.

Особый интерес представляет применение оптимальных методов для ППС в виде множеств АИ, построенных на основе координат ХП и непосредственно отражающих взаимную пространственную связь ХП объекта. При этом множество ППС имеет вид конечного множества числовых троек $\{\alpha_q\}_{q=1}^r$, где r – число ППС [50, 51]. Расстояние между признаками α_q и α_p

для матрицы P можно вычислить как $\rho_{qp} = \sqrt{\sum_{i=1}^3 (\alpha_{qi} - \alpha_{pi})^2}$. Взаимно-однозначное соответствие множеств элементов описаний может быть установлено путем решения задачи поиска взвешенного паросочетания. Решение о принадлежности объекта конкретному эталону принимается

на основе максимального числа установленных соответствий между подмножествами ППС или на основе минимума величины S_2 .

В качестве критерия, характеризующего достоверность классификации, будем использовать величину $\theta = h_1 / h_2$, где h_2 – максимум гистограммы голосов, h_1 – ближайший к нему максимум.

Для базы объектов видео-данных Coil-20 [47] вычислено множество ППС $\{\alpha_q\}$. Эксперимент для традиционного подхода на основе голосования ХП состоял в том, что на вход системы распознавания поступало описание одного из эталонов, которое последовательно сравнивалось со всеми описаниями базы, в результате принималось решение в соответствии с максимумом нормированного количества голосов [50]. Для предложенного оптимального метода сопоставление осуществлялось по принципу «множество-множество» с достижением минимума в выражении (4.4) при условии однозначного соответствия элементов множеств. Для базы Coil-20 с применением структурных описаний в виде множеств ППС достигнуто безошибочное распознавание.

Сравнительные эксперименты показали, что в традиционном подходе величина θ для правильно опознанного класса относительно других эталонов достигает значения $\theta = 0,8$, в то время как для оптимального метода на основе ВМ максимальное из значений $\theta = 0,16$ (рис. 4.3). Это непосредственно доказывает, что достоверность распознавания для оптимального метода существенно лучше, т.к. решение о классе принимается более надежно и фундаментально.

В то же время отметим, что для оптимального метода затраты времени на решение задачи несколько больше, чем для традиционного. При этом временные затраты значительно увеличиваются с увеличением размерности сравниваемых описаний, а применение аппарата ППС по сравнению с обычным представлением в виде множества ХП еще более усиливает временные различия. Например, для $m = 6$ с использованием пространства ППС временные затраты оптимального метода при распознавании оказались в

1,5 раза выше, а при $m = 13$ – в десятки раз больше. Отметим, что эти затраты могут регулироваться за счет процедуры предварительной обработки структурных описаний для конкретно решаемой прикладной задачи.

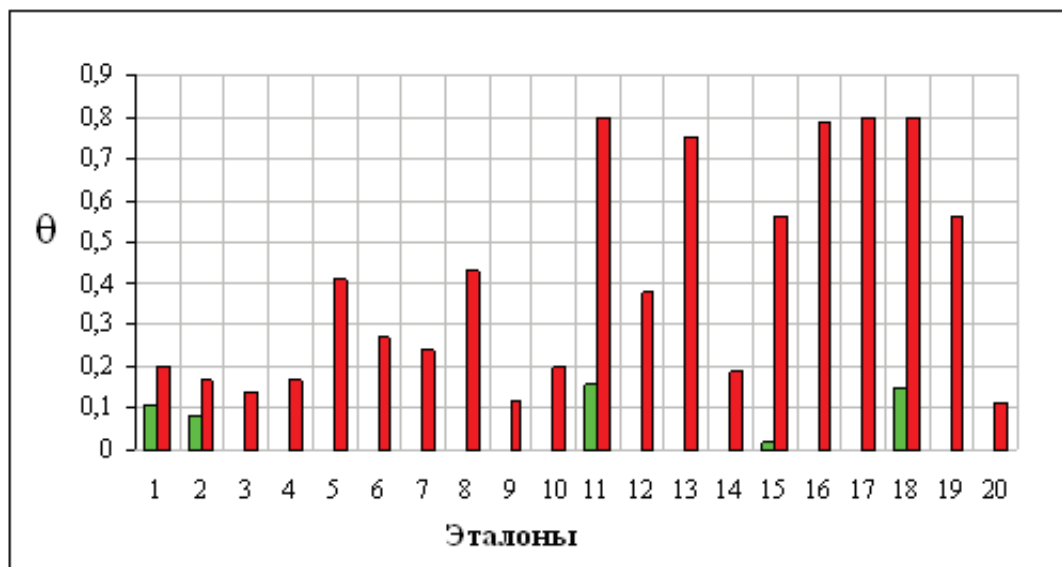


Рисунок 4.3 – Значение критерия θ для сравниваемых методов:

■ – с использованием голосования, ■ – на основе ВМ

Оптимальные методы при сопоставлении структурных описаний объектов за счет более тщательного отбора соответствий обеспечивают лучшую достоверность распознавания по сравнению с традиционными подходами на основе голосования.

В целом полученные характеристики для оптимального метода могут быть оценены как предельно достижимые, в то время как традиционный подход с голосованием отдельных компонентов описания можно считать некоторой практической аппроксимацией меры подобия. Наиболее целесообразно оптимальные методы применять для тех задач, которые требуют наиболее тщательного анализа соответствия описаний.

Результаты исследования показали, что применение оптимальных методов для определения максимального паросочетания в двудольном графе в задаче сопоставления описаний видео-объектов в виде пространственных структур повышает достоверность распознавания за счет однозначного и более точного учета подобия подмножеств описаний.

Практически важным является получение предпочтительных характеристик распознавания по сравнению с известными методами, что говорит о целесообразности развития и применения оптимальных методов сопоставления множеств ППС в задачах компьютерного зрения.

4.2 Результаты экспериментов для моделей классификации на основе разнотипных признаков

Как показали проведенные исследования [48, 49], комплексированные подходы при сопоставлении структурных описаний, предложенные в разделе 3.3, за счет углубленного многостороннего анализа всего разнообразия свойств анализируемых объектов обеспечивают более высокую достоверность распознавания по сравнению с традиционными подходами.

В качестве альтернативы для получения экспериментальных числовых оценок использована традиционная для задач компьютерного зрения база изображений Columbia Object Image Library (COIL-20) Колумбийского университета [89], которая содержит 20 изображений размером 128x128 пикселей (рис. 4.4.)



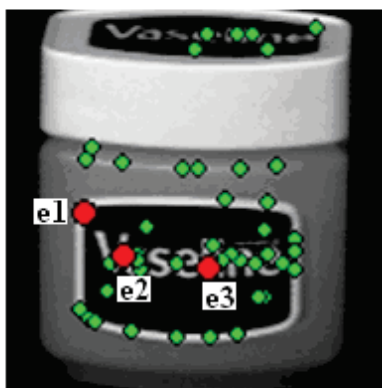
Рисунок 4.4. – База изображений COIL-20

Особенностью данной базы есть смысловая разнотипность видеоданных и достаточная сильная сглаженность яркости, что на практике приводит к незначительному числу ХП. Количество ХП, сформированных детектором SIFT для эталонов базы COIL-20, колеблется в пределах от 6 до 116.

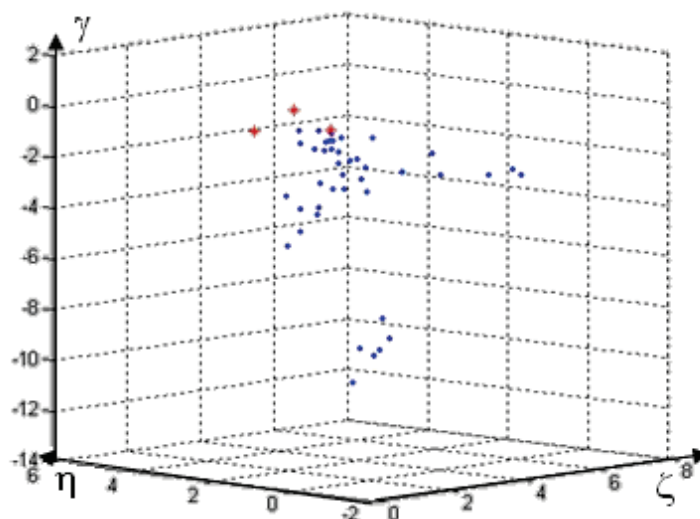
На результат распознавания непосредственно влияют значения порогов точности ε_α , ε_d для пространственных и атрибутивных признаков соответственно, а также порога ε_u для значений ассоциативной метрики (см. раздел 3.3). На практике их выбор зависит от используемой базы видеоданных, а также от уровня аддитивных помех, определяющих допустимое отклонение значений признаков. Обычно величина отклонения задается в виде процента (1-10%) от максимального разброса в значениях признаков или от меры их близости [1,3]. Значения α_q , которые являются барицентрическими координатами для конкретного ХП в определенном базисе и однозначно определяют ХП в пространстве этого базиса, напрямую зависят от удаленности ХП от точек базиса на изображении (рис. 4.5).

Наши экспериментальные исследования по выбору порогов для конкретной базы Coil-20 показывают, что подавляющее большинство значений признаков α_q , сформированных путем анализа координат детектора SIFT, принадлежат диапазону $[-15;15]$ (рис. 4.6), и только отдельные значения не входят в этот интервал [48]. Из этих соображений выберем порог 0,3 как 1% от максимального разброса в диапазоне. Порог ε_d для эквивалентности атрибутов ХП (вектор размерности 128 с нормированными значениями на отрезке $[0,1]$) определялся как 1% от разброса значений евклидовой метрики и выбран равным $\varepsilon_d=0,11$. Значение порога для ассоциативной меры $\varepsilon_u=0,2$ выбрано как среднее арифметическое порогов ε_α , ε_d .

Основная цель проведенных экспериментов – сравнительная оценка эффективности разных видов комплексирования и традиционных подходов в условиях помех.



а)



б)

Рисунок 4.5 – Выделенные на изображении ХП и выбранный для них случайным образом АБ (e_1, e_2, e_3) (а); геометрическая интерпретация распределения значений $\alpha(\zeta, \eta, \gamma)$ в пространстве АБ (e_1, e_2, e_3) (красным – значения, соответствующие точкам базиса, синим – значения $\alpha(\zeta, \eta, \gamma)$, соответствующие остальным ХП в этом базисе (б))

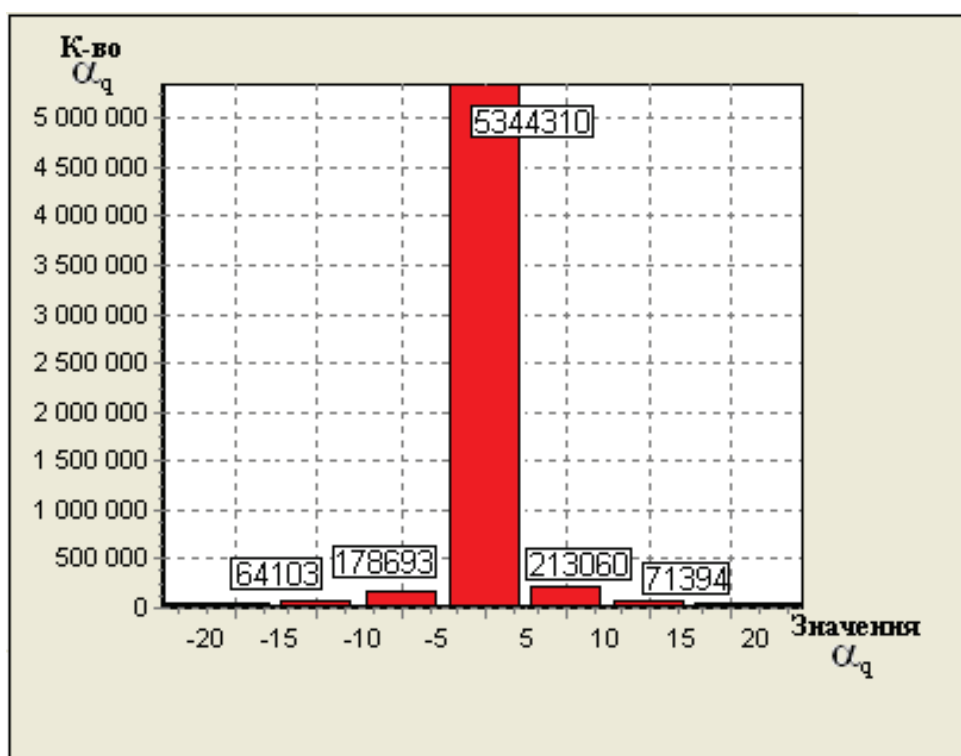


Рисунок 4.6 – Распределение значений α_q для базы изображений Coil-20

Первая задача состояла в оценке работоспособности способа комплексирования на основе (3.9), когда один из типов признаков, в данном случае ГИП, получает преимущество перед другим (значения атрибутивных признаков, вычисленных детектором SIFT), который, в свою очередь, играет роль контроля или подтверждения (подраздел 3.3) при действии на каждый ХП аддитивной помехи с нормальным распределением (матожидание – ноль, дисперсия – σ^2).

Табл. 4.1 содержит оценки значений вероятности P правильного распознавания объектов для базы Coil-20 в зависимости от уровня помехи, задаваемого значением σ среднеквадратического отклонения (СКО). Как видим из табл. 4.1, до уровня $\sigma = 0,012$ такой метод комплексирования обеспечивает вероятность $P = 1$, а при $\sigma = 0,0128$ значение $P \approx 0,9$.

Таблица 4.1 – Вероятность распознавания для аддитивных помех

σ	0,012	0,0125	0,0127	0,0128	0,0129	0,013	0,0135
P	1	0,98	0,92	0,88	0,82	0,67	0,59

Моделирование воздействия пространственных помех, которые искажают значения ГИП, осуществлялось путем перехода ХП в любую из точек его окрестности с вероятностью 0,5, следствием чего является изменение координат отдельных ХП объекта, которые преобразовывались добавлением значений $\pm 1, \pm 2, \pm 3$. Размер окрестности изменения координат отражает уровень помехи. Результаты качества классификации для разных методов: традиционного голосования на основе ГИП, голосования на основе ГИП с пороговой проверкой значения атрибута (комплексирование (3.7)), применением ассоциативной метрики (3.10) отражены в табл. 4.2 (под качеством классификации здесь понимается отношение числа правильно опознанных объектов к общему числу экспериментов). При этом для первого метода применялась только модель пространственных помех, т.к. аддитивная

помеха не изменяет координат ХП, для остальных двух – одновременное действие аддитивной и пространственной помех. Для экспериментов выбран уровень аддитивной помехи $\sigma = 0,0127$, чтобы оценить устойчивость комплексирования в наиболее сложной ситуации влияния помех разных типов.

Таблица 4.2 – Вероятности распознавания для разных методов

Окрестность	0	± 1	± 2	± 3
Голосование α_q	1	0,85	0,76	0,67
Комплексирование	0,92	0,91	0,91	0,89
Ассоциативная мера	0,99	0,99	0,99	0,95

Как показывают экспериментальные значения из табл. 4.2, применение комплексирования имеет существенные преимущества по сравнению с традиционными методами, основанными на раздельном голосовании атрибутов и геометрических признаков АИ. Вероятность распознавания для традиционного метода на основе α_q резко снижается до 0,67 с увеличением окрестности влияния помехи, в то время как комплексированные методы сохраняют достаточно высокую вероятность 0,95 при значительных искажениях. Установлено, что ассоциативная мера более устойчива к пространственным искажениям по сравнению с комплексированием (3.7). Время распознавания для рассмотренных в табл. 4.2 методов (см. раздел 3.3) практически совпадает.

Как видим, в результате экспериментов доказано, что объединение пространств представления структурных признаков в одной мере при сопоставлении описаний в виде ППС улучшает гибкость, универсальность анализа и повышает достоверность распознавания за счет более точного и многостороннего анализа свойств описаний. Практически важным результатом является получение предпочтительных характеристик распознавания в условиях аддитивных и пространственных помех с применением комплексированных мер по сравнению с известными методами,

что говорит о целесообразности их развития и применения в прикладных задачах компьютерного зрения.

4.3 Результаты экспериментов по сопоставлению на основе моделей инвариантных признаков структур

Экспериментальное исследование рассмотренных в разделе 3.1 моделей анализа структурных описаний в виде множеств ППС проведено для базы полутонных видеоданных из базы Coil-20. Суть процедуры фильтрации F_2 в пространстве ХП сводилась к проверке коллинеарности координат базиса, а сравнение осуществлялось путем установления эквивалентности подмножеств $\{\alpha\}_q$ в виде структуры (3.3) объекта и эталона (см. раздел 3.2).

Следующая серия экспериментов была предназначена для оценки влияния порога фильтрации ε_k на число ХП в сжатом описании и оценке при этом достоверности и помехозащищенности на используемой базе данных. Таблица 4.3 содержит номера эталонов базы Coil-20, число точек в их полном описании и после применения процедуры фильтрации с порогом $\varepsilon_k = 3$ на основе проверки условия (3.4) (см. раздел 3.2).

Таблица 4.3 – Число ХП в полном и сжатом описаниях

Номер эталона	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Полное описание	6	6	31	63	58	45	18	25	116	45	5	53	38	27	17	26	14	9	36	30
Сжатое описание ($\varepsilon_k = 3$)	6	5	25	50	47	39	15	21	71	37	5	45	26	22	13	23	12	8	29	25

Из табл. 4.3 видим, что число точек в результате фильтрации сокращается в среднем в 1,2 раза для рассмотренных эталонов. Отметим, что при увеличении значения порога ε_k путем проверки условия коллинеарности количество ХП, а

с ним и число анализируемых АБ будет сокращаться, однако при этом также будет изменяться вероятность правильного опознавания относительно других эталонов. После фильтрации неколлинеарных точек число ХП для базы Coil-20 в среднем уменьшилось: – для порога $\varepsilon_k = 2$ приблизительно в 1,2 раза, – для порога $\varepsilon_k = 5$ – вдвое, для $\varepsilon_k = 10$ – 2,2 раза. При изменении порога в интервале $2 \leq \varepsilon_k \leq 5$, как показало моделирование, резко меняется количество признаков в описании, тогда как при пороге $\varepsilon_k > 5$ число ХП для описаний снижается незначительно. Дальнейшее сокращение числа признаков приводит к существенному снижению достоверности классификации.

Два подмножества ППС $\{\alpha\}_q$ считались эквивалентными, если для объекта найдено более половины эквивалентных элементов, соответствующих эталонному подмножеству. Использовались параметры процедуры фильтрации, для которых вероятность правильной классификации оставалась равной 1. Результаты экспериментов при данных условиях показали, что для отфильтрованного пространства ППС эталона среднее время классификации с применением модели (3.3) примерно в 5,5 раза меньше, чем для полного пространства. При этом объём описания, непосредственно влияющий на затраты памяти, сократился приблизительно в 1,2 раза.

Сравнительный экспериментальный анализ достоверности классификации для множественного и однозначного голосования структурных описаний приведен на рис. 4.7. Значение критерия достоверности θ служит показателем уверенности в принятии решения на основе максимума голосов, отданных за эталон. Глобальное решение в случае меньшего значения θ в большей степени значимо по отношению к локальному максимуму, который соответствует наиболее близкому из остальных (возможно, ложных) классов.

Как видим из рисунка 4.7, значение критерия достоверности классификации θ для нескольких эталонов базы при применении традиционного метода достигает значения $\theta = 0,8$, в то время как для предложенного в монографии метода на основе ППС $\theta = 0,02$. Такое низкое

значение критерия θ для разработанного метода подтверждает его высокий показатель правильной классификации для изучаемой базы видео-информации.

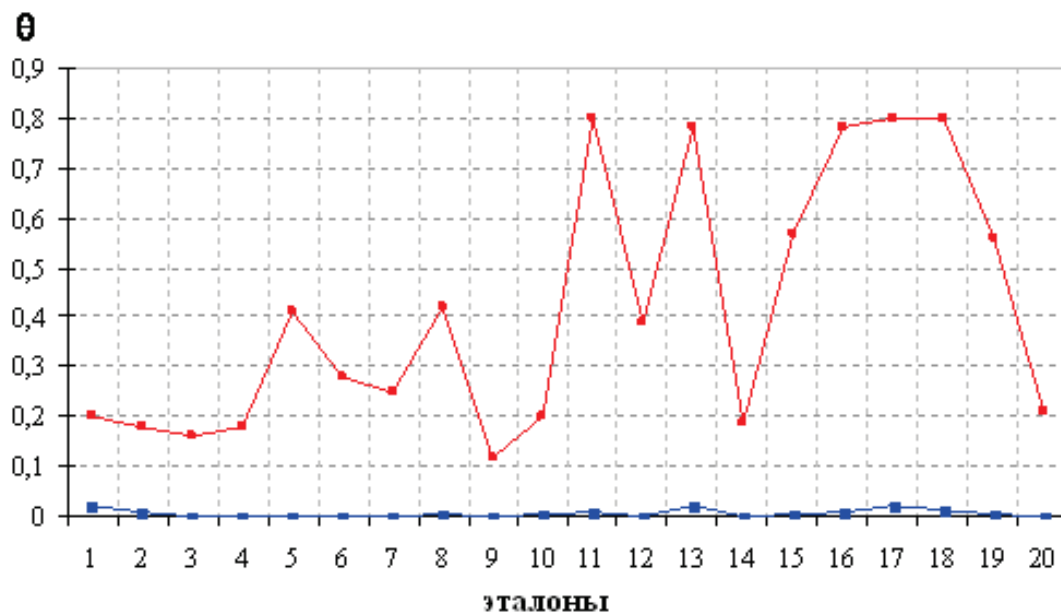


Рисунок 4.7 – Значения критерия θ : —■— множественное голосование с независимыми признаками, —■— однозначное голосование на основе ППС

Для анализа классификации при действии помех производилось моделирование препятствий путем имитации вероятностных переходов ХП объекта в окрестности их истинного местоположения через изменения их координат в пределах ± 1 , ± 2 , ± 3 с некоторой вероятностью p . Расположение точки внутри окрестности выбиралось в соответствии с равномерным распределением.

Наши эксперименты показали, что при отсутствии помех и до уровня $p = 0,2$ множественное голосование признаков, которые рассматриваются как независимые элементы, и однозначное голосование на основе ППС обеспечивают распознавание с вероятностью 1, включая действие помех для наибольшей окрестности искажений ± 3 . Начиная с величины $p = 0,35$, вероятность для однозначного распознавания изменяется следующим образом: при ± 1 составляет 1, при ± 2 — 1, при ± 3 снижается и равна 0,9. При $p = 0,5$ вероятность приобретает значения 0,8; 0,8; 0,6 для соответствующих окрестностей

искажения. При многозначном распознавании вероятность распознавания несколько ниже, например, при $p = 0,35$ она составляет: при $\pm 1 - 0,9$, $\pm 2 - 0,87$, $\pm 3 - 0,83$. Снижение можно объяснить процессом глаживаяния, происходящим при многозначном голосовании, а также тем, что значения ГИП объекта здесь могут проголосовать несколько раз за ложную модель.

Сравнительное исследование быстродействия рассматриваемых подходов показало, что оно возрастает в зависимости от уменьшения числа используемых инвариантов следующим образом. Например, при снижении числа точек с 17 до 13 число инвариантов уменьшается в 6 раз, а среднее время распознавания – в 11 раз, а при снижении числа базисов на 20% (с 14 до 12) быстродействие повышается в 6 раз. Эти результаты в целом подтверждают эффективность предложенного подхода. При этом время распознавания примерно одинаково для обеих модификаций.

В плане сравнения помехозащищенности методов отметим, что при таком существенном повышении быстродействия помехозащищенность при действии помех в окрестности ± 3 практически не изменяется. Путем прореживания точек и исключения близких базисов при фильтрации фактически происходит устранение избыточности имеющейся информации и обеспечивается надежное распознавание на сокращенном описании.

Анализ эффективности применения процедуры фильтрации F_1 (см. раздел 3.1), которая состоит в сокращении числа базисов на эталонном описании, показывает уверенное сохранение достаточного уровня достоверности. Суть F_1 заключалась в отборе не менее 30% базисов, в частности, выбирался каждый третий базис из полного множества.

Эксперименты показали, что в результате применения F_1 время сопоставления сократилось в 2,5 раза. Достоверность классификации по критерию θ с применением структурных моделей пространства ППС значительно улучшилась в сравнении с подходом независимого голосования. Для традиционного подхода (с голосованием полного множества ППС)

вычислена величина критерия $\theta = 0,78$, в то время как для подхода с использованием предложенной модели на базе процедуры F_1 значение критерия составило $\theta = 0,01$, что существенно меньше. Достоверность повышается за счет процедуры группового сопоставления элементов подмножеств объекта и эталона, а быстроедействие – за счет процедур фильтрации и принятия решения по части элементов эталона. Полученные в эксперименте данные говорят о существенных преимуществах обработки, предложенной в данном исследовании. Представление описаний видеообъектов в виде множеств ППС предназначено для учета в процессе классификации геометрических свойств и соотношений между составными частями описания.

Применение рассмотренных моделей для построения структурного пространства признаков визуальных объектов ориентировано, в первую очередь, на сокращение избыточности, что обеспечивает снижение объема вычислительных затрат. В то же время и достоверность на сокращенном описании в пространстве ППС оказывается лучше. Управление объемом структурного описания дает возможность применить на практике другие интеллектуальные подходы к анализу данных, например, связанные с визуализацией информации, с применением кластеризации и нейронных сетей для анализа и др.

Предложенные процедуры фильтрации F_1 , F_2 позволяют управлять размером признакового пространства в зависимости от требований прикладных задач. Такая гибкость может быть одним из путей решения проблемы хранения описаний видеоданных при создании визуальных баз знаний.

Применение разработанных моделей представления пространства геометрических инвариантных признаков, основанных на перегруппировке структурных элементов, обеспечивает повышение показателей достоверности и быстрогодействия классификации. На примере построенного пространства признаков получено экспериментальное подтверждение эффективности предложенных методов.

Полученные в работе методы могут быть успешно применены в прикладных задачах компьютерного зрения, например, при классификации и обнаружении объектов различного вида на географических и имитационных картах, в медицинских применениях для выделения и оценки параметров объектов на полутоновых и цветных изображениях, а также для объектов, состоящих из множеств отдельных точек. Понятно, что в большей степени разработанные подходы применимы в ситуациях, где эффективно функционируют современные методы выделения ХП (SIFT, SURF и др.), т.к. эффективность использования геометрических конструкций полностью определяется устойчивостью при определении ХП, из которых формируются вторичные ППС. В то же время возможно и применение предложенных методов для синтезированных множеств признаков, формируемых, например, в интерактивном режиме [121].

Разработанные метод и модели сопоставления описаний на основе ППС, представленных множеством ГИП, внедрены в строительной компании при идентификации архитектурных стилей строения или его эскиза, определяемых составными конструкциями (форма окна, дверей, крыши и т.п.). Это позволило частично автоматизировать решение задачи идентификации архитектурного стиля отдельных объектов строительства [123]. На этапе архитектурно-строительного проектирования использование предложенного метода на основе ППС обеспечивает возможность идентификации конечного набора архитектурных стилей на основе множества ГИП, построенного для сформированного множества ключевых точек.

Разработанные методы нашли также применение при выполнении исследований при создании базы эталонных описаний образцов технических средств внутренних войск и в технологиях поиска, а также классификации изображений технических средств на основе оптимального сопоставления геометрических структур данных. Это позволяет определить, подсчитать число и устранить дублирование эквивалентных графических объектов, а также предотвратить формирование избыточных структурных описаний с целью

повышения эффективности и полноты информационно-аналитического обеспечения деятельности. Здесь применен анализ синтезированных изображений.

Входное изображение представлено в виде графической электронной карты с нанесенными на нее активными точками – объектами мониторинга. Каждое проводимое мероприятие включает в себя ряд объектов, расположенных в различных частях карты, для которых необходимо учитывать ход выполнения мероприятия и число видов конкретных объектов, которые в нем задействованы.

Другая задача, которая может быть решена разработанными методами – мониторинг пространственного перемещения людей (анализируемых объектов) относительно неподвижных объектов (пунктов слежения за изменением ситуации), координаты которых выступают в качестве точек базисов при построении ППС.

По результатам исследований достигнуто повышение достоверности и быстродействия классификации объектов путем выполнения метода сопоставления визуальных описаний на основе ППС. Применение на практике описанных видов ППС обладает особыми свойствами, эффективными для решения конкретных задач, в то же время применение их в других задачах может быть нецелесообразным. Например, в случае исследований связанных с определением архитектурного стиля объектов, важным является построение ППС вокруг ХП как базовых элементов в определении конструкции зданий, которыми мы не можем пренебречь, поэтому разумно проводить сравнения между сформированными подмножествами ГИП, вычисленными в разных системах координат АБ, но принадлежащим одному ХП. Напротив, сопоставления изображений для отслеживания объектов мониторинга – активных точек на карте, относительно неподвижных пунктов визуального расположения войск рационально применять ППС, сформированные в системах координат неподвижных ХП, что непосредственно повышает оперативность системы классификации за счет рассмотрения малого числа АБ.

Результаты научных исследований, полученных в монографии, используются в учебном процессе кафедры информатики ХНУРЭ при проведении лекционных занятий и лабораторного практикума по дисциплинам «Основи комп'ютерної графіки», «Структурні методи інтелектуальної обробки відеоданих», «Аналіз даних», «Обробка зображень та мультимедіа».

Применение материалов исследований позволяет повысить степень современности, качество преподавания и углубить освоение содержания дисциплин при изучении и применении современных информационных технологий анализа и классификации изображений в интеллектуальных системах компьютерного зрения.

Выводы

1. Предложено применение модели вычисления подобия описаний на основе принципа формирования оптимальных групповых соответствий для определения максимального количества пар признаков объект-эталон в виде ППС на основе Венгерского метода, что повышает достоверность распознавания за счет однозначного и более точного учета подобия подмножеств описаний и обеспечивает минимум критерия суммарного расстояния между элементами описаний объекта и эталона. Экспериментальные исследования показывают предпочтительные характеристики распознавания для оптимального метода, где величина критерия ближайшего максимума для правильно опознанного класса относительно других эталонов не превышает 0,16, в то время как для традиционного метода она достигает 0,8, что говорит о целесообразности развития и применения оптимальных методов сопоставления множеств ППС в задачах компьютерного зрения.

2. Проведенные эксперименты по оценке эффективности разработанных методов подтвердили достаточно высокую их помехозащищенность, которая не уступает известным методам. Вероятность распознавания на основе

пространственных признаков при множественном голосовании снижается до 0,67, в то время как комплексированные методы сохраняют достаточно высокую вероятность 0,89 и 0,95 при этих же искажениях. Наиболее высокая вероятность классификации сохраняется в диапазоне 0,99–0,95 при применении ассоциативной меры подобия, которая учитывает признаки из различных пространств равноценным образом.

3. Сравнение эффективности классификации на основе предложенных моделей построения описания видео-объектов путем применения пространственных структур и процедур фильтрации с существующими моделями на базе полного множества независимых признаков показывает их значительное преимущество в плане достоверности и быстродействия. Эксперименты подтвердили надежность уверенного принятия решений для представленных моделей с использованием процедур сжатия, уровень подобия для ложных классов достигает 0,02 в сравнении с правильным классом за счет учета взаимных расположений и поиска верных соответствий между элементами подмножеств объекта и эталона. Снижение числа элементов на 20% повышает оперативность выполнения процедуры в 6 раз.

4. Внедрение результатов монографии в задачах идентификации на изображении архитектурных стилей строения или его эскиза, определяемых составными конструкциями, а также в технологиях поиска и классификации изображений технических средств подтверждает практическую ценность проведенных исследований.

ВЫВОДЫ

В монографии приведены результаты, которые в совокупности являются решением актуальной научно-практической задачи повышения эффективности классификации изображений за счет использования метода сопоставления структурных описаний на основе моделей пространственных признаков структур, гибридных моделей классификации на основе анализа разнотипных признаков и модели вычисления сходства описаний путем формирования оптимальных групповых соответствий их элементов. Авторами получены следующие научные и практические результаты.

1. Синтезирован новый метод сопоставления структурных описаний видео-объектов на основе построения моделей инвариантных признаков структур с применением голосования, который за счет использования в процессе вычисления сходства новых информационных связей в виде соответствия структур объектов усовершенствует процедуру классификации и обеспечивает более высокий уровень достоверности.

2. Разработана новая модель построения описания видео-объектов путем применения признаков пространственных структур, сгруппированных вокруг координат характерных признаков или аффинных базисов, с последующим сжатием множества структур, обеспечивает повышение показателей достоверности за счет доминирования при принятии решений группой связанных признаков над стандартными решениями, которые принимаются по отдельным элементам на множестве независимых признаков. В результате классификации с применением группировки элементов в виде признаков пространственных структур степень сходства для неправильно классифицированных эталонов не превышает 2%, в то время как для традиционных методов она достигает 80%. Совместное использование структурной группировки и фильтрации множества признаков приводит к значительному сокращению объема описания и вычислительной сложности сопоставления при необходимом для прикладных задач уровне вероятности классификации. В рассмотренных практических примерах объем описания уменьшился в 1,4 раза.

3. Усовершенствованы гибридные модели классификации изображений на основе анализа структурных разнотипных признаков, которые улучшают достоверность классификации за счет многостороннего и углубленного анализа состава описаний. Вероятность распознавания на основе пространственных признаков с применением множественного голосования при увеличении уровня воздействия помех снижается до 0,67, в то время как гибридные модели классификации сохраняют высокую вероятность 0,89 и 0,95 при этих же искажениях. Наиболее высокая вероятность классификации сохраняется в диапазоне 0,99-0,95 при применении ассоциативной меры сходства, которая учитывает признаки из разных пространств равноценным образом.

4. Получила дальнейшее развитие модель вычисления сходства описаний на основе принципа формирования оптимальных групповых соответствий для определения максимального количества пар признаков объект-эталон в виде инвариантных пространственных структур на основе Венгерского метода, которая способствует повышению достоверности распознавания за счет однозначного и более точного учета сходства подмножеств описаний и обеспечивает минимум критерия суммарного расстояния между описаниями объекта и эталона. Экспериментальные исследования показывают предпочтительные характеристики распознавания для оптимального метода, где степень сходства объекта со всеми возможными эталонами не превышает 16%, в то время как для обычного он достигает 80%, что говорит о целесообразности применения оптимальных методов сопоставления множеств в виде признаков пространственных структур в задачах компьютерного зрения.

5. Результаты работы нашли применение в технологиях поиска путем сопоставления и обработки описаний синтезированных изображений технических средств на основе признаков пространственных структур и при автоматизации решения прикладных задач определения архитектурного стиля здания или эскиза путем сопоставления описаний, построенных на основе структур геометрических инвариантных признаков ключевых точек изображений. Результаты использованы в учебном процессе Харьковского национального университета радиоэлектроники.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Путятин Е.П. Обработка изображений в робототехнике / Е.П. Путятин, С. И. Аверин. – М.: Машиностроение, 1990. – 320 с.
2. Ту Дж. Принципы распознавания образов / Дж. Ту, Р. Гонсалес. – М.: Мир, 1978. – 411 с.
3. Путятин Е.П. Нормализация изображений: достижения и нерешенные задачи / Е.П. Путятин // Бионика интеллекта. – 2007. – Вып. 1(66). – С. 5–16.
4. Шапиро Л. Компьютерное зрение / Л. Шапиро, Дж. Стокман.; [пер. с англ. А.А. Богуславского, под ред. С.М. Соколова]. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
5. Форсайт Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс.; [пер. с англ. А.В. Назаренко, И.Ю. Дорошенко]. – М.: Издательский дом «Вильямс», 2004. – 928 с.
6. Шлезингер М. Десять лекций по статистическому и структурному распознаванию / М. Шлезингер, В. Главач. – К.: Наук. думка, 2004. – 535 с.
7. Цвиркун А. Д. Основы синтеза структур сложных систем / А. Д. Цвиркун. – М.: Наука, 1982. – 198 с.
8. Duda R. O. Pattern classification / R. O. Duda, P. E. Hart, D. G. Stork. – 2 ed., Wiley, 2000. – 738 p.
9. Павлидис Т. Алгоритмы машинной графики и обработки изображений / Т. Павлидис. – М.: Радио и связь, 1986. – 400 с.
10. Машталир В. П. Точечно-множественные методы обработки информации / В. П. Машталир. – Харьков: Бизнес Информ, 2001. – 199 с.
11. Шабанов–Кушнарченко Ю. П. Теория интеллекта. Математические средства / Ю. П. Шабанов–Кушнарченко. – Харьков: Вища шк., 1984. – 144 с.
12. Гороховатский В.А. Структурный анализ и интеллектуальная обработка данных в компьютерном зрении: монография / В.А. Гороховатский. – Х.: Компания СМІТ, 2014. – 316 с.

13. Техническое зрение роботов; [под ред. А. Пью] [пер. с англ. Д.Ф. Миронова; под. ред. Г. П. Катуса]. – М.: Машиностроение. – 1987. – 320 с.
14. Прикладная статистика: Классификация и снижение размерности: справ. изд. / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин; под ред. С.А. Айвазяна. – М.: Финансы и статистика, 1989. – 607 с.
15. Дюк В.А. Обработка данных на ПК в примерах / В.А. Дюк. – СПб: Питер, 1997. – 240 с.
16. Lowe D.G. Object recognition from local scale-invariant features / D.G. Lowe // International Conference on Computer Vision. – Corfu, Greece, September 1999. – P. 1150–1157.
17. Башмаков А. И. Интеллектуальные информационные технологии: Учебное пособие / А.И. Башмаков, И.А. Башмаков. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2005. – 304 с.
18. By H. SURF: Speeded Up Robust Features / H. By, A. Ess, T. Tuytelaars, L. Gool // Computer Vision and Image Understanding (CVIU). – 2008. – № 110(3). – P. 346–359.
19. Миркин Б. Г. Анализ качественных признаков и структур / Б.Г. Миркин. – М.: Статистика, 1980. – 319 с.
20. Путятін Є.П. Методи та алгоритми комп'ютерного зору: навч. посіб./ Є.П. Путятін, В.О. Гороховатський, О.О. Матат. – Х.: ТОВ «Компанія СМІТ», 2006. – 236 с.
21. Гороховатский В.А. Распознавание изображений в условиях неполной информации / В.А. Гороховатский. – Харьков: ХНУРЭ, 2003. – 112 с.
22. Журавлев Ю.И. Распознавание образов и распознавание изображений / Ю.И. Журавлев, И.Б. Гуревич // Ежегодник «Распознавание, классификация, прогноз. Математические методы и их применение». – М: Наука, 1989. – 74 с.
23. Гороховатський В.О. Аналіз властивостей, характеристик та результатів застосування новітніх детекторів для визначення особливих точок зображення / В.О. Гороховатський, Д.В. Пупченко, К.Г. Солодченко // Системи управління, навігації та зв'язку. –2018. – №1 (47). – С. 93–98.

24. Яншин В. В. Анализ и обработка изображений: Принципы и алгоритмы / В.В. Яншин. – М.: Машиностроение, 1994. – 240 с.
25. Полякова Т.В. Класифікація зображень на основі ознакових просторових структур : дис... канд. техн. наук: 05.13.23 – системи та засоби штучного інтелекту / Тетяна Вікторівна Полякова; Харк. нац. ун-т радіоелектрон. – Харків, 2014 . – 130 с.
26. Гонсалес Р. Цифровая обработка изображений в среде MATLAB / Р. Гонсалес, Р. Вудс, С. Эддинс. – М.: Техносфера, 2006. – 616 с.
27. Гороховатский В.А. Формирование отношений аффинных базисов при распознавании по структурным признакам / В.А. Гороховатский, Н.Е. Долбня, А.Ю. Русаков // Прикладная радиоэлектроника. Состояние и перспективы развития (МРФ–2008): III Международный радиоэлектронный форум, Харьков, 22 – 24 октября 2008: сборник научных трудов. – Харьков, 2008. – С. 112–114.
28. Фу К.С. Структурные методы в распознавании образов / К.С. Фу. – М.: Мир, 1977. – 372 с.
29. Shapiro V.A. Real-time Hough/Radon transform: algorithm and architectures / V.A. Shapiro, V.H. Ivanov // Proceedings of ICIP'94. – 1994. – № 3. – P. 630 – 634.
30. Theodoridis S. Pattern Recognition / S. Teodoridis, K. Koutroumbus. – Amsterdam: Academic Press, 2003. – 681 p.
31. Раушенбах Г.В. Об измерениях близости между множествами в задачах кластер-анализа / В кн.: Статистика. Вероятность. Экономика. – М.: Наука, 1985. – С. 388–392.
32. Лбов Г.С. Методы обработки разнотипных экспериментальных данных / Г.С. Лбов. – Новосибирск: Наука. – 160 с.
33. Гороховатский В.А. Методы определения релевантности изображений на основе медианной обработки структурных описаний / В.А. Гороховатский // Радиоэлектроника, информатика, управление.–2017. – №1 (40). – С. 100–106.

34. Heikkila J. Multi-scale autoconvolution for affine invariant pattern recognition / J. Heikkila // In: Proceedings of ICPR02. – Qubec, Canada, 2002. – P. 119–122.
35. Гороховатский В.А. Применение самообучения в структурном распознавании изображений / В.А. Гороховатский, А. Е. Берестовский // Матер. XXIII міжн. конф. з автомат. управління (Автоматика–2016), м. Суми, 22–23 вересня 2016.– Суми: Сумський державний університет. – С. 192–193.
36. Lindeberg T. Feature detection with automatic scale selection / T. Lindeberg // International Journal of Computer Vision. – 1998. – № 30 (2). – P. 77–116.
37. Rublee E. ORB: an efficient alternative to SIFT or SURF / E. Rublee, V. Rabaud, K. Konolige, G. Bradski // Computer Vision (ICCV): IEEE International Conference on. IEEE. – 2011. - P. 2564 – 2571.
38. Гаганов В. Инвариантные алгоритмы сопоставления точечных особенностей на изображениях / В. Гаганов // Компьютерная графика и мультимедиа, 2009. – № 1 (17). – [Электронный ресурс]. – Режим доступа: http://cgm.computergraphics.ru/issues/issue17/invariant_features.
39. Leutenegger S. BRISK: Binary Robust Invariant Scalable Keypoints / S. Leutenegger, M Chli, R Siegwart // Computer Vision (ICCV). -2011. –P 2548–2555.
40. Rosten E. Machine learning for high-speed corner detection / E. Rosten, T. Drummond // 9th European Conference on Computer Vision (ECCV): proceedings. – 2006. –P. 430 – 443.
41. Flusser J. Affine moment invariants: a new tool for character recognition / J. Flusser, T. Suk // Pattern Recognition Letters. – 1994. – № 15. – P. 433–436.
42. Calonder M. BRIEF: Binary Robust Independent Elementary Features / M. Calonder, V. Lepetit, C Strecha, P Fua // 11th European Conference on Computer Vision (ECCV): proceedings. – 2010. - P. 778 – 792.
43. Bolles R. C. Recognizing and locating partially visible objects: The local-feature-focus method / R. C. Bolles, R. A. Cain // Internat. J. Robotics Res. – 1982. –№ 1 (3). – P. 57–82.

44. Балк М. Б. Геометрия масс / М. Б. Балк, В. Г. Болтянский. – М.: Наука, 1987. – 160 с.
45. Gorokhovatsky V. A. Image Likelihood Measures of the Basis of the Set of Conformities / V. A. Gorokhovatsky, Ye. P. Putyatin // Telecommunications and Radio Engineering. – 2009. – № 68 (9). – P. 763–778.
46. Постников М. М. Аналитическая геометрия / М. М. Постников. – М.: Наука, Главная редакция физико-математической литературы, 1973. – 751 с.
47. Гороховатский В.А. Повышение быстродействия метода аффинных инвариантов при сопоставлении структурных описаний визуальных объектов / В.А. Гороховатский, Т.В. Полякова // Системи управління, навігації та зв'язку: зб. наук. праць. – Київ, 2009. – № 2 (10). – С. 81–83.
48. Гороховатский В.А. Модели комплексированных мер подобия структурных описаний изображений / В.А. Гороховатский, Т.В. Полякова, Е.П. Путятин // Реєстрація, зберігання і обробка даних. – 2011. – Т. 13. – № 1. – С. 71–76.
49. Гороховатский В.А. Критерии и модели структурной классификации с применением принципа голосования / В.А. Гороховатский, Т.В. Полякова // Системи обробки інформації: зб. наук. праць. – Харків: ХУПС, 2011. – № 2 (92). – С. 12–16.
50. Гороховатский В.А. Оптимальные методы сопоставления структурных описаний видеообъектов / В.А. Гороховатский, Т.В. Полякова // Бионика интеллекта. – 2011. – № 3 (77). – С. 85–88.
51. Полякова Т.В. Анализ и сжатие структурных описаний на основе геометрических инвариантных признаков / Т.В. Полякова // Прикладная радиоэлектроника. – 2012. – Т. 11. – № 3. – С. 417–420.
52. Полякова Т.В. Классификация изображений на основе признаков пространственных структур / Т.В. Полякова, Ю.А. Куликов, В.А. Гороховатский // Вісник Національного технічного університету «Харківський політехнічний інститут»: зб. наук. праць. – Харків: НТУ «ХПІ», 2013. – № 2 (976). – С. 101–109. – Серія «Системний аналіз, управління та інформаційні технології».

53. Gorokhovatskiy V.A. Geometrical Invariant Features Peculiar for the Methods of Structural Classification of Images / V.A. Gorokhovatskiy, T.V. Poliakova // Telecommunications and Radio Engineering. – 2012. – № 71 (17). – P. 1557–1564.

54. Полякова Т. В. Исследование методов сопоставления структурных объектов / Т. В. Полякова // Радіоелектроніка і молодь в ХХІ ст.: 13-й міжнародний молодіжний форум, Харків, 30 березня – 1 квітня 2009 р.: збірник матеріалів. форуму. – Харків: ХНУРЕ, 2009. – С. 337.

55. Гороховатський В. Стиснення описів при застосуванні структурно-ієрархічних методів розпізнавання зображень / В. Гороховатський, Р. Гребінник, Т. Полякова // Комп'ютерні науки та інформаційні технології (CSIT2009): IV міжнародна науково-технічна конференція, Львів, 15–17 жовтня 2009 р.: матеріали конференції. – Львів: НУ «Львівська політехніка», 2009. – С. 79–81.

56. Гороховатский В.А. Метод структурного распознавания изображений с использованием разнотипных признаков / В.А. Гороховатский, Т.В. Полякова // Математичне та програмне забезпечення інтелектуальних систем (MPZIS–2009): VII міжнародна науково-практична конференція, Дніпропетровськ, 25–27 листопада 2009 р.: тези доповідей. – Дніпр-ськ: ДНУ, 2009. – С. 71–72.

57. Гороховатский В.А. Модели признаковых пространств в структурных методах распознавания изображений / В.А. Гороховатский, Ю.А. Куликов, Т.В. Полякова // Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту (ISDMCI'11): міжнародна наукова конференція, Євпаторія, 16-20 травня 2011 р.: матеріали конференції. – Херсон: ХНТУ, 2011. – С. 226–230.

58. Полякова Т.В. Структурная обработка множеств инвариантных геометрических признаков при распознавании видео-объектов / Т.В. Полякова // Актуальні проблеми технічного забезпечення внутрішніх військ МВС України: науково-практичний семінар, Харків, 25 квітня 2012 р.: матеріали семінару. – Харків: АБВ МВС України, 2012. – С. 44 – 48.

59. Полякова Т.В. Групповые представления и анализ описаний видео-объектов в виде множеств структурных геометрических признаков /

Т.В. Полякова, В.А. Гороховатский // Интеллектуальні системи прийняття рішень і проблеми обчислювального інтелекту (ISDMCI'2012): міжнародна наукова конференція, Євпаторія, 27–31 травня 2012 р.: матеріали конференції. – Херсон: ХНТУ, 2012. – С. 398–400.

60. Полякова Т.В. Оптимальные меры подобия структурных описаний изображений / Т.В. Полякова, В.А. Гороховатский // Математичне та програмне забезпечення інтелектуальних систем (MPZIS–2012): X ювілейна міжнародна науково-практична конференція, Дніпропетровськ, 21-23 листопада 2012 р.: тези доповідей. – Дніпр-ськ: ДНУ, 2012. –С. 244–245.

61. Ситников Д. Э. Нахождение минимизированных наборов признаков для описания дискретных объектов в базах данных [Текст] / Д.Э. Ситников, О.А. Романенко, Е.В. Титова // Системи обробки інформації. – 2009. – №. 4(78). – С. 122–126.

62. Mikolajczyk K. Scale and affine invariant interest point detectors / K. Mikolajczyk, C. Schmid. // International Journal of Computer Vision. – 2004. – № 60. – P 63–86.

63. Горелик А. Л. Современное состояние проблемы распознавания / А.Л. Горелик, И.Б. Гуревич, В. А. Скрипкин. – М.: Радио и связь, 1985. – 160 с.

64. Рассел С. Искусственный интеллект: современный подход; 2-е изд./ С. Рассел, П. Норвиг; пер. с англ. – М.: Изд. дом «Вильямс», 2006. – 1408 с.

65. Пападимитриу Х. Комбинаторная оптимизация. Алгоритмы и сложность / Х. Пападимитриу, К. Стайглиц. – М.: Мир, 1985. – 512 с.

66. Баклицкий В. К. Методы фильтрации сигналов в корреляционно-экстремальных системах навигации / В. К. Баклицкий, А. М. Бочкарев, М. П. Мусьяков. – М.: Радио и связь, 1986. – 216 с.

67. Rahtu E. Object classification with multi-scale autoconvolution / E. Rahtu, J. Heikkila // Pattern Recognition (ICPR'04): International Conference, Cambridge, UK, August, 2004: proceedings. – Cambridge, 2004. – № 3. – P. 37–40.

68. Гольштейн Е. Г. Задачи линейного программирования транспортного типа / Е. Г. Гольштейн, Д. Б. Юдин. – М.: Наука, 1969. – 382 с.

69. Lamdan Y. Object recognition by affine invariant matching / Y. Lamdan, S.T. Schwartz, H.J. Wolfson // Computer Vision and Pattern Recognition: IEEE Conference, Ann Arbor, June, 1988: proceedings. – New York, 1988. – № 136. – P. 335–344.
70. Lamdan Y. Geometric hashing: A general and efficient model-based recognition scheme / Y. Lamdan, H.J. Wolfson // Computer Vision: Second International Conference, Tampa, Florida, 5–8 December, 1988: proceedings. – New York, 1988 – P. 238–249.
71. He F. Faster affine invariant point pattern matching using graphical models / F. He, Z. Tian, X. He, L. Yang, Ch. Chai // Computer Science and Automation Engineering (CSAE): International Conference, Shanghai, China, 10–12 June, 2011: proceedings. –2011. – № 4. – P. 32–36.
72. Petrou M. Affine invariant features from the trace transform / M. Petrou, A. Kadyrov // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2004. – № 26. – P. 30–44.
73. Rahtu E. Convexity recognition using multi-scale autoconvolution / E. Rahtu, M. Salo, J. Heikkila // Pattern Recognition: 17th International Conference, Cambridge, UK, August, 2004: proceedings. –Cambridge, 2004.
74. Воробель Р. А. Логарифмична обробка зображень / Р. А. Воробель. – К.: Наук. думка. – 2012. – 231 с
75. Гайдышев И. Анализ и обработка данных: специальный справочник / И. Гайдышев. – СПб.: Питер. – 2001.
76. Gope C. Affine invariant comparison of point-sets using convex hulls and hausdorff distances / C. Gope, N. Kehtarnavaz // Pattern Recognition. – 2007. – № 40. – P. 309–320.
77. Voss K. Affine Point Pattern Matching / K. Voss, H. Suesse // Lecture Notes in Computer Science. – Munich, 2001. – № 2191. – P. 155–162.
78. Weiss I. Geometric invariants and object recognition / I. Weiss // International Journal of Computer Vision. – 1993. – № 10. – P 207–231.

79. Шлезингер М. И. Математические средства обработки изображений / М. И. Шлезингер. – К.: Наук. думка, 1989. – 200 с.
80. Путятин Е.П. Проблема распознавания изображений в системах технического зрения / Е.П. Путятин // Проблемы бионики. – 1999. – Вып. 50. – С. 3–11.
81. Патин М. В. Сравнительный анализ методов поиска особых точек и дескрипторов при группировке изображений по схожему содержанию / М.В. Патин, Д. В. Коробов // Молодой ученый. – 2016. – №11. – С. 214–221. [Электронный ресурс]. – Режим доступа: <https://moluch.ru/archive/115/31188/>.
82. Машталир В. П. Индуцированная согласованность отношений в задачах грануляции информации / В. П. Машталир, В. В. Шляхов // Бионика интеллекта. – 2006. – № 1 (64). – С. 19–26.
83. Раушенбах Г. В. Меры близости и сходства / В кн.: Анализ нечисловой информации в социологических исследованиях. – М.: Наука, 1986. – С. 169–203.
84. Далека В.Д. Модели и структуры данных. Учебное пособие / В.Д. Далека, А.С. Деревянко, О.Г. Кравец, Л.Е. Тимановская // – Харьков: ХГПУ, 2000. – 241 с.
85. Гороховатский В.А. Структурное распознавание изображений на основе моделей голосования признаков характерных точек / В.А. Гороховатский, Е.П. Путятин // Реєстрація, зберігання і обробка даних. – 2008. – Т. 10. – № 4. – С. 75–85.
86. Lei B.J. On feature extraction from images / B.J. Lei, E.A. Hendriks, M.J.T. Reinders // Technical Report, Deliverable 2.1.1.2A+B, MCCWS project, Information and Communication Theory Group. TU Delf. – 1999. – 57 p.
87. Zhang D.S. Shape Retrieval Using Fourier Descriptors / D.S. Zhang, G.J. Lu Fargo // Multimedia and Distance Education: International Conference, Fargo, 1–3 June, 2001: proceedings. – USA, 2001. – P. 312–321.
88. Jain A.K. Algorithms for Clustering Data / A.K. Jain, R.C. Dubes. – Prentice Hall. – 1988. – P. 163–178.

89. Columbia Object Image Library (COIL-20):[Электрон. ресурс].– Режим доступа: <http://www1.cs.columbia.edu/CAVE/software/softlib/coil-20.php>
90. Ling H. Deformation invariant image matching / H. Ling, D. Jacobs // Computer Vision: Tenth IEEE International Conference, Beijing, China, 17-20 October, 2005: proceedings. – 2005. – P. 1466–1473.
91. Tuytelaars T. Matching widely separated views based on affine invariant regions / T. Tuytelaars, L. Van Gool // International Journal of Computer Vision. – 2004. – № 59. – P. 61–85.
92. Ruiz-del Solar J. A new approach for fingerprint verification based on wide baseline matching using local interest points and descriptors / J. Ruiz-del Solar, P. Loncomilla, C. Devia // Advances in Image and Video Technology, Lecture Notes in Comput. Sci. – Springer-Verlag, Berlin, 2007, – № 4872. – P. 586–599.
93. Harris C. A Combined Corner and Edge Detector / C. Harris, M. Stephens // 4th Alvey Vision Conference: proceedings. –1988. - P. 147 – 151.
94. Shalkoff R. J. Digital image processing and computer vision / R. J. Shalkoff. – John Wiley & Sons, New York, 1989. – 489 p.
95. Gold S. New algorithms for 2 D and 3 D point matching: Pose estimation and correspondence / S. Gold, A. Rangarajan, C. P. Lu, S. Pappu, E. Mjolsness // Pattern Recognition. – 1998. –№ 31. – P. 1019–1031.
96. Степаненко А.С. Оценки ускорения вычислений гибридными системами / А.С. Степаненко // Параллельные вычисления и задачи управления: V Международная конференция, Москва, 27-29 октября, 2010: материалы конференции. – М.: ИПУ РАН, 2010. – С. 29–38.
97. Томас Х. Алгоритмы: построение и анализ, 2-е издание / Х. Томас, И. Чарльз, Л. Рональд. – М.: Вильямс. – 2005.
98. Ясницкий Л.Н. Введение в искусственный интеллект: учеб. пособие для ВУЗов / Л.Н. Ясницкий. – М.: Изд. центр «Академия», 2005. – 176 с.
99. Страуструп Бьерн. Программирование: принципы и практика использования C++: исправленное издание / Бьерн Страуструп. – М.: «Вильямс», 2011. – 1248 с.

100. Кнут Дональд. Искусство программирования. Основные алгоритмы, том 1. 3-е изд. / Дональд Кнут. – М.: Вильямс, 2006. – 720 с.
101. Carneiro G. Phase-based local features / G. Carneiro, A. Jepson // In European Conference on Computer Vision. – Copenhagen, Denmark, 2002. – P. 282–296.
102. Torr P. H. S. Feature based methods for structure and motion estimation / P.H.S. Torr, A. Zisserman // In Vision Algorithms: Theory and Practice. – Corfu, Greece, 1999. – P. 278–294.
103. Russmussen C. Texture-based vanishing point voting for road shape estimation / C. Russmussen // British Machine Vision Conference, Kingston, UK, September, 2004: proceedings. – UK, 2004.
104. Suesse H. A Novel Approach for Affine Point Pattern Matching / H. Suesse, W. Ortmann, K. Voss // In: ICIAR 2. – 2006. – № 4142. – P. 434–444.
105. Sprinzak J. Affine point matching / J. Sprinzak, M. Werman // Pattern Recognition Letters. – 1994. – № 15. – P. 337–339.
106. Costa M. Optimal Affine-Invariant Point Matching / M. Costa, R. Haralick, L. Shapiro // SPIE Conference on Applications of Artificial Intelligence, Orlando, Florida, 1989: proceedings. –USA, 1989 – P. 515–530.
107. Szeliski R. Computer Vision: Algorithms and Applications / R. Szeliski. – Springer. – 2010. – 979 p.
108. Bishop C. Pattern Recognition and Machine Learning / C. Bishop. – Springer. – 2007. – 738 p.
109. Малла С. Вэйвлеты в обработке сигналов / С. Малла. – М.: Мир. – 2005. – 672 с.
110. Зарубежная радиоэлектроника / гл. ред. М. К. Размахнин. – М.: Радио и связь, 1987. – № 10. – 96 с.
111. Ситников Д.Э. Семантические свойства информативности ассоциативных зависимостей между признаками объектов в базах данных / Д.Э. Ситников, С. В. Титов, Е. В. Титова // Системи обробки інформації. – 2011. – № 8(98). – С. 18–21
112. Ballard D. H. Generalizing the Hough transform to detect arbitrary shapes/ D. H. Ballard // Pattern recognition. – 1981. – № 13(2) – P. 111–122.

113. Методы компьютерной обработки изображений / [под ред. В.А. Сойфера]. – М.: Физматлит, 2003. – 784 с.
114. Финогеев А.Г. Методика распознавания изображений на основе случайных деревьев в системах автоматизированного проектирования расширенной реальности [Электронный ресурс] / А.Г. Финогеев, М.В. Четвергова // Современные проблемы науки и образования. – 2012. – № 5. – Режим доступа к журн.: <http://www.science-education.ru/105-7110>.
115. Елисеева И.И. Группировка, корреляция, распознавание образов (Статистические методы классификации и измерения связей) / И.И. Елисеева, В.О. Рукавишников. – М.: Статистика, 1977. – 144 с.
116. Беклемишев Д.В. Аналитическая геометрия и линейная алгебра / Д.В. Беклемишев. – М: Высш. шк., 1998. – 320 с.
117. Градосельская Г.В. Сетевой анализ как метод исследования современных трансформаций / Г.В. Градосельская // Социологические опыты: сб. научных статей. – М.: Институт социологии РАН, 2002. – С. 29–40.
118. Ахо А.В. Структуры данных и алгоритмы / А.В. Ахо, Д.Э. Хопкрофт, Д.Д. Ульман. – М.: Вильямс, 2007. – 384 с.
119. Kaula W.M. Theory of Satellite Geodesy / W.M. Kaula. – Waltham, Massachusetts, Blaisdell Publishing Company, 1966.
120. Feasibility of Displacement Monitoring using Low-cost GPS Receivers / H. Jo, S.-H. Sim, A. Takkowski and other // Journal of Structural Control and Health Monitoring. – 2013. – № 20 (9). – P 1240 –1254.
121. Блинова Е.К. Модель семантического архитектурного пространства как инструмент оценки образов ордерных ансамблей / Е.К. Блинова // Известия РГПУ им. А.И. Герцена. – 2010. – № 137. – С. 124–134.
122. Месарович М. Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, И. Такахара. – М: Мир, 1973. – 344 с.
123. Трошкіна О.А. Семантика форми в дизайні представницької архітектури / О.А. Трошкіна // Технічна естетика і дизайн: Міжвідомчий науково-технічний збірник. – № 3–4. – К. : Віпол, 2004. – С. 150–153.

Наукове видання

**Гороховатський Володимир Олексійович,
Полякова Тетяна Вікторівна**

ЗАСТОСУВАННЯ ПРОСТОРОВИХ СТРУКТУР ОЗНАК ДЛЯ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ У КОМП'ЮТЕРНОМУ ЗОРІ

Монографія

Рецензенти:

Путятін Є.П., д-р техн. наук, професор, завідувач кафедри інформатики,
Харківський національний університет радіоелектроніки

Воробель Р.А., д-р техн. наук, професор, завідувач відділу
обчислювальних методів і систем перетворення інформації, Фізико-механічний
інститут ім. Г.В. Карпенко НАН України, м. Львов

Підп. до друку 10.04.18.
Умов. друк. арк. 7,0.
Ціна договірна

Формат 60x84 ¹/₁₆.
Облік. вид. арк. 6,4.

Спосіб друку – ризографія.
Тираж 100 прим.

Віддруковано в типографії ФОП Андреев К.В.

61166, Харків, вул. Серпова, 4,

Свідоцтво про державну реєстрацію №24800170000045020 від 30.05.2003 р.

ep.zakaz@gmail.com

тел. 063-993-62-73