

РОСТОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

ШАПОШНИКОВ ДМИТРИЙ ГРИГОРЬЕВИЧ

ИССЛЕДОВАНИЕ И МОДЕЛИРОВАНИЕ
БИОНИЧЕСКИХ ПРИНЦИПОВ ИДЕНТИФИКАЦИИ И
КОНТЕКСТНОГО ОПИСАНИЯ ИЗОБРАЖЕНИЙ

05.13.18 - Математическое моделирование, численные
методы и комплексы программ

А в т о р е ф е р а т
диссертации на соискание ученой степени
кандидата технических наук

Ростов-на-Дону
2002

Работа выполнена в НИИ нейрокибернетики им. А.Б. Когана
Ростовского государственного университета

Научный руководитель: доктор биологических наук,
профессор Владимирский Б.М.

Официальные оппоненты: доктор технических наук,
профессор Божич В.И.
доктор физико-математических наук,
кандидат технических наук,
с.н.с. Покровский А.Н.

Ведущая организация: Институт прикладной физики РАН
(г. Нижний Новгород)

Защита состоится 26 декабря 2002 г. в 11⁰⁰ часов на заседании
диссертационного совета К.212.208.04 по физико-математическим и
техническим наукам при Ростовском госуниверситете по адресу:
344090, г. Ростов-на-Дону, пр. Стачки, 200/1, корпус 2, ЮГИНФО РГУ.
С диссертацией можно ознакомиться в научной библиотеке РГУ по
адресу: ул. Пушкинская, 148.

Автореферат разослан 22 ноября 2002 г.

Ученый секретарь диссертационного совета,
кандидат физико-математических наук

Муратова Г.В.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследования. Важной проблемой при создании технических систем обработки визуальной информации является инвариантное распознавание изображений по отношению к изменению масштаба, повороту, условиям освещения, уровню сигнал-шум и т.д. [Балухто А. Н., Галушкин А. И., 2002; Журавлев Ю.И., 1998; Ben-Arie J., Nandy D., 1998; Furmanski C. S., Engel S. A., 2000]. В рамках этой проблемы можно выделить несколько принципиальных аспектов, связанных с разработкой алгоритмов, устойчивых к широкому диапазону преобразований изображений, поиском способов для необходимой и достаточной детализации описания изображений и минимизацией вычислительных затрат.

В настоящее время при решении задач сжатия зрительной информации и инвариантного распознавания изображений интенсивно развивается подход, основанный на разработке вычислительных алгоритмов, имитирующих принципы работы реальных зрительных систем (бионический подход), который рассматривается как наиболее перспективный. В рамках бионического подхода особое внимание уделяется разработке алгоритмов и методов определения наиболее информативных областей (**НИО**) изображений для детальной обработки, как аналогов биологических механизмов выбора перцептуально важных фрагментов при осмотре изображений [Hjelmås E., Kee B., 2001; Huang J., Wechsler H., 1999]. Очевидно, что детальная обработка не всего изображения, а отдельных его фрагментов, может существенно уменьшить вычислительные затраты и увеличить эффективность распознавания. Однако, имеющиеся в настоящее время факты и представления о механизмах зрительного восприятия в биологических системах, неполны и недостаточно формализованы для того, чтобы быть детально воспроизведенными в системах искусственного зрения. Кроме того, известные в литературе методы идентификации НИО не обладают свойством инвариантности и/или требуют большого объема вычислений [Wechsler H., 2002]. Поэтому решение задач поиска эффективных алгоритмов детектирования НИО для разных типов изображений по-прежнему остается актуальным.

Принципиально важным остается также и поиск алгоритмов, обеспечивающих сжатое и вместе с тем специфичное описание фрагментов изображений при сохранении инвариантности к различным трансформациям. Одним из широко используемых методов инвариантного описания является представление изображения на нескольких уровнях разрешения [Schwartz E.L. et al, 1995; Берт П.Д.,

1988]. Однако на низких уровнях разрешения может происходить потеря специфических деталей изображения. С другой стороны, алгоритмы, использующие детальное описание каждого фрагмента изображения, становятся очень чувствительными к различным искажениям и требуют больших вычислительных затрат.

Данная работа преемственно связана с исследованиями, проведенными ранее в лаборатории нейроинформатики сенсорных и моторных систем НИИ нейрокибернетики им. А.Б. Когана РГУ, в частности с поведенческой моделью распознавания полутонных изображений [Гусакова В.И. и др., 1994; Rybak I.A. et al, 1998], и направлена на поиск эффективных алгоритмов идентификации, классификации, описания и распознавания изображений различных классов на основе свойств их наиболее информативных областей.

Особенности подхода, использованного при решении этих проблем, состоят в следующем: тесная связь модельных, нейрофизиологических и психофизических исследований; использование бионических принципов для увеличения эффективности распознавания и минимизации вычислительных процедур; применение в качестве первичных одних и тех же локальных признаков и их комбинаций, выделение которых не требует сложных вычислений.

Цель и задачи исследования. Целью работы является разработка и исследование бионических алгоритмов и методов идентификации, классификации и контекстного описания информативных областей изображений, а также создание модели распознавания образов и применение ее в задачах инвариантного распознавания изображений лиц и дорожных знаков.

В процессе работы решались следующие задачи:

1. Разработка алгоритмов идентификации и контекстного описания НИО изображений различных типов.
2. Разработка алгоритмов классификации изображений на основе геометрических свойств НИО.
3. Тестирование разработанных алгоритмов в вычислительных экспериментах для выбора их параметров и оценки диапазона инвариантности к различным преобразованиям изображений.
4. Разработка программно-аппаратного стенда и проведение верификационных психофизических экспериментов по исследованию возможных механизмов выбора точек фиксации взгляда.
5. Создание модели распознавания образов и ее тестирование на примере изображений лиц и дорожных знаков.

Научная новизна результатов исследования. Разработан комплекс новых эффективных алгоритмов и методов идентификации НИО и

классификации на основе их свойств изображений различных типов. Впервые показано, что для НИО характерны специфические комбинации первичных признаков, таких, например, как разноориентированные локальные перепады яркости, выделенные на разных уровнях разрешения. Алгоритм классификации в соответствии с видом изображений, их фрагментов и ракурсом основан на оценке распределения ориентированных элементов, выделенных в пределах НИО, и на оценке пространственного расположения этих областей.

Разработана оригинальная модель инвариантного распознавания изображений лиц и дорожных знаков, основанная на специфичном описании НИО в виде признаков векторов, формирующихся с помощью пространственно-неоднородного сенсора с контекстным описанием признаков в каждом узле.

На качественном уровне получено экспериментальное подтверждение гипотез компьютерного моделирования о наличии локальной пространственной неоднородности остроты зрительного восприятия в периферическом поле зрения человека. Результаты, полученные в психофизических экспериментах, использованы при разработке алгоритма контекстного описания изображений.

Практическая значимость. Разработанные алгоритмы, модели и программы могут быть использованы для решения различных прикладных задач инвариантного распознавания изображений лиц и дорожных знаков, в частности, в системах идентификации личности и системах автоматического управления автомобилем.

Результаты работы использованы при выполнении гранта РФФИ № 98-01-00948 и проекта российско-английского сотрудничества между НИИ нейрокибернетики РГУ и Мидлсекским университетом (Великобритания) "Разработка бионических моделей зрения для распознавания реальных изображений".

По результатам имитационных экспериментов разработаны рекомендации по модификации геометрических свойств предупредительных дорожных знаков.

Апробация работы. Результаты работы представлялись на Международной конференции по нейронным сетям "IJCNN'99" в 1999 г. (г. Вашингтон, США); на Международном совещании по возможностям нейрокомпьютинга в 1999 г. (г. Вашингтон, США); на двух Международных конференциях по нейрокибернетике в 1999 г. и 2002 г. (г. Ростов-на-Дону, Россия); на 5-й и 6-й Международных конференциях "Распознавание образов и анализ изображений" в 2000 г. и 2002 г. (г. Самара, г. Великий Новгород, Россия); на 1-й, 2-й и 4-й Всероссийской конференции "Нейроинформатика" в 1999 г., 2000 г. и

2002 г. (г. Москва, Россия); на семинаре РАСНИ в 2001 г. (г. Москва, Россия); на VIII Всероссийской конференции "Нейрокомпьютеры: разработка и применение" в 2002 г. (г. Москва, Россия); на конференциях молодых ученых в 1996-1999 г. (г. Ростов-на-Дону, Россия); на школе-семинаре "Нейроинформатика, нейрокибернетика, нейрокомпьютеры" по программе "Интеграция" в 1998г. (г. Ростов-на-Дону, Россия); на Российско-финской зимней школе-семинаре "Information transfer, data and bio-organisms: from language to behavior" в 2000г. (г. Хельсинки, Финляндия); на семинаре Лаборатории вычислительного эксперимента в 2002 г. (ЮГИНФО РГУ, г. Ростов-на-Дону); на заседании Санкт-Петербургского физиологического общества им. И.М. Сеченова в 2002 г. (Институт физиологии им. И.П. Павлова РАН, г. Санкт-Петербург).

Публикации. По теме диссертации опубликовано 34 работы, из них 2 в международных журналах, 2 в российских журналах, 18 в тезисах международных конференций, 9 в тезисах российских и студенческих конференций, 2 в различных сборниках и 1 депонирована в ВИНТИ. Основные 7 публикаций, содержащие наиболее полные результаты, приведены в конце автореферата.

Структура диссертации. Работа состоит из введения, четырех глав, заключения, списка литературы, двух приложений и изложена на 121 странице, содержит 28 рисунков и 5 таблиц; список литературы включает 158 наименований.

Положения, выносимые на защиту.

1. Алгоритм признакового описания, состоящий в представлении фрагмента изображения в виде вектора локальных признаков, заполняемого с помощью пространственно-неоднородного сенсора с учетом свойств признаков, выделенных в контекстной области каждого его узла.
2. Каскадный метод идентификации НИО, включающий комплекс последовательно реализуемых алгоритмов и процедур, основанных на выделении специфических комбинаций первичных признаков, характерных для этих областей.
3. Модель распознавания изображений лиц и дорожных знаков, основанная на идентификации НИО и их контекстном описании с помощью пространственно-неоднородного сенсора.

СОДЕРЖАНИЕ

Во введении обоснована актуальность работы, определена цель исследований, отмечена новизна результатов, их научная и

практическая ценность, сформулированы основные положения, выносимые на защиту.

Первая глава, состоящая из 3-х разделов, посвящена аналитическому обзору литературы по проблеме инвариантного описания и распознавания изображений.

В первом разделе описываются общие подходы к решению проблемы инвариантного распознавания изображений. Эти подходы можно условно разделить на два основных класса. В подходах первого класса решение ищется на пути оптимизации стандартных для систем искусственного зрения вычислительных процедур и алгоритмов. К данному подходу относятся, известные классические методы: алгебраические, статистические, потенциальных функций, главных и независимых компонент, использование масок и пр. Проанализированы ограничения стандартных методов при распознавании изображений, полученных в реальных условиях.

В основе подхода второго класса лежит разработка вычислительных алгоритмов и моделей, имитирующих принципы работы реальных зрительных систем (бионический подход). Эти подходы развиваются в нескольких направлениях, различающихся по степени биологической обоснованности, вычислительным алгоритмам, архитектуре систем и пр. Основным их преимуществом по сравнению со стандартными методами является резкое уменьшение объема обрабатываемой информации за счет того, что детально описываются только информативные фрагменты изображения. Однако по ряду причин известные в настоящее время бионические алгоритмы также не обеспечивают приемлемого решения проблемы. В конце раздела делается вывод о том, что поиск решения проблемы инвариантного представления и распознавания изображений должен проводиться с помощью сочетания эффективных алгоритмов, разработанных в рамках обоих подходов.

Во втором разделе представлены известные алгоритмы описания изображений и их отдельных фрагментов. Особое внимание уделено алгоритмам пространственно-неоднородного и контекстного описания изображений. В основе этих алгоритмов лежит тот факт, что в реальной зрительной системе плотность фоторецепторов сетчатки максимальна в центральной области (фовеа) и убывает к периферии, в то время как размер рецептивных полей нейронов, напротив, увеличивается. Сочетание представления изображения на разных уровнях разрешения и контекстного описания отдельных его фрагментов рассматривается в качестве наиболее эффективного способа для обеспечения одновременно инвариантности и

специфичности описания изображений. Анализируются существующие методы и алгоритмы идентификации фрагментов изображения для детального анализа. Показано, что многие алгоритмы и методы либо детектируют значительное количество потенциальных "точек интереса" вне НИО, либо не обеспечивают инвариантной идентификации НИО при изменении ракурса и размера. С другой стороны, для алгоритмов, имеющих высокую детекторную точность, авторы подчеркивают их вычислительную сложность и необходимость включения механизмов внимания и обучения как существенных компонент.

В третьем разделе на основе проведенного аналитического обзора рассматривается постановка задач данного исследования. Обосновывается необходимость разработки эффективных алгоритмов и методов идентификации информативных областей и специфического описания изображений, полученных в реальных условиях, для решения проблем инвариантного распознавания.

Вторая глава, состоящая из 3-х разделов, посвящена описанию разработанных алгоритмов контекстного описания изображений, обеспечивающих специфичное представление фрагментов изображения. Приводятся результаты собственных психофизических экспериментов, позволившие разработать алгоритм контекстного описания признаков фрагментов изображения в каждом узле пространственно-неоднородного сенсора.

В первом разделе описывается математический аппарат предварительной обработки изображения. Исходное изображение рассматривалось как двумерный массив размера $X \times Y$, каждым $I(i, j)$ элементом которого является его яркость. Для представления изображения на разных уровнях разрешения исходное изображение подвергалось стандартной процедуре [Марр Д., 1987; Берт П.Д., 1988] последовательной фильтрации с помощью преобразования Гаусса:

$$I^l(i, j) = \sum_{p, q} \left(g(p, q) \cdot I^{l-1}(i + p, j + q) \right), \quad (1)$$

где $g(p, q) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{p^2 + q^2}{2\sigma^2}\right)$; $p, q = -2, -1, 0, 1, 2$; i, j - координаты

точки изображения; l - уровень разрешения ($l=1, 2, \dots, m$); при этом исходное изображение рассматривалось как изображение наибольшего разрешения.

Для выделения ориентированных перепадов яркости и построения первичных ориентационных $O(i, j)$ и контрастных $E(i, j)$ карт

использовалась разность ориентированных гауссианов со сдвинутыми центрами:

$$G^+(i, j, a) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(i - v_a)^2 + (j - \omega_a)^2}{2\sigma^2}\right) \quad (2)$$

$$G^-(i, j, a) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(i + v_a)^2 + (j + \omega_a)^2}{2\sigma^2}\right)$$

где $\omega_a = \sin\left(\frac{2\pi a}{16}\right)$ $v_a = \cos\left(\frac{2\pi a}{16}\right)$, σ - параметр

преобразования Гаусса.

Ориентация перепадов яркости a определялась с шагом 22.5° и кодировалась как $0, 1, \dots, 15$.

Второй раздел главы посвящен описанию результатов психофизических экспериментов, которые были направлены на верификацию предположений [Podladchikova L.N. et al, 1997] имитационных экспериментов о возможных механизмах контекстного кодирования зрительной информации и наличии локальной пространственной неоднородности остроты восприятия в периферическом поле зрения. Описаны разработанная методика и программно-аппаратный стенд для проведения психофизических экспериментов по определению пространственной динамики одного из параметров зрительного восприятия - остроты сенсорной настройки - в различных частях периферического поля зрения с шагом 2° . Наряду с известным экспоненциальным падением остроты сенсорной настройки от центра поля зрения к периферии в макромасштабе, обнаружена также ее локальная неоднородность, проявляющаяся в чередовании в периферическом поле зрения областей с высокой и низкой остротой зрительного восприятия с периодом $4^\circ \div 10^\circ$. Статистическая оценка полученных результатов показала достоверное ($p < 0,05$) различие между минимумами и максимумами по количеству точек с низкой остротой сенсорной настройки.

Полученные результаты свидетельствуют о локальной неоднородности сенсорной настройки в периферическом поле зрения и наличии микрзон с различной остротой восприятия. Очевидно, в микрообластях с низкой остротой сенсорной настройки затруднено восприятие зрительной информации, и они могут игнорироваться при осмотре изображения. Напротив, части изображения, проецирующиеся

в микрообласти с высокой остротой сенсорной настройки, могут являться потенциальными зонами притяжения взгляда.

В третьем разделе предлагается алгоритм разноуровневой и контекстной обработки зрительной информации в пределах отдельных фрагментов, основанный на специфической структуре пространственно-неоднородного сенсора и каждого его узла, а также на данных психофизических экспериментов.

В основе алгоритма контекстного описания фрагментов изображения лежит формирование признаков векторов в соответствии с топологией пространственно-неоднородного сенсора, являющегося модификацией разработанного ранее "окна внимания" [Rybak I.A. et al, 1998]. Пространственно-неоднородный сенсор представляет собой структуру, содержащую один базовый (центральный) узел и 48 периферических узлов, расположенных на пересечении 3-х концентрических окружностей (разные уровни разрешения), и 16-ти лучей, исходящих из центра (рис. 1а). Для каждой точки фиксации сенсора выделяются ориентированные перепады яркости в центральной области обрабатываемого фрагмента и в контекстных областях периферических узлов сенсора, причем выделение этих перепадов осуществляется на разных уровнях разрешения, соответствующих расстоянию от центра.

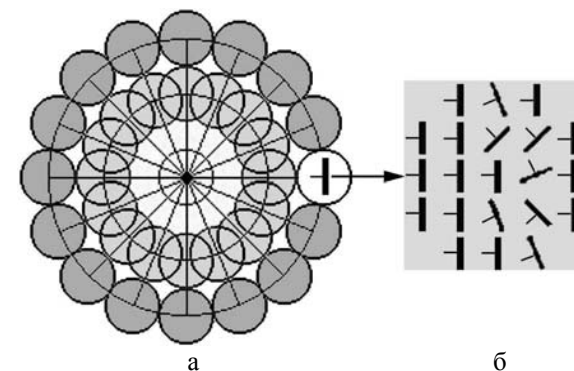


Рис.1. (а) Структура пространственно-неоднородного сенсора. Ориентация перепада яркости, доминирующая в контекстной области одного из узлов сенсора, показана в белом круге. (б) Распределение ориентированных элементов, выделенных в контекстной области данного узла.

Фрагмент изображения в каждой точке фиксации сенсора описывается с помощью признакового вектора $\vec{F}(\alpha, \rho_\alpha)$, компонентами которого являются ориентация α ($\alpha=0,1,\dots,15$), имеющая наибольшее (N_α) представительство в контекстной области (рис. 1б), и доля ρ_α элементов, имеющих эту ориентацию в пределах контекстной области каждого периферического узла. Доля определяется как $\rho_\alpha = \frac{N_\alpha}{S_k}$, где S_k - площадь контекстной области

k -го периферического узла, размер которой пропорционален уровню разрешения и увеличивается в соответствии с расстоянием от центра.

Такой способ описания фрагментов изображения учитывает признаки, выделенные в контекстном окружении каждого периферического узла, что обеспечивает специфичность представления данного фрагмента, сжатие исходной информации, в среднем, в 25 раз и инвариантность к позиционированию в диапазоне, зависящему от размера контекстной области на выбранном уровне разрешения. Разработанная структура пространственно-неоднородного сенсора с контекстным описанием признаков в каждом периферическом узле использовалась при обработке всех тестовых изображений.

Третья глава посвящена описанию методов и алгоритмов инвариантного детектирования и идентификации информативных областей на примере изображений лиц и дорожных знаков. Тестирование разработанных процедур проводилось с использованием изображений лиц из базы ORL¹ (n=400) и базы изображений стандартных дорожных знаков Великобритании (n=100), сканированных из книги правил дорожного движения High Way Code. Глава состоит из двух разделов.

В первом разделе описан каскадный метод идентификации НИО, базирующийся на экспериментальных исследованиях траекторий осмотра изображений лиц, в которых было показано, что НИО являются области глаз, носа и рта [Ярбус А.Л., 1965]. Метод основан на предположении [Головань А.В. и др., 2001], что НИО имеют специфические комбинации первичных признаков, привлекающие зрительное внимание. В качестве первичных признаков были выбраны

¹ База изображений лиц ORL создана в Кембридже (Великобритания) и широко применяется при тестировании различных алгоритмов анализа и распознавания изображений (<http://www.cam-orl.co.uk/facedatabase.html>)

разно-ориентированные контрастные локальные перепады, выделенные на разных уровнях разрешения. Метод включает в себя следующие пять процедур, выполняемых последовательно:

- (1). Выделение ориентированных локальных перепадов яркости, контраст которых превышает порог, равный половине средней величины контраста по всему изображению для каждого уровня разрешения.
- (2). Построение двух частных признаковых карт.

Для получения первой из них по всему изображению оценивается распределение контрастов локальных перепадов, значения которых превышают порог, равный удвоенной величине среднего значения контраста. Для получения второй карты в каждом локальном фрагменте изображения размером 5 x 5 с центром в точке (i,j) оценивается количество перепадов, ориентация которых в некотором диапазоне отличается от ориентации центрального перепада в данном фрагменте. По результатам вычислительных экспериментов был определен оптимальный диапазон изменения относительной ориентации периферического перепада яркости (от 90° до 270°), при котором коэффициент избирательности частной угловой карты был максимален (в среднем 0.55). Коэффициент избирательности определялся как отношение количества "точек интереса", расположенных в пределах НИО, к общему количеству выделенных точек.

- (3). Построение итоговой карты распределения "точек интереса" на изображении.

В процессе получения итоговой карты, для каждого уровня разрешения две частные признаковые карты совмещались "точка в точку". Затем полученные промежуточные карты для разных уровней объединялись по области 3 x 3 для учета возможных смещений первичных признаков, выделенных на разных уровнях разрешения. Описанные процедуры получения итоговой карты обеспечивали относительно высокий коэффициент избирательности (в среднем, 85% "точек интереса" было расположено в пределах НИО) для каждого из трех ракурсов изображения (т.е. детектировали НИО инвариантно ракурсу). Примеры итоговых карт распределения "точек интереса" для изображения одной и той же персоны в трех ракурсах иллюстрирует ряд 1 на рис. 2.

- (4). Группирование "точек интереса" и прореживание по плотности заполнения "областей интереса".

Для группирования строились распределения проекций "точек интереса" по координатным осям. Координаты максимумов по осям X

и Y давали грубую оценку одноименных координат центров "областей интереса", а координаты двух ближайших к каждому максимуму минимумов – грубые оценки одноименных координат границ области. Затем в пределах полученных границ уточнялись координаты центра и размеры каждой из "областей интереса" с помощью метода "ближайшего соседа" [Павлидис Т., 1986]. Прореживание по плотности заполнения "точками интереса" позволяло исключить из рассмотрения "области интереса" с малым их количеством, что увеличивало коэффициент избирательности до 0,9.

(5). Определение типа "областей интереса".

Процедура основана на сравнении гистограмм распределения ориентированных перепадов яркости, выделенных в пределах каждой "области интереса", с усредненными гистограммами каждой НИО, предварительно полученными при участии оператора. Это обеспечивало разделение "областей интереса" на НИО (глаза, нос, рот) и на области, не относящиеся к информативным (уши, воротник, элементы прически) и увеличение коэффициента избирательности до 0,95 (ряд 2 рис.2).

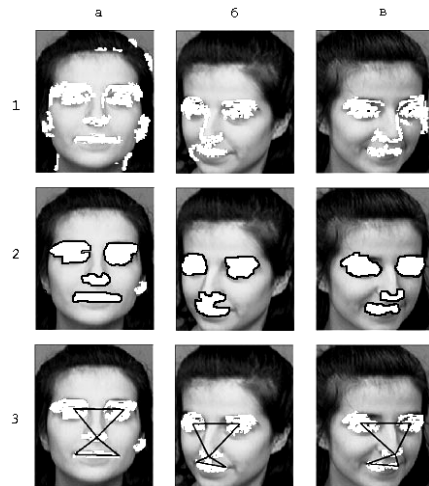


Рис.2. Результаты обработки изображения одной и той же персоны в разных ракурсах (колонки а-в) с помощью последовательных процедур каскадного метода идентификации НИО. Ряд 1 – распределение "точек интереса", ряд 2 – распределение "областей интереса", ряд 3 – схема определения ракурса.

Второй раздел третьей главы посвящен алгоритмам и процедурам идентификации НИО на изображениях дорожных знаков. Очевидно, что для данного типа изображений существенной информативной областью, наряду с внутренней частью знака, является и внешний контур. Определение цвета и границы внешнего контура дорожного знака на изображении проводилось с помощью оценки соотношения цветовых (RGB) компонент (рис.3). Алгоритм определения формы внешнего цветового контура основан на оценке гистограмм ориентированных элементов, выделенных вдоль границы знака. Показано, что "круглые" знаки имеют равновероятное представительство элементов разных ориентаций, тогда как для "прямоугольных" характерно доминирование элементов имеющих горизонтальную и вертикальную ориентацию (в сумме 75% от общего числа выделенных элементов), а для "треугольных" сумма элементов, имеющих характеристическую ориентацию, составила более 85%. Для стандартной базы изображений знаков гистограммный метод обеспечивает безошибочную классификацию в каждой цветовой группе на подгруппы по форме внешнего контура.

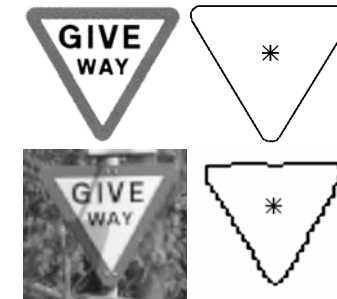


Рис.3. Примеры определения центра и границы внешнего контура (слева) знака для базовых изображений (верхний ряд) и реальных изображений (нижний ряд).

Алгоритм нахождения центра информативной части знака основан на определении геометрического центра цветовых элементов внешнего контура. Он обеспечивает необходимую точность детектирования центра (± 3 пикселя), нормализацию размера знаков и выделение "чистых" знаков (без внешнего фона) из изображений в реальной среде (рис.3, нижний ряд).

Четвертая глава описывает общую структуру разработанной бионической модели распознавания образов и ее модификации для

задач распознавания изображений лиц и дорожных знаков. Глава состоит из трех разделов.

В первом разделе приводится схема разработанной модели, состоящая из трех основных блоков. Первый блок реализует процедуры нижнего уровня обработки информации (получение изображения на разных уровнях разрешения, выделение первичных признаков и детектирование информативных областей) и нормализации изображений. Во втором блоке выполняются процедуры, отнесенные к среднему уровню обработки визуальной информации, в частности – классификация изображений на основе геометрических свойств детектированных информативных областей. Третий блок обеспечивает формирование полного специфического описания изображений на основе детектированных НИО с помощью набора признаковых векторов, заполняемых в топологии пространственно-неоднородного сенсора с контекстным описанием признаков, выделенных в окрестности каждого узла.

Модель предусматривает два режима работы – запоминание и распознавание изображения. Независимо от режима работы модели, каждое изображение подвергается всем процедурам обработки. В первом режиме формируется база эталонов, которая для каждого запоминаемого изображения содержит информацию о взаимном расположении информативных фрагментов, типе каждой НИО и о признаковых векторах, описывающих каждую выделенную область. Распознавание проводится в два этапа. Сначала предъявляемое для распознавания изображение классифицируется по геометрическим свойствам информативных фрагментов, затем вычисляется коэффициент близости признаковых векторов K^b (3), соответствующих одноименным НИО. В качестве прототипа распознаваемого изображения выбирается то изображение из соответствующей подгруппы, сумма коэффициентов близости по всем НИО которого является максимальной. По результатам предварительного тестирования пороговое значение суммы коэффициентов выбиралось равным 25.

$$K^b = \sum_{i=0}^{48} \left[Sg(O_i^b - O_i^{rw}) \cdot \left(1 - \left| \rho_i^b - \rho_i^{rw} \right| \right) \right] \quad (3)$$

где $Sg(x) = \begin{cases} 1, & \text{если } x = 0; \\ 0, & \text{в противном случае} \end{cases}$,

O_i - доминирующая ориентация перепада яркости в контекстной области данного узла сенсора; индекс b присвоен изображениям из

данной подгруппы базы эталонов, индекс rw – изображению, предъявляемому для распознавания; ρ - доля доминирующего перепада яркости в контекстной области.

Программная реализация модели была выполнена в операционной среде Windows 95/98/NT на языке Visual C++, версия 6.0.

Во втором разделе главы приводится описание модификаций концептуальной модели для разработки системы распознавания изображений лиц. Для идентификации областей глаз, носа и рта применялся каскадный метод, подробно описанный в первом разделе третьей главы. В качестве характеристики взаимного расположения НИО были выбраны длины отрезков между центрами выделенных областей, нормированные на расстояние между центрами областей глаз, и углы между отрезками, соединяющими центры НИО (ряд 3 рис.2). Это позволяло с высокой вероятностью ($p = 0,96$) определять ракурс лица на изображении. После идентификации НИО и классификации изображений по ракурсу пространственно-неоднородный сенсор позиционировался в центре каждой из 3-4 выделенных информативных областей. В режиме запоминания изображения классифицировались на три группы – фронтальный ракурс, левый и правый полупрофиль. Тестирование системы проводилось с использованием части изображений базы ORL для десяти персон, каждая из которых представлена в базе 10-ю изображениями в различных ракурсах ($n=100$). При этом 5 изображений каждой персоны в различных ракурсах выбирались для запоминания (эталонные изображения), а оставшиеся 5 использовались для распознавания. Распознавание изображений проводилось с помощью сравнения признаковых векторов одноименных НИО по формуле (3) с учетом ракурса лица на изображении. В проведенных экспериментах **точность распознавания составила около 90%**.

В третьем разделе четвертой главы описаны модификации общей модели для распознавания изображений дорожных знаков с учетом относительной простоты данного типа изображений по сравнению с лицами и принятых стандартов при обработке дорожных знаков. При формировании базы эталонов изображения стандартных знаков Великобритании разбивались на четыре подгруппы, в зависимости от цвета и формы внешнего контура, определяемых с помощью алгоритмов, описанных во втором разделе третьей главы. Формирование специфического описания знака проводилось при однократном позиционировании пространственно-неоднородного сенсора в центре информативной части знака (алгоритм нахождения центра

детально описан во втором разделе третьей главы). Разработанная модель тестировалась при распознавании дорожных знаков, полученных с помощью цифровой фотокамеры при различных погодных условиях (дождь, солнечный день и т.д.) и с различных рядов движения. Выделение знаков из фона проводилось с помощью цветовой сегментации с использованием стандартной модели SIECAM97 [СIE TC1-34, 1998.]. Распознавание дорожных знаков проводилось по отдельным подгруппам, в зависимости от цвета и формы внешнего контура, с помощью сопоставления признаков векторов, описывающих текущее изображение, с хранящимися в базе векторами-эталоны стандартных дорожных знаков. Результаты компьютерных экспериментов показали, что модель обеспечивает быстрое распознавание (время распознавания не превышает 2 сек. на Pentium 166) дорожных знаков, изображения которых получены в реальных условиях, инвариантно к погодным условиям, ракурсу и расстоянию до знака с **уровнем распознавания около 96%**.

В заключении обобщаются полученные результаты, проводится сравнение с известными алгоритмами и методами и рассматриваются перспективы применения разработанных алгоритмов, методов и программ. Описанные алгоритмы и методы дают возможность решать некоторые проблемы специфичного и инвариантного представления фрагментов изображений. Представленная в данной работе модель распознавания различных типов изображений обеспечивает высокий уровень распознавания инвариантно к изменению масштаба, ракурса и к зашумлению. Для практического применения разработанные алгоритмы и модели могут быть оптимизированы. В частности, повышение эффективности распознавания без потери основного свойства – инвариантности к трансформациям изображений – может быть достигнуто за счет увеличения специфичности описания локальных признаков в информативных фрагментах, а также с помощью классификации изображений не только по геометрическим свойствам НИО, но и по их внутреннему содержанию.

ВЫВОДЫ

1. Разработан и исследован алгоритм признакового описания, состоящий в представлении фрагмента изображения в виде вектора локальных признаков, заполняемого с помощью пространственно-неоднородного сенсора с учетом свойств контекстной области каждого его узла. Алгоритм обеспечивает одновременно специфичность и инвариантность отображения фрагментов

изображений в некотором диапазоне их преобразований (изменение масштаба – уменьшение и увеличение в 2 раза от исходного размера; поворот - в пределах 22°; ракурс - от 5° до 30°; уровень шума - до 50% при эмуляции функцией Гаусса)

2. Разработан и исследован каскадный метод идентификации НИО, включающий комплекс последовательно реализуемых алгоритмов и процедур, основанных на выделении специфических комбинаций первичных признаков, характерных для этих областей. Метод обеспечивает идентификацию НИО с точностью 95%.
3. Разработан и исследован алгоритм классификации изображений данного типа, базирующийся на геометрических свойствах НИО, таких как форма и пространственное расположение, и обеспечивающий классификацию изображений на подгруппы с высокой точностью (для изображений лиц – 96%, для изображений знаков – 98%).
4. Разработана модель распознавания лиц (распознавание на уровне 90%), основанная на идентификации НИО и их контекстном описании с помощью пространственно-неоднородного сенсора, а также на классификации базы изображений в соответствии с ракурсом.
5. Разработана модель распознавания дорожных знаков (распознавание на уровне 96%), основанная на одной фиксации пространственно-неоднородного сенсора в центре информативной части знака и классификации по цвету и форме внешнего контура

РЕЗУЛЬТАТЫ ДИССЕРТАЦИИ ОПУБЛИКОВАНЫ В СЛЕДУЮЩИХ ОСНОВНЫХ РАБОТАХ:

1. Шапошников Д.Г., Подладчикова Л.Н., Гао К. Классификация изображений на основе свойств информативных областей. // 6-я Межд. конф. Распознавание образов и анализ изображений: новые информационные технологии: Тез. док. – Великий Новгород, 2002. – Т.2. – С.614-618.
2. Шапошников Д.Г., Головань А.В., Подладчикова Л.Н., Шевцова Н.А., Гао К., Гусакова В.И., Гизатдинова Ю.Ф. Применение поведенческой модели зрения для инвариантного распознавания лиц и дорожных знаков. // Нейрокомпьютеры: разработка и применение. – 2002. – 7-8. – С.21-33.
3. Shaposhnikov D.G., Podladchikova L.N., Golovan A.V., Shevtsova N.A., Hong K., Gao X. Road Sign Recognition by Single Positioning of

- Space-Variant Sensor Window. // Proc. of 15th Int. Conf. on Vision Interface. – 2002. – P.213-217.
4. Головань А.В., Шевцова Н.А., Подладчикова Л.Н., Маркин С.Н., Шапошников Д.Г. Детектирование информативных областей лиц с помощью локальных признаков. // Нейрокомпьютеры: разработка и применение. – 2001. – 1. – С.50-57.
 5. Shaposhnikov D. G., Gizatdinova Yu. F., and Podladchikova L. N. The Peculiarities of the Visual Perception in the Peripheral Field of Vision. // Pattern Recognition and Image Analysis. – 2001. – V.11. – 2. – P.376–378.
 6. Шапошников Д.Г. Алгоритмы выделения информативных областей при обработке изображений лиц. // Труды аспирантов и соискателей Ростовского государственного университета. – Ростов-на-Дону: Изд. РГУ, 2001. – Т.VII. – С.69-71.
 7. Podladchikova L.N., Guskova V.I., Shaposhnikov D.G., Faure A., Golovan A.V. and Shevtsova N.A. MARR: Active vision model. // Intelligent Robots and Computer Vision XVI: Algorithms, Techniques, Active Vision, and Materials Handling: Proc. SPIE. – 1997. – V.3208. – P.418-425.

Выражаю признательность за интересную тему исследований и помощь в работе моему научному руководителю профессору Б.М.Владимирскому и куратору работы зав. лабораторией нейроинформатики НИИ НК РГУ Л.Н. Подладчиковой.