

На правах рукописи



Степанов Дмитрий Юрьевич

**МЕТОДЫ И АЛГОРИТМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ
С ИСПОЛЬЗОВАНИЕМ ДРЕВОВИДНЫХ ПРЕДСТАВЛЕНИЙ
МНОГОКАНАЛЬНЫХ ИЗОБРАЖЕНИЙ**

Специальность 05.13.01 - «Системный анализ, управление и обработка
информации»

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Москва – 2013

Работа выполнена на кафедре «Прикладная синергетика» Федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Московский государственный технический университет радиотехники, электроники и автоматики».

Научный руководитель: кандидат технических наук, старший научный сотрудник, ведущий научный сотрудник Вычислительного центра им. А.А. Дородницына Российской академии наук
Ланге Михаил Михайлович

Официальные оппоненты: **Местецкий Леонид Моисеевич**
доктор технических наук, профессор
Московского государственного университета им. М.В. Ломоносова

Гришин Владимир Александрович
кандидат технических наук,
доцент, старший научный сотрудник
Института космических исследований
Российской академии наук

Ведущая организация: Федеральное государственное бюджетное учреждение науки Институт прикладной математики им. М.В. Келдыша Российской академии наук

Защита состоится «31» октября 2013 г. в 15:00 на заседании диссертационного совета Д212.131.03 при Московском государственном техническом университете радиотехники, электроники и автоматики (МГТУ МИРЭА) по адресу: 119454, г.Москва, пр-т Вернадского, д.78, ауд.Г-412.

С диссертацией можно ознакомиться в библиотеке МГТУ МИРЭА.

Отзывы на автореферат в двух экземплярах, заверенные печатью, просим отправлять по адресу: 119454, г.Москва, пр-т Вернадского, д.78.

Автореферат разослан «13» сентября 2013 г.

Ученый секретарь диссертационного совета Д212.131.03, д.т.н., проф.



О.А. Тягунов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Исследование и разработка методов и алгоритмов автоматического анализа и понимания изображений продиктованы активным развитием информационных технологий, используемых в системах искусственного интеллекта. Развитие таких технологий стимулируется растущим быстродействием вычислительной техники, способной быстро и с высокой точностью обрабатывать большие объемы данных. Одно из ведущих направлений в области анализа и понимания изображений составляет распознавание образов, которое тесно связано с проблематикой технического зрения. Значительный вклад в развитие методов и алгоритмов, используемых для решения задач классификации и анализа изображений, внесли как отечественные ученые Ю.И. Журавлев, К.В. Рудаков, В.Н. Вапник, А.Я. Червоненкис, Н.Г. Загоруйко, Ю.П. Пытьев, А.И. Чуличков, К.В. Воронцов, Л.М. Местецкий, В.В. Моттль, В.В. Рязанов, так и зарубежные специалисты R. Duda, R. Gonzalez, J. Serra, R. Haralick, L. Shapiro, T. Huang, A. Rosenfeld, L. Kuncheva.

В диссертационной работе исследуется задача распознавания объектов, заданных многоканальными изображениями. В этом случае объекты задаются наборами образов, которые выделяются на изображениях, получаемых по различным информационным каналам. Особенность рассматриваемой задачи характеризуется большим числом классов. Примером такого источника являются изображения биометрических данных (лиц, отпечатков пальцев, радужных оболочек глаз, подписей, рисунков ладоней и т.п.), которые используются в системах идентификации личности. Число классов в таких системах определяется числом персон, информация о которых хранится в базе данных, и может составлять величины порядка тысяч и более. Подобный источник образуют также цветные изображения, которые являются трехканальными изображениями в стандартных моделях RGB, HSI и др.

Многоканальные изображения обладают существенно большим количеством информации о распознаваемых объектах и позволяют повысить достоверность классификации за счет уменьшения доли ошибочных решений. Однако многоканальность источника в сочетании с большим числом классов порождает большие объемы данных, которые в реальных системах должны обрабатываться за достаточно малое время. Поэтому указанная специфика задачи распознавания, наряду с высокими требованиями к достоверности решений, выдвигает особые требования к быстродействию классификатора, причем показатели достоверности и быстродействия, как правило, связаны обратным соотношением.

Для достижения компромисса между достоверностью и быстродействием задачу распознавания целесообразно решать в терминах соотношения характеристик качества и вычислительной сложности, требуя минимизации вероятности ошибок при заданном ограничении на объем вычислений. В такой постановке актуальной задачей является разработка и исследование универсальной модели классификатора с метрическим решающим правилом в пространстве структурных представлений образов с многоуровневым разрешением. Универсальность предполагает высокую эффективность такого классификатора для различных источников многоканальных изображений, а структура представлений образов должна обеспечить высокое быстродействие за счет иерархического поиска решений в базе эталонов с многоуровневым разрешением.

Один из первых подходов к построению структурных описаний геометрических форм на основе их иерархической декомпозиции предложен Н. Row и G. Medioni (1993). Метод рекурсивной декомпозиции использован S. Berretti и A. Del Vimbo (2004) для представления двумерных форм древовидно-структурированными наборами спектральных признаков. Для распознавания геометрических форм А. Torsello и E. Hancock (2004) исследовали процедуры установления соответствия иерархически-структурированных признаков. Подходы и методы, рассмотренные этими авторами, ориентированы на построение эффективного пространства описаний (представлений) образов, выделенных на бинарных изображениях. К подобным структурным описаниям геометрических форм относятся также скелетные представления, предложенные Л.М. Местецким (2000), и древовидные представления наборами эллиптических примитивов, разработанные М.М. Ланге и С.Н. Ганебных (2004). Древовидные представления эллиптическими примитивами обобщены М.М. Ланге и С.Н. Ганебных (2007-2009) для широкого класса образов, заданных одноканальными полутоновыми изображениями. На множестве представлений введена мера различия полутоновых образов, и показана эффективность полученного пространства представлений для распознавания жестов руки и подписей.

Развитием структурных методов распознавания является обобщение метода древовидного представления эллиптическими примитивами для источников многоканальных изображений и построение эффективного и быстрого классификатора в пространстве обобщенных древовидных представлений. Актуальность этого подхода обусловлена его универсальностью и широкими практическими приложениями.

Цель работы и задачи исследования. Цель диссертационного исследования заключается в разработке универсального метрического классификатора в пространстве древовидных представлений образов с многоуровневым разрешением для источников многоканальных

изображений. Область исследования включает разработку методов и алгоритмов решения задачи распознавания образов и проведение вычислительных экспериментов по распознаванию объектов, заданных изображениями. В соответствии с поставленной целью решаются следующие задачи:

1. Разработка модели метрического классификатора объектов по многоканальным изображениям. Построение многослойных древовидных представлений образов и введение меры различия объектов на множестве представлений. Формализация критерия классификации.

2. Построение процедуры обучения классификатора на основе отбора эталонов и оптимизации параметров. Получение сравнительных оценок вычислительной сложности иерархического и переборного алгоритмов поиска решений.

3. Подготовка базы изображений, разработка комплекса программ и проведение экспериментов по идентификации личности с использованием многоканальных данных, заданных полутоновыми изображениями подписей и цветными изображениями лиц.

Основные положения, выносимые на защиту.

1. Модель метрического классификатора по критериям взвешенного голосования и ближайшего эталона в пространстве многослойных древовидных представлений образов, заданных многоканальными изображениями. Модель включает способ представления объектов, меру различия на множестве представлений и критерии классификации.

2. Процедура обучения классификатора на основе построения древовидно-структурированного покрытия обучающего множества сферами в пространстве представлений с заданной мерой. Функция ошибок и численная оптимизация параметров покрытия и меры методом скользящего контроля.

3. Способ комплексирования решений по обобщенной мере различия объектов на множестве многослойных древовидных представлений.

4. Программный комплекс для реализации процедур обучения и тестирования классификатора. Экспериментальные оценки качества идентификации личности по полутоновым изображениям подписей, цветным изображениям лиц и по ансамблю изображений подписей и лиц.

Методы исследования. В выполненных разработках и исследованиях использовались методы теории информации, теории вероятности и теории распознавания образов, элементы математического моделирования, численных методов оптимизации и прикладного программирования.

Научная новизна работы состоит в построении пространства универсальных древовидно-структурированных представлений двумерных объектов, заданных многоканальными изображениями; разработке метрической модели классификатора в указанном пространстве представлений; получении оценок вычислительной сложности, демонстрирующих выигрыш в быстродействии иерархического алгоритма поиска решения по сравнению с алгоритмом полного перебора; предложении процедуры обучения метрического классификатора; разработке способа объединения решений по обобщенной мере различия объектов; получении экспериментальных оценок эффективности разработанного классификатора для источников многоканальных биометрических изображений.

Теоретическая и практическая ценность. Теоретическую ценность работы составляют многослойное представление образов с многоуровневым разрешением, мера различия на множестве многослойных представлений и модель метрического классификатора. Теоретические результаты работы включены в курс лекций по дисциплине «Синергетическая информатика» на кафедре «Прикладная синергетика» в МГТУ МИРЭА и использованы в научных отчетах по проекту Российского фонда фундаментальных исследований №09-01-00573-а, выполненному в Вычислительном центре им. А.А. Дородницына РАН.

Практическую ценность выполненных исследований составляют результаты экспериментов по классификации биометрических изображений. Полученные оценки качества распознавания демонстрируют высокую эффективность разработанной модели классификатора и уменьшение доли и дисперсии ошибок при расширении состава источников (увеличении числа информационных каналов). Разработанный комплекс программ внедрен в системе контроля доступа в помещения организации ОАО «Чувашнефтепродукт».

Достоверность полученных результатов обоснована математическими выкладками, модельными экспериментами и согласованностью полученных оценок качества распознавания с аналогичными характеристиками известного классификатора на основе метода опорных векторов. Теоретическая и практическая ценность выполненных разработок подтверждена актами о внедрении.

Апробация работы. Результаты диссертационной работы обсуждены и получили положительные отзывы на следующих научных конференциях, конкурсах и семинарах:

- 11-й научно-практической конференции «Современные информационные технологии в управлении и образовании», г.Москва, ФГУП НИИ «Восход», 2012 г. (работа отмечена дипломом 2-й степени и премией имени д.т.н. В.С. Корсакова-Богаткова);

- научном семинаре «Морфологический анализ данных» под руководством д.ф.-м.н., проф. Ю.П. Пытьева, г.Москва, Физический ф-т МГУ, 2011 г.;
- 14, 15-й Всероссийской конференции «Математические методы распознавания образов», г.Суздаль, 2009 г., г.Петрозаводск, 2011 г.;
- 10-й Международной конференции «Распознавание образов и анализ изображений: новые информационные технологии», г.Санкт-Петербург, 2010 г.;
- 16-й Международной научно-практической конференции студентов и молодых ученых «Современные техника и технологии», г.Томск, 2010 г.;
- 3-й Всероссийской конференции студентов, аспирантов и молодых ученых «Искусственный интеллект: философия, методология, инновации», г.Москва, 2009 г. (работа отмечена дипломом);
- 58, 59, 60-й научно-технической конференции МИРЭА, г.Москва, МИРЭА, 2009-2011 гг.;
- конкурсе «Лучшая научная работа студентов и молодых ученых МИРЭА 2009», г.Москва, МИРЭА, 2009 г. (работа отмечена почетной грамотой);
- открытом конкурсе инновационных идей «Стремление», г.Москва, МИРЭА, 2009 г. (работа отмечена дипломом 2-й степени);
- Всероссийском смотре-конкурсе научно-технического творчества студентов и аспирантов высших учебных заведений «Эврика-2009», г.Новочеркасск, ЮРГТУ НПИ, 2009 г. (работа отмечена дипломом 3-й степени).

Диссертационная работа поддержана грантами Российского фонда фундаментальных исследований №09-01-16059-моб-з и 10-01-16048-моб-з.

Публикации. По материалам диссертационной работы опубликовано 12 печатных работ, в том числе 4 в журналах из перечня ВАК, получены 2 свидетельства о государственной регистрации программного продукта (№50200900489, 50201001618).

Структура и объем работы. Диссертационная работа состоит из введения, пяти глав, заключения, библиографического списка из 80 наименований и двух приложений. Работа содержит 170 страниц основного текста, 64 рисунка и 22 таблицы.

ОСНОВНЫЕ ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

Пусть изображение определяется целочисленной функцией, заданной на двумерной прямоугольной решетке с конечным числом элементов (пикселей). Под информационным каналом наблюдения

понимается способ отображения реальных предметов в их изображения. Изображение на выходе заданного канала называется одноканальным изображением. Набор одноканальных изображений, полученных по различным каналам наблюдения, образует многоканальное изображение. Источник определяется множеством одноканальных или многоканальных изображений семантически однотипных предметов, которые принадлежат заданному числу классов, снабженных семантическими признаками (именами). Подмножество изображений, принадлежащих одному классу, образует кластер. Полутонные изображения предметов образуют источники одноканальных изображений, цветные изображения содержат три канала и являются примером многоканальных изображений. В общем случае любой набор источников порождает многоканальные изображения.

Образом называется информативный набор пикселей, выделенный на одноканальном изображении, полученном от заданного канала наблюдения. Объект определяется совокупностью образов, выделенных на наборе одноканальных изображений, образующих многоканальное изображение. Для источников одноканальных изображений понятия образа и объекта тождественны. Представление объекта определяется набором описаний образов, составляющих этот объект, в заданной системе признаков. Представление объекта, заданного одноканальным изображением, определяются описанием образа этого объекта на изображении.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертационной работы, сформулирована цель исследования, отмечена научная новизна и практическая ценность результатов, выделены основные положения, выносимые на защиту. Приведены сведения об апробации и использовании результатов и данные о структуре диссертации.

В первой главе выполняется обзор наиболее известных методов распознавания, вводятся основные понятия теории распознавания образов, выявляются типовые требования к методам распознавания, и формулируется перечень задач, исследуемых в диссертационной работе.

В разделе 1.1 рассмотрены особенности байесовских, метрических и линейных методов классификации. Отмечено, что основным преимуществом байесовских методов является простота реализации, а недостатком - необходимость использования обучающего множества большой мощности. Метрические схемы допускают небольшие обучающие выборки, однако, результаты классификации могут быть не устойчивы к случайным выбросам. Линейные методы обеспечивают высокое качество распознавания, но вычислительно трудоемки.

В разделе 1.2 более детально изложена проблематика метрической модели. Рассмотрены вопросы формирования пространства представлений образов и выбора меры. Обсуждены характеристики качества классификации в терминах ложных отказов (False Rejection Rate) и ложных распознаваний (False Acceptance Rate) и соотношение качества и вычислительной сложности. Выбран подход к снижению вычислительной сложности в пространстве представлений с многоуровневым разрешением.

В разделе 1.3 введены основные термины и определения, сформулированы задачи, исследуемые в диссертационной работе.

Во второй главе дано описание способа древовидного представления объектов многоканальных изображений. Введена мера различия объектов на множестве представлений. Построена модель метрического классификатора, и оценена вычислительная сложность иерархического и переборного алгоритмов поиска решений.

В разделе 2.1 изложен способ представления объектов многоканальных изображений. Пусть \mathbf{A} – множество объектов, в котором каждый объект $A \in \mathbf{A}$ задан набором образов, выделенных на N изображениях (от N каналов наблюдения). Предполагается, что представление образа от каждого q -го канала ($q=1, \dots, N$) может быть построено с помощью известной процедуры рекурсивного разбиения образа на непересекающиеся сегменты и аппроксимации сегментов эллиптическими примитивами [М.М. Ланге, С.Н. Ганебных, 2009]. Такое представление имеет вид полного бинарного дерева глубины L , узлы которого содержат описания эллиптических примитивов. Тогда для объекта $A \in \mathbf{A}$, заданного N -канальным изображением, предлагается использовать многослойное представление

$$A^L = \{A_q^L = (a_q^0, \dots, a_q^l, \dots, a_q^L)\}_{q=1}^N, \quad (1)$$

в котором q -й слой A_q^L является полным бинарным деревом глубины L , содержащим уровни a_q^l , $l = 0, \dots, L$. Подпоследовательность $A_q^l = (a_q^0, \dots, a_q^l)$ в (1) образует поддерево глубины l . Уровень a_q^l задан набором эллиптических примитивов

$$a_q^l = \{Q_{qn} = (\mathbf{r}_{qn}, \mathbf{u}_{qn}, \mathbf{v}_{qn}, z_{qn})\}_{n=2^l-1}^{2^{l+1}-2}, \quad (2)$$

где Q_{qn} - примитив q -го слоя с номером n , \mathbf{r}_{qn} - вектор центра, $(\mathbf{u}_{qn}, \mathbf{v}_{qn})$ - векторы ориентации, z_{qn} - значение яркости примитива.

На рис.1 даны примеры однослойного древовидного представления изображения подписи по каналу I и трехслойного древовидного представления изображения лица по каналам модели HSI.

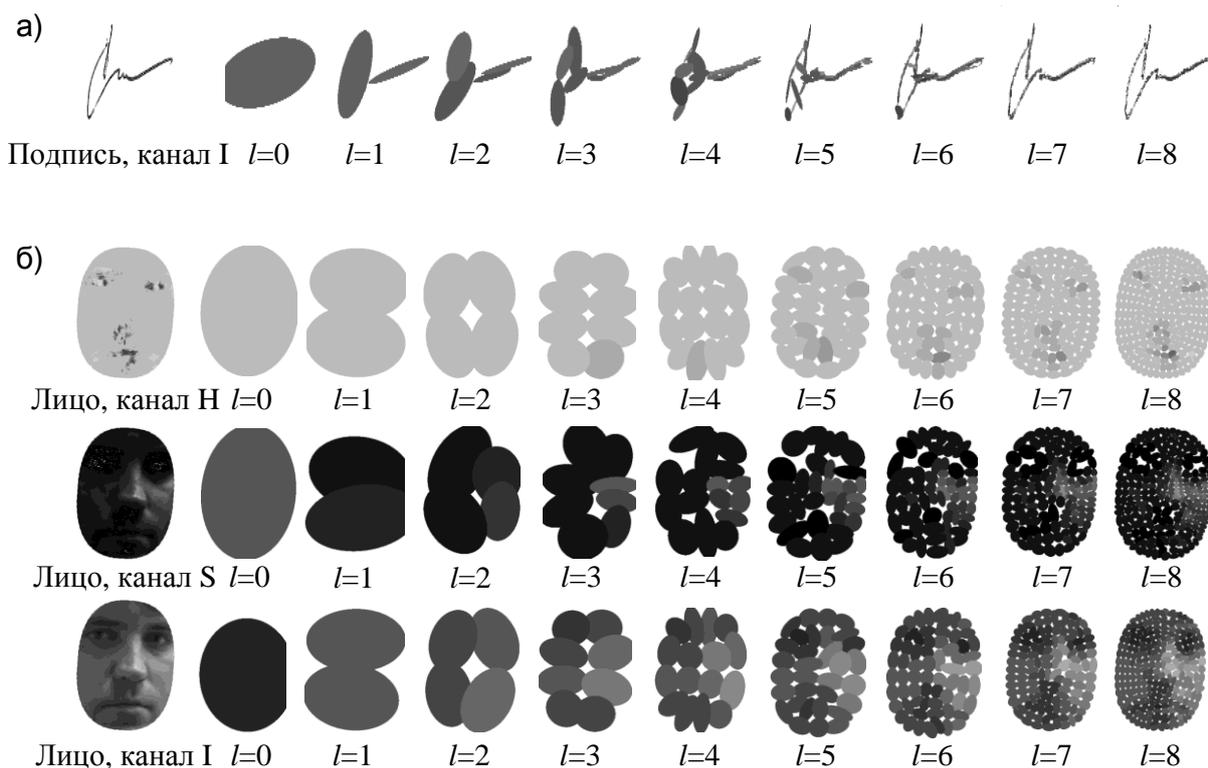


Рис.1. Древовидно-структурированное представление с уровнями разрешения $l=0, \dots, 8$: а) подписи; б) лица

Построение эллиптических примитивов каждого q -го слоя A_q^L в собственных координатах аппроксимируемых сегментов и нормировка параметров всех примитивов относительно корневого примитива с номером $n = 0$ позволяют сформулировать следующее утверждение.

Утверждение 1. Многослойное представление, заданное соотношениями (1) и (2), практически инвариантно к преобразованиям поворота, смещения, изменения масштаба и уровня яркости объекта при достаточно малом размере пикселей и большом числе уровней квантования значений пикселей.

В разделе 2.2 введена мера различия пары объектов (A, \hat{A}) по любому q -му слою их представлений A_q^l и \hat{A}_q^l , построена обобщенная мера по многослойным представлениям $A^l = \{A_q^l\}_{q=1}^N$ и $\hat{A}^l = \{\hat{A}_q^l\}_{q=1}^N$, и разработана модель метрического классификатора по введенной обобщенной мере. Для любой пары соответственных примитивов $(Q_{qn} \in A_q^l, \hat{Q}_{qn} \in \hat{A}_q^l)$ с номером $n > 0$ вводятся неотрицательные функции потерь

$$\rho_1(Q_{qn}, \hat{Q}_{qn}) = \frac{\|\mathbf{r}_{qn} - \hat{\mathbf{r}}_{qn}\|}{\max(\|\mathbf{r}_{qn}\|, \|\hat{\mathbf{r}}_{qn}\|)}, \quad (3a)$$

$$\rho_2(Q_{qn}, \hat{Q}_{qn}) = \frac{1}{2} \left(\frac{\|\mathbf{u}_{qn} - \hat{\mathbf{u}}_{qn}\|}{\max(\|\mathbf{u}_{qn}\|, \|\hat{\mathbf{u}}_{qn}\|)} + \frac{\|\mathbf{v}_{qn} - \hat{\mathbf{v}}_{qn}\|}{\max(\|\mathbf{v}_{qn}\|, \|\hat{\mathbf{v}}_{qn}\|)} \right), \quad (36)$$

$$\rho_3(Q_{qn}, \hat{Q}_{qn}) = \frac{|z_{qn} - \hat{z}_{qn}|}{\max(z_{qn}, \hat{z}_{qn})}, \quad (37)$$

где $\|\mathbf{f}\|$ - евклидова норма вектора \mathbf{f} , а $|f|$ - модуль значения f . С учетом функций (3) для объектов A и \hat{A} определяются три компоненты меры различия l -го порядка по q -м слоям A_q^l и \hat{A}_q^l

$$d_{lk}^{(q)}(A, \hat{A}) = \sum_{n=1}^{2^{l+1}-2} w_n \rho_k(Q_{qn}, \hat{Q}_{qn}), \quad k=1,2,3, \quad (4)$$

и мера различия l -го порядка вида

$$d_l^{(q)}(A, \hat{A}) = \sum_{k=1}^3 \omega_k^{(q)} d_{lk}^{(q)}(A, \hat{A}) \quad (5)$$

на уровнях с номерами $l=1, \dots, L$. Весовые коэффициенты $w_n \geq 0$ в (4) определяются номером n соответственных примитивов Q_{qn} и \hat{Q}_{qn} , а

коэффициенты $\{\omega_k^{(q)} \geq 0: \sum_{k=1}^3 \omega_k^{(q)} = 1\}_{k=1}^3$ в (5) оцениваются на этапе

обучения. С учетом (5) для пары объектов (A, \hat{A}) , представленных в форме (1), введена обобщенная мера различия l -го порядка

$$d_l(A, \hat{A}) = \sum_{q=1}^N \gamma^{(q)} d_l^{(q)}(A, \hat{A}), \quad \gamma^{(q)} \geq 0, \quad (6)$$

где $\{\gamma^{(q)} \geq 0: \sum_{q=1}^N \gamma^{(q)} = 1\}_{q=1}^N$ - нормированные весовые коэффициенты,

оценки которых также строятся на этапе обучения.

Предложена следующая модель классификатора. Пусть множество $\mathbf{A} = \{\mathbf{A}_i\}_{i=0}^c$ содержит объекты, допускающие представление (1) и принадлежащие $c+1$ классам. Каждый класс \mathbf{A}_i с номером $i \neq 0$ включает семантически однородные объекты, а класс \mathbf{A}_0 объединяет все прочие объекты. При известных априорных вероятностях классов

$$P_i = P(\mathbf{A}_i): \sum_{i=0}^c P(\mathbf{A}_i) = 1, \quad \text{вероятности} \quad P_{own} = \sum_{i=1}^c P(\mathbf{A}_i) \quad \text{и}$$

$P_{alien} = P_0 = 1 - P_{own}$ дают распределение «своих» (*own*) и «чужих» (*alien*) объектов на множестве \mathbf{A} .

Для обучения используется обучающее множество

$$\mathbf{B} = \{\mathbf{B}_i = \{B_{ij}\}_{j=1}^{m_i}\}_{i=1}^c \subset \mathbf{A},$$

составленное из семантически однородных кластеров $\mathbf{B}_i \subset \mathbf{A}_i$ мощности $m_i = \|\mathbf{B}_i\|$. На обучающем множестве отбирается множество эталонов

$$\hat{\mathbf{B}} = \{\hat{\mathbf{B}}_i = \{\hat{B}_{ij}\}_{j=1}^{\hat{m}_i}\}_{i=1}^c \subset \mathbf{B}, \quad (7)$$

в котором $\hat{m}_i \leq m_i$ и каждый эталонный объект \hat{B}_{ij} , представленный поддеревом \hat{B}_{ij}^l , является центром сферы

$$S_l(\hat{B}_{ij}, D_l(\hat{B}_{ij})) = \{B \in \mathbf{B} : d_l(B, \hat{B}_{ij}) \leq D_l(\hat{B}_{ij})\} \quad (8)$$

с радиусом $D_l(\hat{B}_{ij})$, вычисляемым по мере l -го порядка вида (6). Используя эталоны (7), заданные представлениями (1), и их сферы (8), взятые при $l=L$, введена мера сходства L -го порядка между любым объектом $A \in \mathbf{A}$ и подмножеством эталонов $\hat{\mathbf{B}}_i$. Эта мера имеет вид

$$\mu_L(A, \hat{\mathbf{B}}_i) = \sum_{j=1}^{\hat{m}_i} q(\hat{B}_{ij}) 2^{\frac{-d_L(A, \hat{B}_{ij})}{D_L(\hat{B}_{ij})}} [d_L(A, \hat{B}_{ij}) \leq D_L(\hat{B}_{ij})], \quad (9)$$

где $\{q(\hat{B}_{ij}) \geq 0 : \sum_{j=1}^{\hat{m}_i} q(\hat{B}_{ij}) = 1\}$ - нормированные весовые коэффициенты эталонов, $[(*)]$ - индикатор условия (*). Мера (9) дает следующий критерий установления номера класса i^* для объекта $A \in \mathbf{A}$

$$i^* = k[\mu_L(A, \hat{\mathbf{B}}_k) > 0], \quad k = \arg \max_{i=1}^c \mu_L(A, \hat{\mathbf{B}}_i), \quad (10)$$

который при нулевом значении индикатора порождает отказ ($i^* = 0$), соответствующий классу \mathbf{A}_0 . В общем случае критерий (10) определяет правило взвешенного голосования (Weighted Vote) эталонов, дополненное функцией отказа. При замене суммы (9) наибольшим членом, критерий (10) эквивалентен правилу ближайшего эталона (Nearest Neighbor).

В разделе 2.3 оценена вычислительная сложность алгоритмов принятия решений (10) на основе стратегии иерархического поиска и полного перебора. Вычислительная сложность решающих алгоритмов измеряется числом пар вершин-примитивов, обрабатываемых при сравнении предъявляемого и эталонных объектов в форме (1). Иерархический поиск (Guided Search) базируется на сокращении числа анализируемых классов в последовательных уровнях базы эталонов, представленных сферами (8) порядка $l=1, 2, \dots, L$. Число классов, анализируемых на l -м уровне (по представлениям глубины l), задается

функцией

$$c_l = \lfloor c 2^{-\beta(l-1)} \rfloor, \quad l = 1, 2, \dots, L,$$

где коэффициент $\beta \geq 0$. Из c_l классов отбираются c_{l+1} (для анализа на $(l+1)$ -м уровне) с наибольшими значениями меры сходства l -го порядка вида (9), которая вычисляется по представлению A^l предъявляемого объекта и сферам $S_l(\hat{B}_{ij}, D_l(\hat{B}_{ij}))$ эталонных объектов при единичном значении индикатора. Алгоритм полного перебора (Exhaustive Search) анализирует все c классов эталонов по представлениям глубины L . Следующее утверждение содержит сравнительные оценки вычислительной сложности переборного и иерархического алгоритмов.

Утверждение 2. При $\beta = \frac{1}{L} \log_2 c \geq 1$ вычислительная сложность переборного и иерархического алгоритмов поиска решения удовлетворяет оценкам

$$C_{ES} = 2N(2^L - 1) \sum_{i=1}^c \hat{m}_i = O(c^2),$$

$$C_{GS} \leq 2NcL \max_{i=1}^c \hat{m}_i = O(c \log_2 c).$$

Из утверждения 2 следует, что при большом числе классов c алгоритм Guided Search дает выигрыш $C = \frac{C_{ES}}{C_{GS}}$ в быстродействии по сравнению с алгоритмом Exhaustive Search порядка $\frac{c}{\log_2 c}$.

Раздел 2.4 посвящен обсуждению результатов второй главы. Основной вывод - предложенная модель представления многоканальных данных позволяет уменьшить объем вычислительных операций за счет многоуровневой структуры, удобной для иерархического поиска решений.

В третьей главе описана процедура отбора эталонов с помощью построения древовидно-структурированных покрытий (Tree-Structured Covers) кластеров обучающего множества сферами, введена функция ошибки скользящего контроля, предложен способ оптимизации TSC-покрытий на основе минимизации функции ошибки скользящего контроля и способ оценивания параметров меры на множестве древовидных представлений объектов.

В разделе 3.1 предложена процедура построения древовидно-структурированных покрытий, состоящая в следующем. На множестве представлений деревьями глубины L для каждого кластера $\mathbf{B}_i \subset \mathbf{B}$

мощности $m_i = \|\mathbf{V}_i\|$, $i = 1, \dots, c$, выполняется дихотомическое разбиение на непересекающиеся сегменты \mathbf{V}_{ij} размера $m_{ij} = \|\mathbf{V}_{ij}\|$

$$\mathbf{V}_i = \bigcup_{j=1}^{\hat{m}_i} \mathbf{V}_{ij}, \quad m_i = \sum_{j=1}^{\hat{m}_i} m_{ij},$$

и выбор для каждого сегмента \mathbf{V}_{ij} покрывающей сферы вида (8) с центром

$$\hat{B}_{ij} = \arg \min_{\hat{B} \in \mathbf{V}_{ij}} \max_{B \in \mathbf{V}_{ij}} d_L(B, \hat{B}) \quad (11)$$

и радиусом

$$D_L(\hat{B}_{ij}) = (1 - \alpha_i) D_L^{(\min)}(\hat{B}_{ij}) + \alpha_i D_L^{(\max)}(\hat{B}_{ij}). \quad (12)$$

Здесь α_i - параметр, выбираемый на отрезке $[0, 1]$; $D_L^{(\min)}(\hat{B}_{ij})$ - расстояние по заданной мере между центром $\hat{B}_{ij} \in \mathbf{V}_{ij}$ и наиболее удаленным объектом $B \in \mathbf{V}_{ij}$; $D_L^{(\max)}(\hat{B}_{ij})$ - наибольшее значение из пары величин: значения $D_L^{(\min)}(\hat{B}_{ij})$ и величины расстояния между центром $\hat{B}_{ij} \in \mathbf{V}_{ij}$ и наименее удаленным объектом $B \in \mathbf{V} \setminus \mathbf{V}_i$. На каждом шаге дихотомии выполняется разбиение сегмента \mathbf{V}_{ij} , $\mathbf{V}_{ij} \rightarrow (\mathbf{V}_{ij'}, \mathbf{V}_{ij''})$, имеющего наибольшее рассеяние $\max_{B \in \mathbf{V}_{ij}} d_L(B, \hat{B}_{ij})$ среди сегментов текущего разбиения кластера \mathbf{V}_i . Сегменты $\mathbf{V}_{ij'}$ и $\mathbf{V}_{ij''}$ включают объекты, которые наиболее близки по заданной мере различия L -го порядка к наиболее удаленным друг от друга опорным объектам $B_{ij'} \in \mathbf{V}_{ij'}$ и $B_{ij''} \in \mathbf{V}_{ij''}$. Пример TSC-покрытия кластера $\mathbf{V}_i \subset \mathbf{V}$ сферами дан на рис.2.

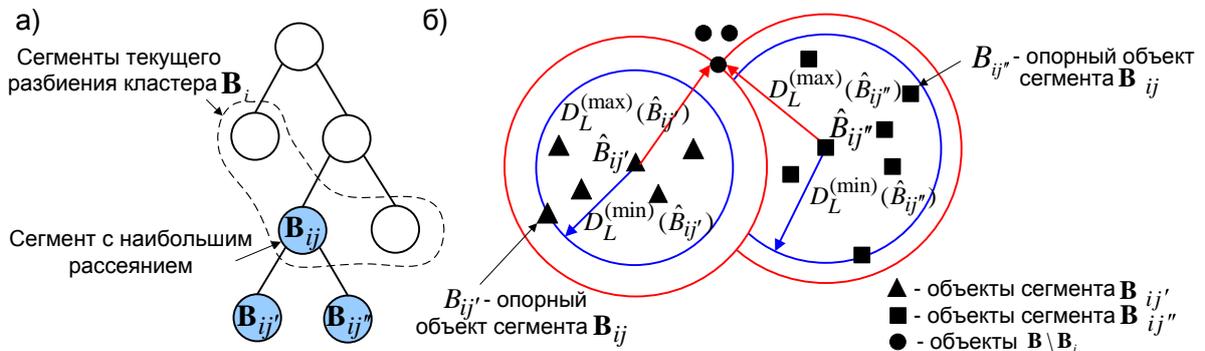


Рис.2. Покрытие кластера \mathbf{V}_i сферами: а) бинарное дерево сегментов; б) сферы покрытия сегментов

Утверждение 3. Множество сфер с центрами вида (11) и радиусами вида (12) образует древовидно-структурированное покрытие мощности \hat{m}_i кластера \mathbf{B}_i . Объединение покрытий всех кластеров порождает TSC-классификатор по критерию (10).

Покрытие кластера \mathbf{B}_i образует пару множеств $(\hat{\mathbf{B}}_i, \mathbf{D}_L(\hat{\mathbf{B}}_i))$, заданных соответственно центрами и радиусами покрывающих сфер i -го кластера. Объединение покрытий c кластеров порождает пару множеств $\hat{\mathbf{B}} = \bigcup_{i=1}^c \hat{\mathbf{B}}_i$ и $\mathbf{D}_L(\hat{\mathbf{B}}) = \bigcup_{i=1}^c \mathbf{D}_L(\hat{\mathbf{B}}_i)$. Проекция $(\hat{\mathbf{B}}, \mathbf{D}_l(\hat{\mathbf{B}}))$, $l=1, \dots, L-1$, этой пары совместно с $(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ образуют L -уровневую базу эталонов.

В разделе 3.2 получена оценка вероятности ошибки обучения классификатора, и предложен способ оптимизации параметров покрытий кластеров на основе минимизации функции ошибки скользящего контроля.

Утверждение 4. Вероятность ошибки обучения TSC-классификатора с параметрами $(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ оценивается средним значением

$$\varepsilon_{\mathbf{B}}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) = \frac{1}{c} \sum_{i=1}^c \varepsilon_{\mathbf{B}}^{cv}(\hat{m}_i, \alpha_i), \quad (13)$$

где

$$\begin{aligned} \varepsilon_{\mathbf{B}}^{cv}(\hat{m}_i, \alpha_i) = & \left(\frac{(c-1)}{c} \varepsilon_{\mathbf{B} \setminus \mathbf{B}_i}^{cv}(\hat{m}_i, \alpha_i) + \varepsilon_{\mathbf{B}_i}^{cv}(\hat{m}_i, \alpha_i) \right) P_{own} + \\ & + \varepsilon_{\mathbf{B} \setminus \mathbf{B}_i}^{cv}(\hat{m}_i, \alpha_i) P_{alien} \end{aligned} \quad (14)$$

функция ошибки скользящего контроля, $\varepsilon_{\mathbf{B}_i}^{cv}(\hat{m}_i, \alpha_i)$ и $\varepsilon_{\mathbf{B} \setminus \mathbf{B}_i}^{cv}(\hat{m}_i, \alpha_i)$ - доли ложных отказов и ложных распознаваний при предъявлении покрытию с параметрами (\hat{m}_i, α_i) «своих» (\mathbf{B}_i) и «чужих» ($\mathbf{B} \setminus \mathbf{B}_i$) объектов.

Минимум оценки (13) по параметрам $\{\hat{m}_i, \alpha_i\}_{i=1}^c$ достигается в результате независимой минимизации функций (14) по переменным \hat{m}_i и α_i . Оптимизация сводится к нахождению параметров

$$(\hat{m}_i, \alpha_i) = \arg \min_{1 \leq \hat{m}_i' \leq m; 0 \leq \alpha_i' \leq 1} \varepsilon_{\mathbf{B}}^{cv}(\hat{m}_i', \alpha_i'), \quad (15)$$

где m определяет наибольшее заданное число эталонов.

В разделе 3.3 получены оценки параметров меры в соотношениях (5) и (6) при $l=L$. В качестве коэффициентов $\omega_1^{(q)}$, $\omega_2^{(q)}$, $\omega_3^{(q)}$ в мере (5) выбраны оценки

$$\omega_k^{(q)} = \frac{\log_2 \varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))}{\sum_{k=1}^3 \log_2 \varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))}, \quad k=1,2,3,$$

где $\varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))$ - доля ошибок вида (13), полученная на покрытиях с параметрами (15) по мере $d_{Lk}^{(q)}(B, \hat{B})$, $B \in \mathbf{B}$, $\hat{B} \in \mathbf{B}$. Аналогичный способ использован для оценивания коэффициентов $\gamma^{(q)}$, $q=1, \dots, N$, в обобщенной мере (6). В качестве оценок выбраны величины

$$\gamma^{(q)} = \frac{\log_2 \varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))}{\sum_{q=1}^N \log_2 \varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))},$$

где $\varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ - доля ошибок вида (13), вычисляемая на покрытиях с параметрами (15) по мере $d_L^{(q)}(B, \hat{B})$, $B \in \mathbf{B}$, $\hat{B} \in \mathbf{B}$.

Предложенный способ вычисления обобщенной меры различия положен в основу стратегии объединения классификаторов по ансамблю источников, которая названа GDM-стратегией (General Dissimilarity Measure). Наряду с предложенной GDM-стратегией, исследована известная WMV-стратегия (Weighted Majority Vote) объединения решений по отдельным каналам [L. Kuncheva, 2004].

Раздел 3.4 посвящен обсуждению результатов третьей главы. Основной вывод состоит в том, что предложенные способы отбора эталонов и оценивания параметров меры применимы к различным источникам и могут использоваться для объединения классификаторов.

Четвертая глава содержит описание баз изображений подписей и лиц и реализованного комплекса программ для проведения экспериментов с разработанным TSC-классификатором и известным SVM-классификатором (Support Vector Machine) в версии библиотеки OpenCV.

В разделе 4.1 рассмотрены основные этапы подготовки базы подписей. Выполнен анализ доступных баз изображений подписей. Используя открытую базу подписей [Signature Verification Competition, 2004], сформирована база полутоновых I-изображений подписей от $c = 25$ персон по $m_i = 40$ реализаций каждой персоны. Для улучшения качества применена процедура пороговой фильтрации полученных изображений.

В разделе 4.2 приведены основные этапы формирования базы изображений лиц. Проведена съемка, и сформирована база цветных HSI-изображений от $c = 25$ персон по $m_i = 40$ реализаций от каждой персоны с вариацией ракурса съемки. Предложена процедура выделения

информативной области лица на изображении. К найденным областям применены процедуры нормализации и эквализации изображений.

В разделе 4.3 описано назначение и основные функции разработанного комплекса программ. Предложена трехуровневая структура программ, обеспечивающая гибкость и отсутствие дублирования в описании программных модулей, а также централизованная схема вызова приложений, позволяющая объединять программные модули с помощью динамически соединяемых библиотек. Программный комплекс разработан в среде Turbo C++ с использованием библиотеки OpenCV.

В разделе 4.4 обсуждены результаты четвертой главы. Основной вывод заключается в том, что предложенные трехуровневая структура программ и централизованная схема вызова приложений позволили ускорить разработку программного комплекса и обеспечить гибкость замены модулей при проведении экспериментов.

В пятой главе приведены результаты экспериментов по распознаванию полутоновых изображений подписей, цветных изображений лиц и объединенного источника подписей и лиц.

В разделе 5.1 описана схема проведения экспериментов. В эксперименте использованы подготовленные базы изображений подписей и лиц. Эксперименты проводились по схеме скользящего контроля с 30-кратным разбиением множества изображений источника на равные по объему обучающие и тестовые выборки (30 times 2 fold cross-validation) [S. Theodoridis, K. Koutroumbas, 2009]. На обучающих выборках моделировались классификаторы при значениях априорной вероятности $P_{own} = 1, 0.75, 0.5$. На тестовых выборках оценивались доли ошибок распознавания и их среднеквадратические отклонения при использовании алгоритмов Exhaustive Search и Guided Search в пространстве древовидных представлений с параметром разрешения $l = 9$.

В разделе 5.2 описаны результаты моделирования процедур обучения TSC-классификатора по критерию взвешенного голосования и ближайшего эталона (TSC_{wv} и TSC_{nn}) и SVM-классификатора для источников подписей, лиц и ансамбля указанных источников. Следует отметить, что в используемой версии SVM-классификатора отсутствует функция отказа, поэтому его моделирование выполнено для источников с априорной вероятностью $P_{own} = 1$.

В разделе 5.3 продемонстрированы результаты тестирования классификаторов. Экспериментальные характеристики распознавания полутоновых изображений подписей по каналу I с использованием алгоритма Exhaustive Search приведены в табл.1.

Таблица 1. Средние значения ε и среднеквадратические отклонения σ ошибок классификации подписей по каналу I (алгоритм Exhaustive Search)

Канал		$P_{own} = 1$			$P_{own} = 0.75$		$P_{own} = 0.5$	
		TSC _{NN}	TSC _{wv}	SVM	TSC _{NN}	TSC _{wv}	TSC _{NN}	TSC _{wv}
I	ε	0.013	0.005	2E-4	0.015	0.006	0.018	0.010
	σ	0.005	0.003	6E-4	0.007	0.007	0.009	0.007

Согласно табл.1, при $P_{own} = 1$ SVM-классификатор обеспечивает меньшую долю ошибок по сравнению с TSC_{wv} - классификатором (2E-4 против 0.005). При значении $P_{own} = 1$ доля ошибок TSC_{wv} - классификатора значительно ниже доли ошибок TSC_{NN} - классификатора (0.005 и 0.013 соответственно). При значениях $P_{own} = 0.75$ и 0.5 указанное соотношение ошибок для TSC_{wv} и TSC_{NN} - классификаторов сохраняется.

Аналогичные характеристики распознавания лиц по каналам HSI-модели и по объединению этих каналов с использованием GDM и WMV-стратегий и алгоритма Exhaustive Search приведены в табл.2.

Таблица 2. Средние значения ε и среднеквадратические отклонения σ ошибок классификации лиц в модели HSI (алгоритм Exhaustive Search)

Канал		$P_{own} = 1$			$P_{own} = 0.75$		$P_{own} = 0.5$	
		TSC _{NN}	TSC _{wv}	SVM	TSC _{NN}	TSC _{wv}	TSC _{NN}	TSC _{wv}
H	ε	0.015	0.008	0.002	0.016	0.010	0.018	0.012
	σ	0.005	0.006	0.001	0.007	0.006	0.007	0.007
S	ε	0.017	0.012	0.009	0.018	0.015	0.020	0.017
	σ	0.006	0.003	0.004	0.008	0.005	0.008	0.008
I	ε	0.022	0.012	0.017	0.025	0.014	0.027	0.020
	σ	0.006	0.005	0.006	0.008	0.005	0.008	0.007
HSI _{GDM}	ε	0.007	0.001	0.001	0.012	0.005	0.015	0.007
	σ	0.005	0.002	0.001	0.006	0.004	0.007	0.005
HSI _{WMV}	ε	0.010	0.005	0.006	0.015	0.007	0.015	0.009
	σ	0.005	0.003	0.005	0.007	0.005	0.007	0.005

Из табл. 2 следует, при $P_{own} = 1$ объединение классификаторов отдельных каналов цветовой модели HSI на основе GDM-стратегии (HSI_{GDM}) позволило уменьшить долю ошибок до 0.001. При $P_{own} = 1, 0.75$ и 0.5 результаты объединения с помощью WMV-стратегии (HSI_{WMV}) уступают GDM-стратегии.

Таблица 3. Средние значения ε и среднеквадратические отклонения σ ошибок идентификации личности по ансамблю изображений подписи по каналу I и лица в модели HSI (алгоритм Exhaustive Search)

Источник		$P_{own} = 1$			$P_{own} = 0.75$		$P_{own} = 0.5$	
		TSC _{NN}	TSC _{WV}	SVM	TSC _{NN}	TSC _{WV}	TSC _{NN}	TSC _{WV}
Ансамбль (I, HSI) _{GDM}	ε	0.003	3E-4	7E-5	0.007	0.001	0.010	0.003
	σ	0.004	7E-4	4E-4	0.004	0.001	0.005	0.002

Характеристики качества идентификации личности по ансамблю источников подписей и лиц с использованием GDM-стратегии объединения и алгоритма полного перебора (Exhaustive Search) даны в табл.3. Приведенные характеристики демонстрируют увеличение достоверности распознавания по сравнению с данными табл.1-2.

Экспериментальные результаты распознавания полутонных изображений подписей (канал I), цветных изображений лиц (модель HSI_{GDM}) и ансамбля изображений «подписи, лица» (I, HSI)_{GDM} с использованием алгоритма иерархического поиска решения (Guided Search) даны в табл.4. Результаты получены для TSC-классификатора в пространстве древовидных представлений объектов с уровнями разрешения $l = 1, \dots, 9$. Получение аналогичных характеристик SVM-классификатора требует модификации используемой версии SVM, что является предметом отдельного исследования.

Таблица 4. Среднее значение ошибок ε и вычислительного выигрыша

$$C = \frac{C_{ES}}{C_{GS}} \text{ при использовании алгоритма Guided Search}$$

Источник		$P_{own} = 1$		$P_{own} = 0.75$		$P_{own} = 0.5$	
		TSC _{NN}	TSC _{WV}	TSC _{NN}	TSC _{WV}	TSC _{NN}	TSC _{WV}
Подписи I	ε	0.025	0.015	0.031	0.020	0.034	0.025
	C	5.228	5.362	5.631	5.975	6.722	6.274
Лица HSI _{GDM}	ε	0.017	0.012	0.022	0.018	0.030	0.020
	C	5.975	5.228	7.605	8.366	8.714	8.496
Ансамбль (I, HSI) _{GDM}	ε	0.015	0.010	0.020	0.014	0.027	0.018
	C	5.542	5.400	7.967	8.184	8.556	8.366

Согласно данным табл.4, доли ошибок TSC - классификатора при использовании алгоритма Guided Search превышают аналогичные показатели алгоритма Exhaustive Search (табл.1-3). Однако при различных априорных распределениях «своих» и «чужих» объектов иерархический поиск решения продемонстрировал более чем пятикратный выигрыш в быстродействии по сравнению с алгоритмом полного перебора.

В заключении приведены основные результаты работы и рекомендации по их использованию. Обозначены возможные направления дальнейших исследований.

Приложение содержит копии актов об использовании теоретических и практических результатов диссертационной работы в учебно-производственных организациях.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Разработан способ древовидного представления многоканальных изображений многослойными наборами эллиптических примитивов. На множестве представлений введена мера различия объектов, и построена модель метрического TSC-классификатора по критерию взвешенного голосования и ближайшего эталона.

2. Получены оценки вычислительной сложности иерархического и переборного алгоритмов поиска решений. Показано, что при большом числе классов c иерархический алгоритм обеспечивает в $\frac{c}{\log_2 c}$ раз большее быстродействие по сравнению с алгоритмом полного перебора.

3. Разработана процедура обучения TSC-классификатора, включающая отбор эталонов и оценивание параметров. Отбор эталонов базируется на покрытии обучающего множества сферами и оптимизации покрытия (по числу и радиусам сфер) методом скользящего контроля. Предложен способ объединения классификаторов.

4. Реализован программный комплекс в среде Turbo C++ для проведения экспериментов с различными источниками изображений. Комплекс включает модули предобработки и представления данных, обучения и тестирования классификаторов. Предложена трехуровневая структура программ и централизованная схема вызова приложений, удобная для интеграции программных модулей.

5. Выполнены эксперименты по распознаванию полутоновых изображений подписей, цветных изображений лиц и ансамбля изображений «подписи, лица». Для указанных источников проведено моделирование процедур обучения и тестирования разработанного TSC-классификатора и известного SVM-классификатора. Получены экспериментальные оценки качества классификации и вычислительного выигрыша иерархического алгоритма относительно переборного алгоритма.

6. Разработанная модель классификатора универсальна и применима для широкого класса источников многоканальных изображений, которые содержат образы в виде двумерных тел с идентифицируемой системой собственных координат.

7. В перспективе планируется исследовать модификации схемы отбора эталонов и алгоритма иерархического поиска решения, а также расширить класс анализируемых источников изображений.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Степанов Д.Ю. Предобработка и представление данных для решения задачи распознавания лиц: Сборник трудов 58-й научно-технической конференции МИРЭА, г.Москва, 13-27 мая 2009 / Московский государственный институт радиотехники, электроники и автоматики (технический университет). – М.: МИРЭА, 2009. - с.116-121.

2. Степанов Д.Ю. Выделение контура лица для решения задачи распознавания лиц / М.: Свидетельство о государственной регистрации программного продукта №50200900489 от 02.06.2009.

3. Ланге М.М., Степанов Д.Ю. Многослойное древовидное представление объектов многоканальных изображений: 14-я Всероссийская конференция «Математические методы распознавания образов», г.Суздаль, 21-26 сентября 2009.: Сборник докладов. - М.: МАКС Пресс, 2009. - с.376-378.

4. Степанов Д.Ю. Идентификация личности на основе метода древовидных представлений: Сборник конкурсных работ Всероссийского смотра-конкурса научно-технического творчества студентов высших учебных заведений «Эврика-2009», г.Новочеркасск, декабрь 2009 / Южно-Российский государственный технический университет (Новочеркасский политехнический институт). – Новочеркасск: Лик, 2010. - с.54-57.

5. Степанов Д.Ю. Распознавание лиц в пространстве многослойных древовидных представлений изображений цветовой модели HSI // Труды Института системного анализа Российской академии наук (ИСА РАН) «Динамика неоднородных систем». - 2010. - т.53, №2. - с.305-314.

6. Степанов Д.Ю. Отбор эталонов на основе покрывающих сфер на этапе обучения классификатора // Естественные и технические науки. - 2010. - т.48, №4. - с.370-372.

7. Stepanov D.Y. Data preprocessing for person identification based on color face images: Proceedings of 16th International Scientific and Practical Conference of Students, Post-graduates and Young Scientists «Modern Techniques and Technologies», Tomsk, 12-16 April, 2010 / Tomsk Polytechnic University. - Tomsk: TPU Press, 2010. - pp. 121-123.

8. Степанов Д.Ю. Обучение классификатора на основе сфер покрытия / М.: Свидетельство о государственной регистрации программного продукта №50201001618 от 05.10.2010.

9. Lange M.M., Ganebnykh S.N., Stepanov D.Y. Multiresolution network of templates for fast pattern recognition: Proceedings of 10th International Conference «Pattern Recognition and Image Analysis: New Information Technologies», St. Petersburg, 5-12 December, 2010. – SPb.: Politechnika, 2010. vol.I - pp.149-152.

10. Степанов Д.Ю. Идентификация личности на основе распознавания лиц по древовидным представлениям изображений // Информационно-измерительные и управляющие системы. - 2011. - т.9, №7. - с.13-24.

11. Степанов Д.Ю., Ланге М.М. Распознавание лиц по многослойным древовидным представлениям цветных изображений: 15-я Всероссийская конференция «Математические методы распознавания образов», г. Петрозаводск, 11-17 сентября 2011.: Сборник докладов. - М.: МАКС Пресс, 2011. - с.330-333.

12. Ganebnykh S.N., Lange M.M., Stepanov D.Y. Metric classifier using multilevel network of templates // Pattern Recognition and Image Analysis. - 2012. - vol.22, №2. - pp. 265-277.

Личный вклад автора. В работе [3] автору принадлежит мера на множестве многослойных древовидных представлений, алгоритм выделения информативной области лица на изображении и экспериментальные оценки вероятности ошибки распознавания лиц по цветным изображениям. В [9] и [11] имеет место неразделимое соавторство. В [12] автору принадлежит схема проведения экспериментов и экспериментальные оценки вероятности ошибки распознавания подписей и лиц.

Степанов Дмитрий Юрьевич

МЕТОДЫ И АЛГОРИТМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ
С ИСПОЛЬЗОВАНИЕМ ДРЕВОВИДНЫХ ПРЕДСТАВЛЕНИЙ
МНОГОКАНАЛЬНЫХ ИЗОБРАЖЕНИЙ

Автореферат диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 28.08.2013.

Формат 60x84/16.

Уч.-изд. л. 1. Усл. печ. л. 1.

Тираж 100 экз. Заказ 18

Отпечатано на ротапринтах
в Федеральном государственном бюджетном учреждении науки
Вычислительный центр им. А.А. Дородницына Российской академии наук
119333, г.Москва, ул.Вавилова, д.40