

Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Национальный исследовательский
Томский политехнический университет»

На правах рукописи



Буй Тхи Тху Чанг

**АЛГОРИТМЫ РАСПОЗНАВАНИЯ ЛИЦ И ЖЕСТОВ НА
ОСНОВЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ И МЕТОДА
ГЛАВНЫХ КОМПОНЕНТ**

05.13.11 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

**Диссертация
на соискание ученой степени
кандидата технических наук**

Научный руководитель –
доктор технических наук,
профессор В.Г. Спицын

Томск – 2014

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	5
ГЛАВА 1. АНАЛИТИЧЕСКИЙ ОБЗОР МЕТОДОВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ	13
1.1 Применение и развитие методов распознавания объектов	13
1.2 Анализ подходов при решении задач распознавания объектов.	17
1.2.1 Формирование признаков изображения.....	17
1.2.2 Классификация методов распознавания объектов.....	21
1.2.3 Вейвлет-анализ	29
1.3 Обнаружение лиц на изображениях.....	37
1.4 Слежение за объектами	40
1.5 Цель и задачи исследования	42
1.6 Основные результаты и выводы по главе 1.....	45
ГЛАВА 2. АЛГОРИТМЫ РАСПОЗНАВАНИЯ ЛИЦ И ЖЕСТОВ НА ОСНОВЕ ПРИМЕНЕНИЯ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ И МЕТОДА ГЛАВНЫХ КОМПОНЕНТ	46
2.1 Способ выделения признаков объектов на изображениях	46
2.2 Алгоритм распознавания лиц и жестов на статических изображениях.....	49
2.2.1 Алгоритм распознавания объектов при применении PCA ...	49
2.2.2 Принципы и схемы работы предложенного алгоритма распознавания объектов на изображениях	51
2.3 Экспериментальное оценивание точности распознавания объектов на основе применения вейвлет-преобразований и PCA.....	54
2.4 Алгоритм распознавания лиц на видеопоследовательности.....	63
2.4.1 Интегральное представление изображений.....	65

2.4.2	Хаар-подобные признаки.....	66
2.4.3	Метод построения классификатора на основе алгоритма бустинга	69
2.4.4	Метод комбинирования классификаторов в каскадную структуру	70
2.5	Алгоритм распознавания жестов на видеопоследовательности	71
2.6	Основные результаты и выводы по главе 2.....	75
ГЛАВА 3. ТЕСТИРОВАНИЕ И ПРИМЕНЕНИЕ РАЗРАБОТАННЫХ АЛГОРИТМОВ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ И ЖЕСТОВ НА ИЗОБРАЖЕНИЯХ И ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ		76
3.1	Постановка экспериментов	76
3.2	Тестирование разработанного ПО.....	77
3.2.1	Распознавание лиц на статических изображениях.....	77
3.2.2	Распознавание жестов на статических изображениях.....	87
3.2.3	Распознавание множества лиц на видеопоследовательности... ..	94
3.2.4	Распознавание жестов на видеопоследовательности	98
3.3	Перспективы использования полученных результатов в жизни	99
3.4	Основные результаты и выводы по главе 3.....	101
ГЛАВА 4. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ И ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ .		102
4.1	Выбор средств разработки	102
4.1.1	Структура классов библиотеки	104
4.2	Программные средства для распознавания объектов на изображениях и видеопоследовательностях	116

4.2.1	Приложение для распознавания объектов на статических изображениях.....	117
4.2.2	Приложение для распознавания лиц на видеопоследовательностях.....	120
4.2.3	Приложение для распознавания жестов на видеопоследовательностях.....	123
4.3	Основные результаты и выводы по главе 4.....	125
ЗАКЛЮЧЕНИЕ		126
ОБОЗНАЧЕНИЯ.....		127
СПИСОК СОКРАЩЕНИЙ.....		128
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ И ЛИТЕРАТУРЫ.....		129
СПИСОК ПУБЛИКАЦИЙ АВТОРА		142
ПРИЛОЖЕНИЕ		145

ВВЕДЕНИЕ

Актуальность работы. В настоящее время успешно развивается направление, связанное с интеллектуализацией методов обработки и анализа данных. Задача распознавания лиц и жестов является одной из первых практических задач, которая послужила стимулом для развития теории распознавания объектов [1, 27–29, 42]. Распознавание лиц и жестов находит применение в различных сферах человеческой деятельности. Данное направление появилось в начале 1980-х годов, но его развитие началось в 1990-х годах в процессе создания информационно-поисковых систем распознавания лиц для идентификации личности.

В развитие теории и практики распознавания лиц и жестов значительный вклад внесли зарубежные ученые *Jonathon Phillips (NIST, США)*, *Harry Wechsler* (университет *George Mason*), работающие в области распознавания образов уже 15 лет, а также *Takeo Kanade*, *Matthew Turk* [113], *Paul Viola* [116–118], *Michael J. Jones* [116–118], *Bruce D. Lucas*, *Jing Xiao*, *Simon Baker*, *Dorin Comaniciu* [57–60], *Gary R. Bradski* [54] и другие. Интересные результаты в области распознавания образов и распознавания объектов на изображениях были получены российскими учеными: *Ванником В.Н.* [15], *Журавлёвым Ю.И.* [27, 28], *Сойфером В.А.* [42, 43], *Загоруйко Н.Г.* [29], *Местецким Л.М.* [31], *Алпатовым Б.А.* [1], *Гореликим А.Л.* [20, 21], *Гуревичем И.Б.* [21], *Сергеевым В.В.* [38], *Васиным Ю.Г.* [115] и другими. Ценный вклад в решение проблемы распознавания лиц внесли российские ученые: *Желтов С.Ю.* [25], *Визильтер Ю.В.* [25], *Крылов А.С.* [32], *Вежневцев В.П.* [16], *Шерстобитов А.И.* [46], *Пахирка А.И.* [34] и другие.

Распознавание лиц заключается в идентификации одного или нескольких человек путем сравнения входных изображений с имеющимися изображениями в базе данных (БД) [30]. Методы, применяемые для решения задачи распознавания лиц и жестов, должны обеспечивать приемлемую точность распознавания и высокую скорость обработки видеопоследовательностей.

Распознавание объектов состоит из четырех этапов: захвата объектов, извлечения признаков, сравнения полученных признаков с шаблонами и поиска соответствия. Большой интерес представляют задачи распознавания объектов. Решение этих задач востребовано в связи с возрастающими практическими потребностями правоохранительных органов для осуществления процедур верификации, криминалистической экспертизы и т.д. [33].

Проблема автоматизированного распознавания лиц и жестов является относительно новой и до сих пор не решена в полном объеме. За последние годы был предложен ряд различных методов и алгоритмов по обработке, локализации и распознаванию лиц и жестов на статических изображениях, таких как «собственные лица» (метод главных компонент, *Principal Component Analysis*, PCA), нейронные сети, эволюционные алгоритмы, алгоритм *AdaBoost*, метод опорных векторов и т.д. Однако эти подходы для распознавания объектов обладают недостаточной точностью, надежностью и скоростью в сложной реальной обстановке, характеризующейся присутствием шумов на изображениях и на видеопоследовательностях.

Следует отметить существование ряда факторов, усложняющих распознавание объектов на статических изображениях и видеопоследовательностях. К ним относятся: изменение условий освещенности в процессе съемки, низкое качество видеоизображений, сложность отделения объекта от фона, присутствие в видеокадре множества объектов и т.д.

Таким образом, необходимо совершенствовать подходы, методы и алгоритмы распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях в интересах разнообразных применений. В частности, разрабатываемые алгоритмы могут найти применение при создании человеко-машинных интерфейсов, которые расширяют интеллектуальные возможности компьютерных систем.

Целью диссертационной работы является разработка алгоритма, позволяющего распознавать лица и жесты на статических изображениях и видеопоследовательностях в режиме реального времени.

Для достижения указанной цели были поставлены следующие **основные задачи:**

1. Разработать способ выделения признаков объектов на изображениях, обеспечивающий высокую точность при решении задач распознавания лиц и жестов.

2. Разработать алгоритм на основе предложенного способа выделения признаков объектов, позволяющий распознавать лица и жесты в присутствии шума на статических изображениях.

3. Разработать алгоритм на основе предложенного алгоритма распознавания лиц на изображениях, позволяющий распознавать лица на видеопоследовательностях в режиме реального времени.

4. Разработать алгоритм на основе предложенного алгоритма распознавания жестов на изображениях, позволяющий распознавать жесты на видеопоследовательностях в режиме реального времени.

5. Реализовать разработанные алгоритмы в виде программных систем (ПС), предназначенных для распознавания лиц и жестов, и провести вычислительные эксперименты с целью оценки их качества и эффективности.

Степень достоверности результатов проведённых исследований подтверждается результатами численных экспериментов, согласующимися с результатами, полученными другими авторами.

Внедрение работы. Результаты работы внедрены в Томском политехническом университете на кафедре вычислительной техники при подготовке специалистов по магистерской программе «Компьютерный анализ и интерпретация данных» по курсу «Методы интеллектуальной обработки и анализа изображений».

Реализация результатов работы. Способы, алгоритмы и программные средства, разработанные в диссертационной работе, использовались при выполнении работ по гранту РФФИ № 09-08-00309 «Создание программного комплекса автоматизированной обработки изображений и распознавания образов на основе применения искусственных нейронных сетей, регуляторных сетей и эволюционных алгоритмов» (2009–2011 гг.); в проекте «Создание комплексных технологий распознавания объектов на изображении на основе применения моделей зрительного восприятия и методов вычислительного интеллекта», поддержанном грантом РФФИ № 12-08-00296 (2012–2014 гг.).

Предметом исследования являются алгоритмы для распознавания лиц и жестов на статических изображениях и видеопоследовательностях в режиме реального времени.

Теоретическая значимость результатов диссертационного исследования состоит в том, что разработаны, исследованы и апробированы новые алгоритмы для распознавания лиц и жестов на изображениях и видеопоследовательностях в режиме реального времени.

Апробация работы. Результаты диссертационной работы обсуждались и докладывались на следующих симпозиумах, семинарах и конференциях: VIII, IX Всероссийские научно-практические конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии» (Томск, 2010, 2011); XVI, XVII Международные научно-практические конференции студентов, аспирантов и молодых ученых «Современные техника и технологии» (Томск, 2010, 2011); III Всероссийская научно-практическая конференция «Научная инициатива иностранных студентов и аспирантов российских вузов» (Томск, 2010); VI Международная научно-практическая конференция «Электронные средства и системы управления» (Томск, 2010); XIX Всероссийской семинар «Нейроинформатика, ее приложения и анализ данных» (Красноярск, 2011); XIV Всероссийский с международным участием научный симпозиум по

теории и приложениям непараметрических и робастных статистических методов «НЕПАРАМЕТРИКА-XIV» (Томск, 2012); The 7th International Forum on Strategic Technology IFOST (Томск, 2012).

Кратко изложим содержание **основных разделов работы**.

В **первой главе** представлен аналитический обзор алгоритмов и методов при решении задач распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях в режиме реального времени, обозначены их преимущества и недостатки.

Проводится обсуждение одномерного и двумерного вейвлет-преобразований. Проанализированы возможности совместного использования вейвлет-преобразований и РСА при решении задач распознавания лиц и жестов как на статических изображениях, так и на видеопоследовательностях в режиме реального времени.

Во **второй главе** описываются алгоритмы для распознавания объектов на статических изображениях и динамических объектов на видеопоследовательностях в режиме реального времени, разработанные в соответствии с требованиями, сформулированными в главе 1. Предложены алгоритмы, успешно работающие как со статическими изображениями, так и с видеопоследовательностями. Особенностью предложенных алгоритмов является совместное применение вейвлет-преобразований, РСА, метода Виолы-Джонса и алгоритма *CAMShift*.

В **третьей главе** излагаются результаты применения разработанных алгоритмов к решению задачи распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательности в режиме реального времени. Представлены данные для обучения и тестирования, а также таблицы с результатами распознавания объектов. Произведено тестирование разработанного программного обеспечения (ПО) и сопоставление предложенных алгоритмов распознавания объектов с известными аналогами.

Четвёртая глава посвящена разработке ПО на основе предложенных во 2 главе алгоритмов. Выбран язык программирования для реализации библиотеки. Описываются используемые библиотеки *OpenCV*, *Emgu CV* и *AForge.Net*. Представлена структура классов библиотеки.

Научная новизна результатов диссертационной работы заключается в следующем:

1. Предложен способ выделения признаков объектов на изображениях, основанный на совместном применении вейвлет-преобразований Хаара и Добеши, позволяющий эффективнее распознавать объекты по сравнению с применением вейвлет-преобразования Хаара или Добеши в отдельности.

2. Разработан алгоритм распознавания лиц и жестов на статических изображениях в присутствии шума, основанный на предложенном способе выделения признаков и PCA, обеспечивающий высокую точность распознавания лиц и жестов.

3. Предложен алгоритм, основанный на совместном применении предложенного алгоритма распознавания лиц на изображениях и метода Виолы-Джонса, позволяющий распознавать множество лиц на видеопоследовательности.

4. Разработан алгоритм, основанный на применении предложенного алгоритма распознавания жестов на изображениях, метода Виолы-Джонса и алгоритма *CAMShift*, дающий возможность распознавания жестов на видеопоследовательности в режиме реального времени.

Научную ценность работы представляет вклад в развитие в области распознавания объектов, заключающийся в предложенном алгоритме, основанном на совместном применении вейвлет-преобразований, PCA, метода Виолы-Джонса и алгоритма *CAMShift*, увеличивающем точность и скорость распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательности в режиме реального времени.

Практическая ценность. Разработанные в диссертации алгоритмические и программные средства могут найти применение в системах правоохранительных органов для верификации личности, осуществления криминалистической экспертизы, а также при проведении телеконференций.

Реализованные в ходе диссертационной работы алгоритмы предназначены для решения задач распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях.

Методы исследования. При выполнении диссертационной работы использовались вейвлет-преобразования, РСА, метод Виолы-Джонса, алгоритм *CAMShift*, алгоритмы цифровой обработки изображений, технология разработки ПО, а также методы теории вероятностей и математической статистики для количественной обработки данных экспериментов.

Личный вклад автора. Постановка задач исследования по теме диссертации выполнена автором совместно с научным руководителем, д.т.н., профессором В.Г. Спицыным. Основные теоретические и практические результаты, представленные в диссертации, получены лично автором.

Основные положения, выносимые на защиту:

1. Способ выделения признаков объектов на изображениях на основе совместного применения вейвлет-преобразований Хаара и Добеши.
2. Алгоритм распознавания лиц и жестов в присутствии шума на статических изображениях, основанный на предложенном способе выделения признаков объектов на изображениях и РСА.
3. Алгоритм, основанный на совместном применении предложенного алгоритма распознавания лиц на изображениях и метода Виолы-Джонса, позволяющий распознавать множество лиц на видеопоследовательности.
4. Алгоритм, основанный на применении предложенного алгоритма распознавания жестов на изображениях, метода Виолы-Джонса и алгоритма

CAMShift, дающий возможность распознавания жестов на видеопоследовательности в режиме реального времени.

Автор выражает глубокую благодарность научному руководителю, профессору, доктору технических наук В.Г. Спицыну за помощь в написании работы, ценные замечания и доброжелательную критику. Также автор благодарит за ценные замечания доктора технических наук, профессора, заведующего кафедрой Вычислительной техники Н.Г. Маркова, кандидатов технических наук Ю.Р. Цоя, А.А. Белоусова, Ю.А. Болотову. Автор выражает благодарность за ценные замечания и помощь в выполнении работы членам научной группы профессора В.Г. Спицына.

Структура и объем работы. Диссертация содержит введение, четыре главы, заключение и список использованной литературы, содержащий 124 наименования. Общий объем диссертации составляет 145 страниц машинописного текста, иллюстрированного 82 рисунками и 20 таблицами.

ГЛАВА 1. АНАЛИТИЧЕСКИЙ ОБЗОР МЕТОДОВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

В данной главе представлен аналитический обзор алгоритмов и методов распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях в режиме реального времени. Приведены их преимущества и недостатки. Рассматривается возможность использования вейвлет-преобразований и PCA при решении задач распознавания лиц и жестов как на статических изображениях, так и на видеопоследовательностях.

1.1 Применение и развитие методов распознавания объектов

В книге Гонсалеса [19] предложено понятие образа: под *образом* понимается некоторая упорядоченная совокупность признаков. Совокупность образов, обладающих некоторыми общими свойствами, называется *классом*. Под машинным распознаванием образов понимаются методы, позволяющие относить образы к тем или иным классам – автоматически или с минимальным вмешательством человека.

Распознавание объектов включает в себя постановку задачи построения и применения формальных операций над числовыми или символьными изображениями объектов. Результаты распознавания отражают отношения эквивалентности между этими объектами и объектами, имеющимися в БД [40].

Задача распознавания лиц и жестов является одной из самых востребованных в таких областях как: цифровая обработка изображений, компьютерное зрение, биометрия, организация видеоконференций, создание интеллектуальных систем безопасности и т.п.

Люди, как правило, легко решают задачу распознавания объектов, но практика показала, что эта задача для компьютера не такая уж простая. Распознавание объектов представляет собой процесс отнесения исходных

данных к определенному классу с помощью выделения признаков, которые характеризуют эти данные, из общей массы несущественных данных [15, 20].

За последние годы распознавание лиц находит применение во многих сферах человеческой деятельности. Человек использует распознавание лиц для идентификации личности. Достижения в области вычислительной техники за последние несколько десятилетий позволяют осуществлять подобные операции автоматически.

Раньше в алгоритмах распознавания лиц использовались простые геометрические модели. На сегодняшний день процесс распознавания лиц осуществляется на основе сложных математических представлений соответствующих процессов. Например, в ПО для управления веб-альбомами *iPhoto*, *Picasa* и др. используется функция идентификации людей на фотографиях [23].

Проблема автоматизированного распознавания лица является относительно новой. В разработанной в 1960-х годах первой полуавтоматической системе для распознавания лиц требуется администратор, чтобы выделить особенности (такие как глаза, уши, нос и рот) на фотографиях. Затем рассчитываются расстояния от выделенных характерных точек до общей контрольной точки, которые сравниваются с исходными данными.

В 1970-х годах *Goldstein*, *Harmon* и *Lesk* использовали 21 субъективный маркер, в частности, цвет волос и толщину губ для автоматизации процесса распознавания [67]. Основным недостатком в указанных способах решения являлось то, что определение местоположения характерных точек на изображениях и расстояний между ними производилось вручную.

В 1988 г. *Sirovich* и *Kirby* применили PCA для решения проблемы распознавания лиц [104]. Этот подход явился важной вехой, поскольку в указанном методе осуществлялось уменьшение размерности данных с потерей наименьшего количества информации.

В 1991 г. Turk и Pentland открыли, что при использовании метода построения «собственных лиц» (*eigenfaces*) остаточная ошибка может быть применена для обнаружения лица на изображениях [113]. Это открытие позволило создать надежные автоматизированные системы распознавания лиц в реальном режиме времени. Хотя такой подход был несколько ограничен фактором наличия определенных свойств окружающей среды, он, тем не менее, продемонстрировал необходимость дальнейшего развития автоматизированной технологии распознавания лиц.

Указанная технология впервые привлекла внимание общественности во время тестовых испытаний программной реализации в январе 2001 году (*Super Bowl*) [100]. Представленная система собирала наблюдаемые изображения и сравнивала их с БД цифровых фотографий.

Распознавание объектов на статических изображениях и видеопоследовательностях можно разделить на два основных этапа.

- Первый этап – нахождение и выделение объектов на изображениях. Такой этап можно назвать «детектированием объектов».
- Второй этап включает в себя непосредственно распознавание объекта, т.е. установление степени сходства с одним из объектов, «известных» системе.

Практическая разработка системы классификации осуществлялась по следующей схеме (рис. 1.1).



Рис. 1.1. Основные элементы системы распознавания объектов [31]

В связи с развитием технологий задача обнаружения движущихся или появляющихся объектов на видеопоследовательности в режиме реального времени является чрезвычайно актуальной для потребителей систем обнаружения и сопровождения объектов [30, 68, 92].

Задача распознавания объектов на видеопоследовательности может быть сформулирована следующим образом:

- определить в данном кадре последовательности наличие или отсутствие объектов;
- если ответ был «Да», то найти прямоугольную область, в которой содержится объект.

Схема процесса распознавания объектов на видеопоследовательности представлена на рис. 1.2.



Рис. 1.2. Схема процесса распознавания объектов на видеопоследовательности [70]

В настоящее время достигнуты значительные успехи при решении задач обнаружения неподвижного объекта или по-другому – обнаружения объекта на статических изображениях. Однако существует несколько нерешенных проблем при распознавании объектов на видеопоследовательности.

Факторы, усложняющие процесс выделения лиц на видеопоследовательностях [30]:

- низкое разрешение видеоизображения;

- изменение изображения лица существенно зависит от условий освещения, угла поворота, позы, присутствия солнечных очков, перекрытия части лица другими объектами;
- присутствие объектов на заднем плане;
- детектирование лиц на видеопоследовательности должно осуществляться в реальном режиме времени, что накладывает жесткие ограничения по сравнению с методами выделения объектов на статических изображениях. Требуется высокая скорость детектирования лиц на каждом кадре;
- число лиц на изображении заранее неизвестно.

1.2 Анализ подходов при решении задач распознавания объектов

Большой интерес представляют задачи распознавания лиц и жестов. Решение этих задач востребовано в связи с возрастающими практическими потребностями правоохранительных органов для осуществления процедур верификации, криминалистической экспертизы и т.д.

При всем многообразии методов и алгоритмов распознавания изображений, типичный подход к решению проблемы распознавания объектов состоит из следующих этапов: преобразование исходного изображения в начальное представление, выделение ключевых характеристик и осуществление классификации [37].

1.2.1 Формирование признаков изображения

Как известно, изображение можно представить в виде массива пикселей. При этом величина значения одного пикселя изображения практически ничего не значит. Это делает такое представление изображения избыточным и неэкономичным. При решении задачи распознавания объектов требуется намного меньше информации. Таким образом, для эффективного

распознавания объектов необходимо разработать некоторый компактный и удобный формат представления изображений.

Предположим, что задано множество объектов A и множество допустимых ответов B . Пусть $g:A\rightarrow B$ называется решающей функцией. Решающая функция g должна допускать эффективную компьютерную реализацию; по этой причине её также называется алгоритмом. *Признак (feature) f* объекта a – отображение $f:A\rightarrow D_f$, где D_f – множество допустимых значений признака. В частности, любой алгоритм $g:A\rightarrow B$ также можно рассматривать как признак. Если задан набор признаков $f_1, \dots, f_n$, то вектор $x=(f_1(a), \dots, f_n(a))$ называется признаковым описанием объекта $a\in A$. Признаковые описания допустимо отождествлять с самими объектами. При этом множество $A=(D_{f_1}\times \dots \times D_{f_n})$ называют *признаковым пространством* [31].

Измерения, используемые для классификации образов, называются *признаками*. *Вектором значений признаков* является совокупность признаков, относящихся к одному образу. Векторы значений признаков принимают значения в пространстве признаков. При решении задачи распознавания каждому образу ставится в соответствие единственное значение вектора признаков [31].

Правило отнесения образа к одному из классов на основании его вектора значений признаков называется *классификатором* или *решающим правилом*. *Признаком изображения* является численная характеристика, описывающая те или иные свойства изображения [36].

Признаки играют важную роль при сегментации изображений, а также при решении задач выделения, обнаружения и распознавания объектов. Общеизвестным является тот факт, что решение задач классификации зависит от удачно выбранной системы признаков. Из-за сложности процесса построение признаков изображения до настоящего времени остается процедурой эвристической [42].

Существует много причин использования признаков, а не значений пикселей непосредственно. Основной причиной является то, что признаки могут описать те знания о классе объектов, которые трудно выявить на конечном числе обучающих данных. Вторая важная причина применения признаков состоит в том, что системы, построенные на их основе, работают намного быстрее, чем системы, работающие напрямую с пикселями.

Признаки можно разделить на две группы: *естественные* и *искусственные* [1]. В первой группе признаки могут быть установлены непосредственным анализом изображения. Набор естественных признаков, которыми обладает изображение, определяется типом используемого регистратора изображений.

Искусственные признаки получаются путем обработки естественных признаков. По характеру пространственного описания изображения искусственные признаки можно разделить на две группы: локальные и глобальные. Локальные признаки изображения обычно рассчитываются для каждой точки изображения с учетом значений из некоторой окрестности этой точки. Глобальные признаки характеризуют изображение в целом и вычисляются с участием всех точек изображения.

По характеру описания временных процессов, происходящих в последовательности изображений, признаки можно разделить на две группы: *статические* и *динамические*.

В первой группе признаки описывают лишь пространственные свойства изображения. Для их вычисления используется лишь одно изображение.

Динамические признаки характеризуют процессы, протекающие с изображением во времени. Динамические признаки рассчитываются на основе последовательности изображений.

Извлечение признаков изображений (*feature extraction*)

Поскольку объем обрабатываемой информации быстро возрастает, то извлечение наиболее значимой информации из данных в виде признаков

стало очень востребованным. Одним из преимуществ этого подхода является то, что он позволяет людям лучше и быстрее понять основные свойства зашумленных наборов данных. Если характерный признак информации найден, то он также может использоваться для сжатия, сравнения или выявления некоторых особенностей соответствующей информации.

Процедура извлечения признаков находит одно из основных применений в анализе изображений. Ввиду существенного увеличения количества изображений и видео в нашей жизни, все большую значимость приобретает применение интеллектуальных программ и систем для идентификации объектов, представленных на изображении или для обнаружения объектов и слежения за ними на видеопоследовательностях в реальном режиме времени.

Обеспечение адекватной сложности системы признаков для получаемой точности описания изображения и ценности извлекаемой информации играет главную роль в задаче распознавания объектов. Предложено множество алгоритмов и методов для представления изображений. Одним из самых распространенных методов является *частотное преобразование*. Значение интенсивности пикселя переводится в набор частотных коэффициентов, малозначащие коэффициенты (соответствующие высоким частотам) отбрасываются. В данном случае высокие частоты характеризуют мелкие шумовые вариации изображения, поэтому ими можно пренебречь [26].

Выбор признаков

Допустим, что дан набор пикселей, которые предположительно связаны вместе и должны быть классифицированы. Какие признаки следует представить классификатору? Одно из возможных решений – представить значения всех пикселей. Это даст классификатору максимально возможный объем информации о наборе пикселей, но создаст множество проблем [45].

Во-первых, для пространства высокой размерности потребуется большое число образцов для простого представления доступных

возможностей. Например, лицо при низком разрешении имеет довольно простую структуру. В то же время, если рассматривать лица с высоким разрешением, то может быть затруднительно поддерживать достаточное число образцов, чтобы определить, является ли данная структура существенной и что незначительные изменения текстуры кожи несущественны. Взамен этого удобно выбрать пространство признаков, которое сделает эти свойства очевидными, для этого, как правило, вводится некоторая структура образцов.

Во-вторых, заранее могут быть известны некоторые свойства образов. Лучше использовать признаки, согласующиеся с нашими знаниями об образе. Такое решение может включать использование предварительно обработанных областей или выбор признаков, инвариантных при определенных преобразованиях.

Если значение одного признака можно точно предсказать по значению других, то этот признак очевидно является избыточным и его можно опустить. Отбрасывание признаков – это не единственная доступная возможность, например, можно создавать новые признаки, являющиеся функциями старых.

1.2.2 Классификация методов распознавания объектов

Предложено множество алгоритмов и методов для решения задач распознавания объектов. Они могут быть классифицированы по разным критериям. В работе [67] предлагается разделить методы распознавания лиц на две группы: аналитические и холистические. Первая группа состоит из методов, основанных на выделении геометрических признаков лица, описывающих его индивидуальные особенности. Для распознавания лиц используется вектор значений признаков, сформулированных на основе местоположения различных черт лица и особенностей его контура.

Во второй группе рассматриваются общие свойства изображений человеческих лиц. В методах данной группы используется представление

изображения всего лица на основе набора интенсивностей составляющих его пикселей.

В методах распознавания, основанных на сопоставлении, каждый класс представляется вектором значений признаков объекта, являющегося прототипом этого класса. Неизвестный объект приписывается к тому классу, прототип которого оказывается ближайшим в смысле заранее заданной метрики. Простейший подход заключается в использовании классификатора, который вычисляет Евклидовы расстояния между вектором значений признаков неизвестного объекта и каждым вектором возможного прототипа. Принадлежность объекта к известному классу определяется по минимальному расстоянию Евклида [19].

В работе [34] приведена классификация методов распознавания лиц. Указанные методы делятся на три категории. В первой категории рассматриваются глобальные свойства изображения лица. Во второй категории содержатся методы, основанные на использовании геометрических свойств. Третья категория является объединением методов первых двух категорий.

Распознавание лиц на видеопоследовательности обладает преимуществом по сравнению с распознаванием лиц на статических изображениях, так как анализ движения объектов позволяет получить дополнительную информацию [55, 78]. До настоящего времени разработано небольшое количество подходов для распознавания лиц на видеопоследовательности в реальном режиме времени.

В книге Лукьяницы [30] приведена категория методов, основанных на динамической информации. Распознавание объектов на видеопоследовательности в режиме реального времени состоит из двух этапов. Первый этап – трекинг и детектирование для определения местоположения объекта в текущем кадре. Второй этап – распознавание объектов.

В данной диссертационной работе методы распознавания объектов делятся по используемым признакам объектов (табл. 1.1).

Таблица 1.1. Классификация методов распознавания объектов

Категории	Методы
Адаптивные методы распознавания объектов и классификации	• Метод опорных векторов • Алгоритм <i>Adaboost</i> • Искусственные нейронные сети
Линейные преобразования	• Линейный дискриминантный анализ • Метод главных компонент

Адаптивные методы распознавания объектов основаны на автоматической подстройке к свойствам обрабатываемых данных, что позволяет разработать распознающую систему с приемлемыми характеристиками.

Ко второй категории относятся методы, характеризующиеся сокращением размерности данных.

а. Адаптивные методы распознавания объектов

Сложность задачи распознавания объектов заключается в том, что характерные признаки являются многомерными векторами, в совокупности описывающими некоторые свойства рассматриваемых объектов, например: текстуры, набор моментов, гистограммы и т.п. Среди адаптивных методов можно выделить следующие разновидности [48]:

- метод опорных векторов;
- искусственные нейронные сети;
- алгоритм *Adaboost*.

Метод опорных векторов

Метод опорных векторов (*Support Vector Machines, SVM*) был предложен отечественными учеными В. Вапником и А. Червоненкисом в 1995 году [114]. Он используется для решения задач классификации и

регрессионного анализа. В работе *Joachims* в 1998 году впервые применен SVM к задаче классификации текстов [74]. Особые свойства SVM – непрерывное уменьшение эмпирической ошибки классификации и увеличение зазора.

Применение SVM к решению задачи обнаружения лица заключается в поиске гиперплоскости в признаковом пространстве, отделяющей класс изображений «лицо» от класса изображений «не лицо».

В работе [83] представлена комбинация Байесовского подхода и SVM для распознавания лиц. Используются две БД *FERET* и *XM2VTS* для тестирования работоспособности алгоритма. В работе [108] показана эффективность применения SVM для распознавания жестов, информация о которых поступает из зафиксированных веб-камер. Использование SVM и мешка признаков (*Bag-of-Features*) оказалось эффективным в задачах обнаружения и распознавания жестов на видеопоследовательности [61, 92], а также при решении других задач [72, 81].

К основному недостатку SVM относятся существенные временные затраты при осуществлении процесса настройки. Из-за этого скорость распознавания объектов уменьшается. Кроме того, SVM требует большого объема памяти.

Алгоритм AdaBoost

Роберт Шапиро разработал алгоритм *AdaBoost* (адаптивность и усиление) в 1996 году. Этот алгоритм был успешно использован во многих областях, в частности при решении задач поиска лиц на изображениях.

Система поиска объектов на изображениях *Viola-Jones Object Detector*, основанная на *AdaBoost*, показывает высокую эффективность как по уровню распознавания, так и по скорости работы.

В качестве характерных признаков в простейшем случае можно использовать вейвлеты Хаара [116]. Однако размерность признаков, получающихся в этом случае, слишком велика. Таким образом, представляется целесообразным применение алгоритма бустинга для

уменьшения размерности данных и выбора основных признаков. После получения заданного числа характерных признаков с помощью алгоритма *AdaBoost* к ним может быть применен PCA для дальнейшего снижения размерности [109].

Успешное использование бустинга в задачах обнаружения объектов на изображениях позволяет рассмотреть возможность их применения при распознавании лиц. К основным преимуществам *AdaBoost* относятся высокая эффективность распознавания и простота реализации [30]. Недостатком алгоритма *AdaBoost* является требование неограниченного числа примеров для обучения, что не всегда можно выполнить на практике.

Искусственные нейронные сети

С точки зрения машинного обучения, *искусственные нейронные сети* (ИНС) являются вычислительными моделями, которые способны решать задачи распознавания образов. В нейронных сетях часто рассматривается задача определения принадлежности объекта к одному из заранее определенных классов на основе анализа совокупности его признаков.

Нейронные сети успешно применяются для решения задач распознавания объектов. При решении задачи обнаружения лица применялось большое количество нейронных сетей различных архитектур [98], в частности: вероятностные нейронные сети (*probabilistic decision-based neural networks*) [84], многослойные персептроны [75] и т.д.

Способность адаптации видеосъемки к сложным условиям является преимуществом при использовании нейронных сетей для распознавания объектов. Однако выбор оптимальной архитектуры сети требует немало времени. К недостатку следует отнести необходимость в тщательной и кропотливой настройке нейронной сети для получения удовлетворительного результата классификации.

Следует отметить, что нейронная сеть используется в задаче распознавания и отслеживания жестов на видеопоследовательности [103, 107, 111].

в. Линейные преобразования

Изображение рассматривается как матрица, значения элементов которой равны интенсивностям соответствующих пикселей. Такую матрицу можно представить в виде одномерного вектора. Полученный вектор характеризуется большой размерностью, а для анализа необходимо получить данные меньшей размерности. Таким образом, основной целью в данном разделе методов является нахождение оптимального уменьшения размерности множества признаков, позволяющего представить основные особенности каждого объекта.

Линейный дискриминантный анализ

При применении метода линейного дискриминантного анализа (*Linear Discriminant Analysis*, LDA) осуществляется проекция пространства изображений на пространство признаков [52, 69]. Выбор проекции удовлетворяет условию минимизации внутриклассового и максимизации межклассового расстояния в пространстве признаков.

LDA применяется для уменьшения размерности пространства. Указанный метод используется для построения достаточно простых классификаторов. При этом создается обучающая последовательность изображений лиц, в которой известна принадлежность каждого изображения к одному из имеющихся классов. LDA максимизирует отношение дисперсии между различными классами и внутриклассовой дисперсией.

LDA используется для нахождения проекции в пространстве, в которой разница между различными классами объектов была бы максимальной. Это требование формулируется как получение максимально компактных кластеров, соответствующих различным классам, удаленным на максимально возможное расстояние. При применении LDA получается подпространство небольшой размерности, в котором минимально пересекаются кластеры изображений «лицо» либо «не лицо» [123].

В работе [52] содержится обучающий набор лиц при базовых условиях освещённости, на основе которых при помощи линейных комбинаций можно

получить любые другие условия освещённости. LDA обладает высокой точностью распознавания при различных условиях освещённости, различных эмоциях, наличии или отсутствии очков.

При решении задач распознавания с использованием LDA возникают следующие вопросы:

- 1) можно ли применять LDA в больших базах изображений?
- 2) будет ли LDA работоспособен при одновременном изменении ракурса и условий освещения.

Метод главных компонент

Метод главных компонент (Principal Component Analysis, PCA) является статистическим методом. PCA находит применение во многих областях, таких как распознавание образов, компьютерное зрение, сжатие данных и т.п. Вычисление главных компонент сводится к вычислению собственных векторов и собственных значений ковариационной матрицы исходных данных или к сингулярному разложению матрицы данных.

Основной идеей применения PCA для распознавания лиц является введение понятия «*собственных лиц*» (*eigenfaces*), представляющих изображения лиц в виде набора главных компонент изображений [113].

Определяющее свойство «собственных лиц» заключается в том, что изображение, соответствующее каждому такому вектору, имеет лицеподобную форму. Для получения главных компонент вычисляются собственные векторы и собственные значения ковариационной матрицы, которые рассчитываются на основе изображения. Реконструкцией изображения является сумма главных компонент, умноженных на соответствующие собственные векторы.

Главные компоненты вычисляются для каждого изображения лица. Остальные компоненты кодируют мелкие различия между лицами и шум. Процесс распознавания лиц заключается в сравнении главных компонент неизвестного изображения с главными компонентами всех известных изображений. Изображения лиц, соответствующие одному человеку,

группируются в кластеры в пространстве собственных лиц. Из БД выбираются изображения-кандидаты, имеющие наименьшее расстояние от неизвестного входного изображения [8].

Основные преимущества применения РСА проявляются при хранении и поиске изображений в больших базах изображений, а также при реконструкции изображений [88]. Недостатки РСА – высокие требования к условиям съёмки изображений.

Следует отметить, что при решении задач распознавания объектов на основе применения РСА в работе [91] получена высокую точность. В работе [34] предлагается алгоритм для распознавания объектов (рис. 1.3), который позволяет уменьшить размерность данных с минимизацией потерь информации.

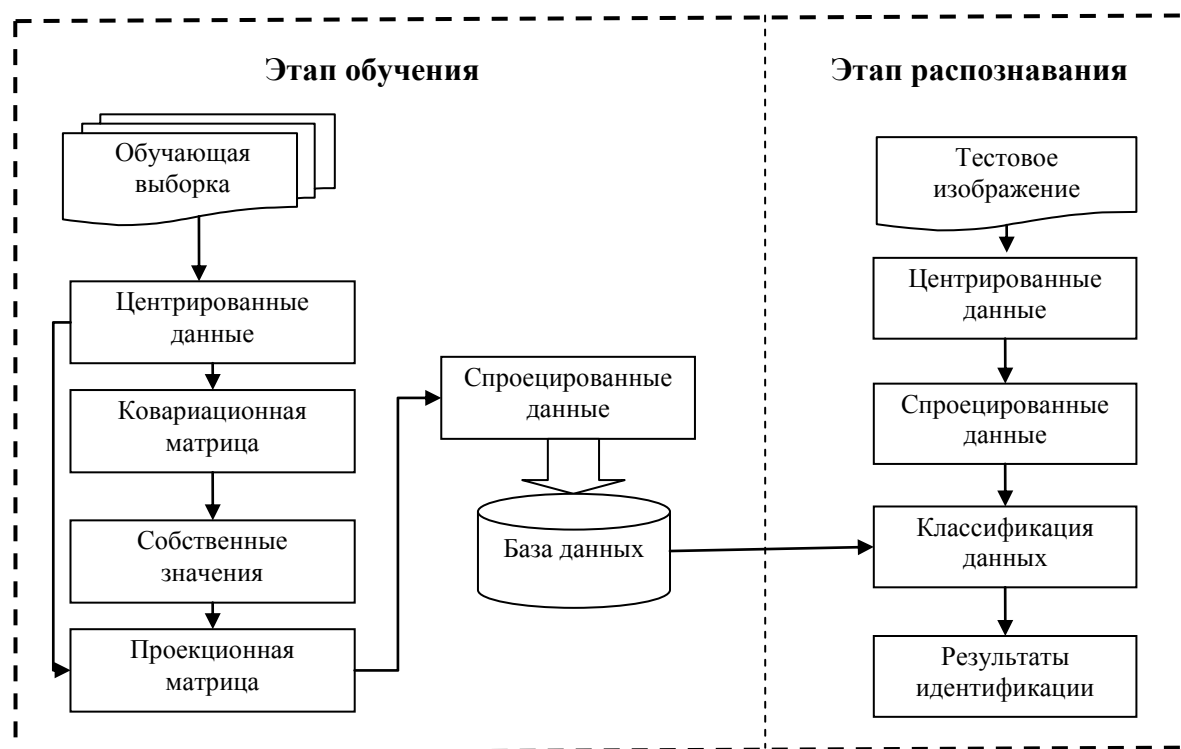


Рис. 1.3. Распознавание лиц на основе РСА [34]

Указанный подход часто используется при решении задачи распознавания лиц на статических изображениях, например, в следующих работах [56, 79, 86, 90]. В работе [71] приведена таблица сравнения методов для классификации облачных изображений (табл. 1.2). Как следует из

приведенной таблицы, PCA является наилучшим методом для классификации изображения.

Таблица 1.2. Результаты распознавания различными методами

Метод	Точность, %	Ошибка, %
PCA	88,85	1,2
<i>Back Propagation Neural Networks</i>	71,80	–
<i>Fisher Discriminant Analysis</i>	64,00	–
<i>Вейвлет-преобразование</i>	78,30	3,9
<i>Non-Negative Matrix Factorization</i>	69,94	8,1
<i>Radial Basis Function Neural Network</i>	73,20	7,3
<i>Метод опорных векторов</i>	84,11	–
<i>Probabilistic Neural Networks</i>	86,01	–
<i>K-SOM (K-Self Organizing Maps) Neural Networks</i>	80,00	–
<i>Fuzzy Logic Neural Networks</i>	81,00	3,4

1.2.3 Вейвлет-анализ

В последнее время возникло научное направление, связанное с так называемым *вейвлетами*¹. Вейвлеты широко применяются в различных областях знаний при решении ряда задач, таких как: распознавание образов, обработка и синтез сигналов, сжатие и обработка изображений, обучение нейросетей и т.д. Обычно под вейвлетами понимаются функции, сдвиги и растяжения которых образуют базис многих важных пространств. Эти функции являются компактными как во временной, так и в частотной области. Вейвлеты могут быть ортогональными, полуортогональными и биортогональными. *Вейвлет-коэффициенты* определяются интегральным преобразованием сигнала [19].

¹«Термин *вейвлет* (от английского слова «*wavelet*» – маленькая, небольшая в смысле продолжительности волна), утвердился в русской литературе, относящейся к областям обработки информации, сигналов и изображений. В математической литературе получил распространение предложенный К.И. Осколковым термин *всплеск*» – Прим. перев. [Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.].

Вейвлет-преобразование (wavelet transform) является инструментом, разбивающим данные или функции или операторы на составляющие с разными частотами, каждая из которых затем изучается с разрешением, подходящим масштабу [62].

Вейвлет-преобразование можно разделить на *непрерывное вейвлет-преобразование* (НВП) и *дискретное вейвлет-преобразование* (ДВП) [35].

НВП (Continuous wavelet transform) – это преобразование, отображающее данную вещественнозначную функцию $x(t)$, определенную на временной оси переменной t , в функцию двух переменных τ и s :

$$\gamma(\tau, s) = \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t - \tau}{s} \right) dt,$$

$$\psi_{\tau, s}(x) = \frac{1}{\sqrt{|s|}} \psi \left(\frac{x - \tau}{s} \right),$$

где параметр τ представляет параллельный перенос; s – масштаб; $\psi(t)$ – материнский вейвлет [18].

Разложение в вейвлет-ряд ставит в соответствие функции непрерывного аргумента некоторую последовательность коэффициентов. В случае, когда подлежащая разложению функция является последовательностью чисел, таких как отсчеты непрерывной функции $f(x)$, получаемая последовательность коэффициентов называется ДВП (*Discrete Wavelet Transform*) функции $f(x)$ [19].

ДВП преимущественно применяется при решении задач кодирования сигналов, распознавания образов и компьютерной графики. НВП применяется для анализа структуры сложных сигналов.

Хотя преобразование Фурье является основой трансформационных методов обработки изображений еще с конца 1950-х годов, применение более современного преобразования, называемого вейвлет-преобразованием, упрощает сжатие, передачу и анализ изображений. В отличие от преобразования Фурье, базисными функциями которого являются

гармонические функции, вейвлет-преобразования основаны на разложениях по малым волнам, называемым вейвлетами, изменяющейся частоты и ограниченными во времени [19].

На сегодняшний день вейвлет-преобразование используется для анализа многих нестационарных процессов. Оно показало свою эффективность при решении широкого класса задач, связанных с обработкой изображений и распознаванием объектов на изображениях. Коэффициенты вейвлет-преобразования включают информацию об анализируемом процессе и используемом типе вейвлета. Каждый вейвлет имеет временные и частотные особенности.

Вейвлет-преобразование успешно используется для удаления шумов на изображениях. Изображение претерпевает вейвлет-преобразование, фильтрацию и обратное вейвлет-преобразование (рис. 1.4) [4].

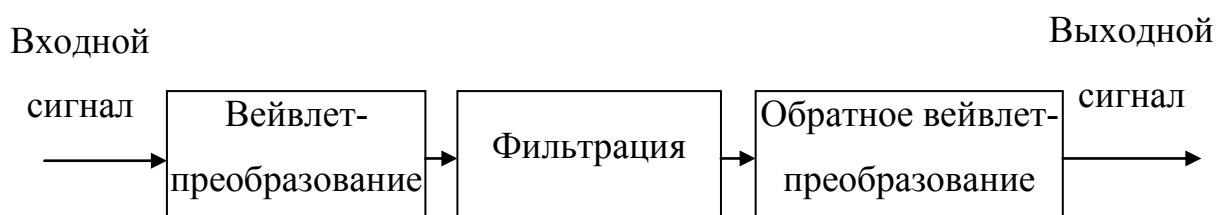


Рис. 1.4. Обобщенная схема подавления шума на основе вейвлет-преобразования [4]

Входным сигналом является яркостная составляющая изображения. Происходит преобразование изображения из формата RGB в цветовое пространство YUV.

Фильтрация применяется для очистки изображений от шума. При этом происходит удаление вейвлет-коэффициентов, соответствующих высокочастотным компонентам разложения. Затем осуществляется обратное вейвлет-преобразование.

После обратного вейвлет-преобразования изображение вновь переводится в формат RGB-кодирования. В результате получается изображение, очищенное от шумов.

а. Одномерное вейвлет-преобразование

Вейвлет-преобразование одномерного сигнала состоит в его разложении по базису, сконструированному из обладающей определенными свойствами солитоноподобной функции (вейвлета) посредством масштабных изменений и переносов. Каждая из функций этого базиса характеризует как определенную пространственную (временную) частоту, так и ее локализацию в физическом пространстве (времени) [13, 14].

б. Двумерное вейвлет-преобразование

Двумерное дискретное вейвлет-преобразование состоит из поочередного одномерного вейвлет-преобразования строк и столбцов этой матрицы. Сначала выполняют одномерные вейвлет-преобразования каждой строки в отдельности, при этом преобразованная строка записывается на прежнее место. Элементы нумеруются способом, указанным в предыдущих разделах. Далее вейвлет-преобразования применяются ко всем столбцам. В результате применения однократного вейвлет-преобразования изображение разбивается на четыре равные части (квадранты): LL, LH, HL и HH (рис. 1.5). Квадрант LL соответствует низкочастотным вейвлет-коэффициентам, а HH – высокочастотным (буква L означает Low, H – High) [47].

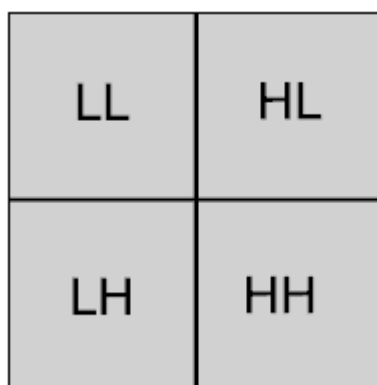


Рис. 1.5. Однократное применение двумерного вейвлет-преобразования к квадратному изображению

Если не оговорено противное, под N -кратным двумерным вейвлет-преобразованием понимается применение N раз двумерного вейвлет-преобразования, причём очередное двумерное вейвлет-преобразование

применяется к младшей четверти матрицы (квадрант LL на рис. 1.5). В итоге N -кратное преобразование выглядит так, как показано на рис. 1.6 (при $N=3$).

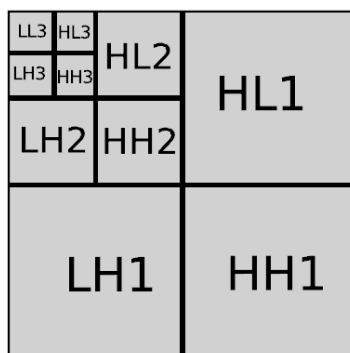


Рис. 1.6. Трёхкратное применение двумерного вейвлет-преобразования

На рис. 1.6 показаны принятые стандартные обозначения квадрантов изображения. Квадранты N -кратного двумерного вейвлет-преобразования имеют аналогичное обозначение.

Обратное двумерное вейвлет-преобразование рекурсивно восстанавливает младший квадрант. В случае, представленном на рис. 1.6, для получения (восстановления) нового квадранта LL2 используются квадранты LL3, LH3, HL3 и HH3. Далее для восстановления квадранта LL1 используются квадранты LL2, LH2, HL2, HH2 и т.д. Аналогично выполняется N -кратное обратное вейвлет-преобразование.

Заметим, что указанное преобразование является иерархическим, т.е. если при применении обратного вейвлет-преобразования вычисляются не все уровни, а меньшее их количество, то в квадранте LL образуется уменьшенная копия изображения, как показано на рис. 1.7. В частности, если вообще не используется обратное вейвлет-преобразование, то самый младший квадрант тоже является уменьшенной копией изображения.

Благодаря этому свойству, обратное вейвлет-преобразование позволяет вырезать фрагменты изображений при различных масштабах. Следует отметить, что, во-первых, доступные масштабы определяются количеством уровней вейвлет-преобразования и, во-вторых, масштабы не произвольны, а отличаются увеличением в два раза.

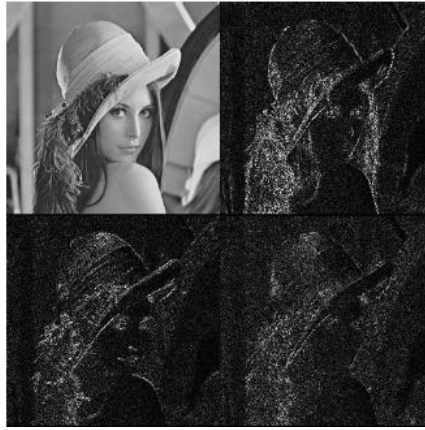


Рис. 1.7. Однократное применение двумерного вейвлет-преобразования

с. Вейвлет Хаара

Вейвлетом Хаара (Haar) являются кусочно-постоянные функции, заданные на нечетных интервалах различных масштабов и принимающие два значения $\{-1; +1\}$ [51]. Вейвлет Хаара единичного масштаба и нулевого смещения (материнский вейвлет Хаара) представляет собой функцию, равную $+1$ на интервале $[0; 1/2)$ и -1 на интервале $[1/2; 1)$. Вейвлет Хаара хорошо зарекомендовал себя в практических задачах обработки дискретных сигналов, таких как массивы отсчетов аудиосигналов и цифровые фотографии. Характерная особенность преобразования Хаара заключается в том, что оно является разделимым и легко вычисляется.

Вейвлет Хаара является ортогональным вейвлетом, определяющимся соотношением [3]:

$$\psi^H(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2} \leq t < 1, \\ 0, & t < 0, t \geq 1. \end{cases}$$

Любые две функции ψ_{jk}^H, ψ_{lm}^H , полученные из этого вейвлета по формуле:

$$\psi_{jk}(t) = 2^{\frac{j}{2}} \psi(2^j t - k), j, k \in I$$

с помощью масштабных преобразований $\frac{1}{2^j}, \frac{1}{2^l}$ и сдвигов $\frac{k}{2^j}, \frac{m}{2^l}$ ортогональны и имеют единичную норму.

Функция Хаара не обладает хорошей частотно-временной локализацией, ее преобразование Фурье $\hat{\psi}(\xi)$ убывает как $|\xi|^{-1}$ при $\xi \rightarrow \infty$.

Преобразование Хаара (ПХ) является простейшим базисным вейвлет-преобразованием. Пусть имеется одномерный дискретный сигнал $f(f_1, f_2, \dots, f_n)$. ПХ разлагает каждый сигнал на два компонента, один из которых называется средним или аппроксимацией (*approximation*), а другой известен как различие (*difference*) или деталь (*detail*) [13, 14].

Формула для вычисления среднего значения подсигнала (*subsignal*) выражается как $a^1 = (a_1, a_2, \dots, a_{N/2})$ на первом уровне для одного сигнала длины N , т.е. $f(f_1, f_2, \dots, f_n)$ имеет вид:

$$a_n = \frac{f_{2n-1} + f_{2n}}{\sqrt{2}}, \quad n = 1, 2, 3, \dots, N/2.$$

Соответствующее соотношение для детализирующей компоненты подсигнала $d^1 = (d_1, d_2, \dots, d_{N/2})$, на этом же уровне представляется в виде:

$$d_n = \frac{f_{2n-1} - f_{2n}}{\sqrt{2}}, \quad n = 1, 2, 3, \dots, N/2.$$

Эти значения формируют два новых сигнала: $a \{a_n\} n \in Z$ и $d \{d_n\} n \in Z$, один из которых является огрубленной версией исходного сигнала (каждой паре элементов f соответствует их среднее арифметическое), а другой содержит информацию (будем называть ее детализирующей), необходимую для восстановления исходного сигнала. Действительно:

$$f_{2n-1} = \frac{a_n + d_n}{\sqrt{2}},$$

$$f_{2n} = \frac{a_n - d_n}{\sqrt{2}}.$$

d. Вейвлеты Добеши

Для построения различных вейвлетных разложений, связанных с равномерной сеткой, исходным моментом является кратномасштабное уравнение. Для получения вариантов вейвлетных разложений задаются числами d_j и рассматривают кратномасштабное уравнение вида [24]:

$$\omega(t) = d_0\omega(2t) + d_1\omega(2t-1) + d_2\omega(2t-2) + \dots + d_N\omega(2t-N).$$

Масштабирующую функцию $\omega(t)$ иногда подчиняют тем или иным дополнительным условиям. Примерами таких условий являются условие нормировки и условие ортогональности:

$$\int_{-\infty}^{+\infty} |\omega(t)|^2 dt = 1, \quad \int_{-\infty}^{+\infty} \omega(2t-j)\omega(2t-k) dt = 0 \text{ при } j \neq k.$$

Отыскание функции ω из приведенных выше условий является непростым делом. Применение преобразования Фурье к тождеству является основным приемом решения этой задачи. Это приводит к представлению образа Фурье функции ω в виде бесконечного произведения, откуда можно, по крайней мере, извлечь ее значения в фиксированных точках вещественной оси (т.е. в точках, являющихся обыкновенными дробями со знаменателем 2^s , где s – целое число).

В частности, если взять:

$$d_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad d_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, \quad d_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad d_3 = \frac{1-\sqrt{3}}{4\sqrt{2}},$$

а остальные коэффициенты d_j положить равными нулю, то получается вейвлет, называемый *D4-вейвлетом Добеши (Daubechies)*.

Заметим, что этот вейвлет равен нулю вне промежутка $[-1, 2]$, он непрерывен, но не дифференцируем. В точках вида $\frac{k}{2^s}$ (где k и s – целые) у него имеются только левая производная, а правой производной нет.

е. Примеры применения вейвлет-преобразования для выделения признаков объектов на изображениях

В работах [9, 10] представлены разложение изображения и выделение его признаков для классификации изображений самолетов на основе применения вейвлет-преобразования Хаара и многослойной нейронной сети. В работе [11] используются вейвлет-преобразования Хаара и Добеши для выделения признаков изображения конкретных объектов в виде лиц и жестов. Примеры применения вейвлет-преобразования Хаара и Добеши для выделения признаков изображения лица представлены на рис. 1.8.

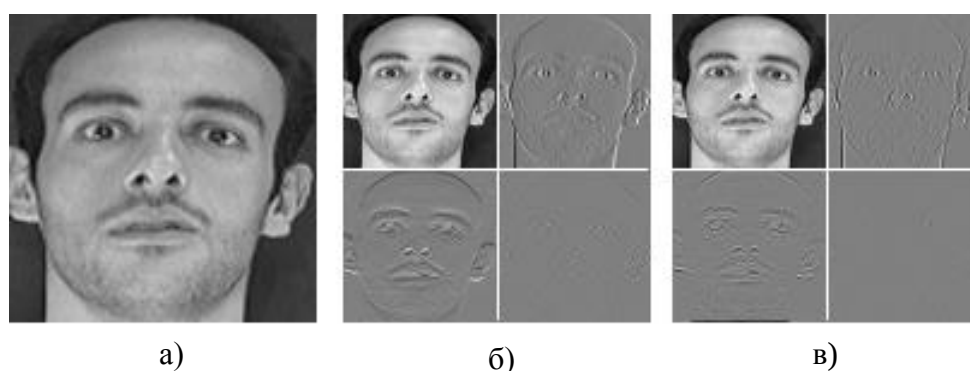


Рис. 1.8. Примеры выделения признаков лица: а) исходное изображение лица; б) результат применения вейвлет-преобразования Хаара; в) результат применения вейвлет-преобразования Добеши

1.3 Обнаружение лиц на изображениях

Выделение границ (выделение краёв) играет важную роль в анализе изображений и распознавании образов при решении проблемы компьютерного зрения. Края являются точками перепада значений яркости в полутоновом изображении и относятся к важным характеристикам изображения. Представление краев объектов на изображении позволяет уменьшить количество данных, сохраняющих важную информацию о форме объектов в сцене. Выделение краев является наиболее используемой технологией в процессе обработки цифрового изображения. При определении границ объекта производится вычисление двумерного пространственного градиента на изображении и определение областей,

соответствующих краям. При этом происходит оценка модуля градиента в каждой точке полутонового изображения [101].

Выделение краев заключается в выделении точек изображения, в которых резко изменяется яркость или есть другие виды неоднородностей. Результатом выделения является набор связанных кривых, обозначающих границы объектов, граней и оттисков на поверхности, а также кривые, которые отображают изменения положения поверхностей. В работе [12] анализируются методы выделения краёв на цифровых изображениях, такие как: Собеля, Лапласа и алгоритмы математической морфологии.

Корректное обнаружение и выделение лиц на изображениях играет важную роль при решении задач распознавания. Выделение лиц для сравнения фотографий на документах не представляет особого труда, поскольку при этом нет необходимости в детектировании. В аэропортах при детектировании лиц может использоваться движение объекта, полученное из видеопоследовательности.

В настоящее время для обнаружения лиц на видеоизображениях предлагается использовать информацию о цвете. При этом выделяется прямоугольная область, в которой содержится лицо. После этого, как правило, применяются методы, разработанные для анализа статических изображений [48].

Задача обнаружения объектов и методы ее решения

Цель обнаружения – определить наличие объекта на изображениях и найти его положение в системе координат пикселей исходного изображения. Положение объекта в зависимости от выбора алгоритма детектирования может определяться координатами прямоугольника, окаймляющего объект, либо контуром этого объекта, либо координатами точек, наиболее характерных для объекта [5].

Обнаружение лиц можно рассматривать как особый случай задачи распознавания, в котором существует всего два класса – «лицо» либо «не лицо». При решении этой задачи используются специфичные методы для

детектирования и выделения лиц. Результатом работы детектора является множество особых точек, выделенных на анализируемом изображении.

Среди алгоритмов для обнаружения лиц на изображениях, наиболее важными являются алгоритм обнаружения объекта Виолы-Джонса [116–118], статистический подход обнаружения Шнейдермана и Канады [99] и обнаружение лица на основе нейронной сети [102]. Эти алгоритмы имеют высокий уровень эффективности при обнаружении лиц, но только метод Виолы-Джонса оказался достаточно быстрым и точным для обнаружения объектов на статических изображениях и видеопоследовательности в режиме реального времени [93].

Метод Виолы-Джонса был разработан и представлен Полом Виолой (*Viola*) и Майклом Джонсом (*Jones*) в 2001 году [117, 118]. Метод хорошо обнаруживает объекты при наблюдении под небольшим углом (примерно до 30°), на его работу слабо влияют условия освещённости и он позволяет обнаруживать не только одно, но и несколько лиц на изображении.

Метод Виолы-Джонса работает следующим образом.

Шаг 1. Представление изображения в интегральном виде.

Шаг 2. Применение признаков Хаара для поиска объекта.

Шаг 3. Использование бустинга для выбора наиболее подходящих признаков для искомого объекта на данной части изображения.

Шаг 4. Все признаки поступают на вход классификатора, который определяет, является ли предъявленный объект «лицом».

Шаг 5. Использование каскада признаков для быстрого отбрасывания окон, в которых не найдено лицо.

Пример работы метода Виолы-Джонса для обнаружения множества лиц представлен на рис. 1.8.

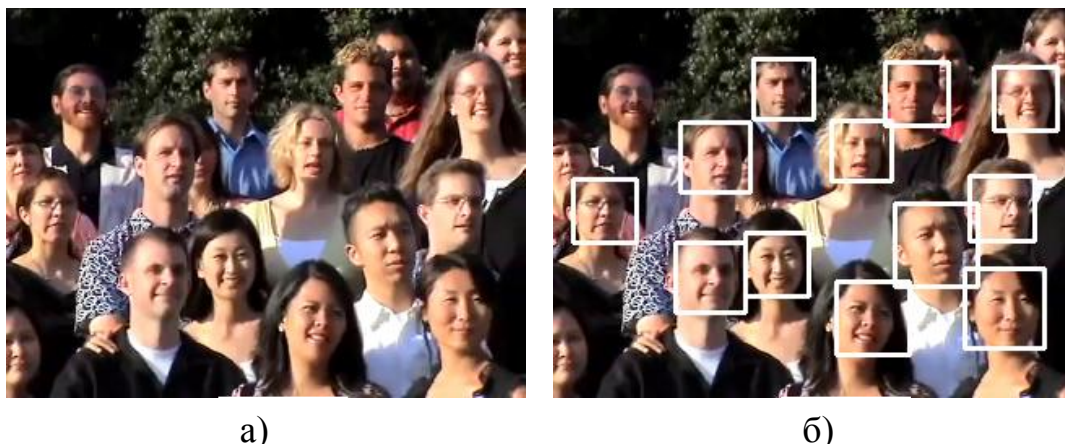


Рис. 1.8. Пример работы метода Виолы-Джонса: а) исходное изображение; б) обнаружение лиц

1.4 Слежение за объектами

Слежение за объектом на видеопоследовательности — это задача последовательного определения положения одного объекта или нескольких объектов, независимо от движения объекта или камеры [82].

Целью слежения за объектами является установление соответствия между объектами в видеопоследовательности, определение скорости движения объектов и их траекторий. Основной проблемой в слежении за объектами является сопоставление положения целевого объекта на кадрах видеопоследовательности.

Факторы, влияющие на сопровождение выделенных объектов в видеопоследовательностях [30]:

- изменение освещенности сцены, наличие шума камеры;
- изменяющиеся формы объектов;
- наличие движущихся объектов, имеющих близкие характерные признаки;
- неправильная сегментация объектов на предыдущих этапах обработки;
- необходимость осуществлять слежение в реальном режиме времени.

Поскольку слежение за объектами проводится в реальном режиме времени, то необходимо использовать алгоритмы или методы, которые не требуют больших вычислительных ресурсов. Выбор алгоритма и метода для слежения за объектами зависит от способа его представления.

Алгоритм *Mean Shift*

Алгоритм *Mean Shift* является эффективным подходом для слежения за объектами. Он позволяет быстро находить максимумы функции плотности распределения на динамических изображениях как для полутоновых, так и для цветных изображений [57–60]. При этом необходимо преобразовывать цвета каждого кадра в процессе слежения за цветными объектами на видеопоследовательности в распределение вероятности. Для достижения этой цели применяются гистограммы цветов. Распределение цвета, основанное на всей видеопоследовательности, меняется со временем [30].

Алгоритм *Mean Shift* используется для нахождения (*climbing*) градиента распределения вероятностей с целью определения ближайшего доминирующего пика.

Алгоритм *CAMShift*

Брадски (*Bradski*) в 1998 году создал алгоритм *CAMShift* (*Continuously Adaptive Mean-Shift*), который на основе цветовой информации был способен отслеживать объекты [54]. В нем осуществляется комбинация алгоритма отслеживания объекта *Mean Shift*, основанного на карте вероятности цвета кожи, с адаптивным шагом изменения размера области отслеживания. Вероятность цвета кожи каждого пикселя изображения определяется методом *Histogram Backprojection*, основанным на цвете, представленном в виде цветового тона (*Hue*) модели HSV (*Hue Saturation Value*). Так как алгоритм *CAMShift* способен отслеживать лица на основе вероятности цвета кожи, то он может применяться и для отслеживания руки.

Преимуществами данного алгоритма являются низкие требования к вычислительным ресурсам, гибкие настройки точности позиционирования и возможность работы в различных условиях освещённости. Кроме этого,

дополнительным преимуществом алгоритма является возможность работы в условиях частичного перекрытия отслеживаемого объекта. Указанные выше свойства алгоритма обусловлены использованием модели объекта, построенной на основе гистограммы яркости и цвета, а также использованием процедуры *Mean Shift* для точного позиционирования положения объекта.

1.5 Цель и задачи исследования

Учет значения каждого пикселя как координаты отдельной оси является простейшим вариантом для представления любого цифрового изображения в виде вектора в пространстве признаков. Поскольку цифровое изображение рассматривается как матрица пикселей $M \times N$, то размерность такого пространства будет очень большой. С учетом того, что все человеческие лица схожи между собой, все векторы, описывающие изображения лиц, будут размещаться в узко ограниченной области указанного пространства. Поэтому при решении задачи идентификации человека по фотопортрету описание и хранение всего векторного пространства избыточно. Первоочередной задачей является построение пространства меньшей размерности, в котором описываются только изображения человеческих лиц.

Обеспечить уменьшение времени обработки изображений можно выбором минимального количества признаков изображений объектов. При этом необходимо учитывать ограничения, которые возникают при уменьшении количества используемых признаков объектов для распознавания:

- 1) уменьшение статистической устойчивости результатов распознавания;
- 2) зависимость продолжительности работы по измерению значений признаков (затрат времени) от применяемой технологии распознавания.

Для надежного распознавания объектов на основе всей имеющейся на изображении информации, нужно выделить ее наиболее релевантную часть и представить ее в виде, позволяющем существенно упростить задачу сравнения текущего изображения объекта с изображениями объекта, имеющимися в БД.

В Массачусетском технологическом институте введен подход к извлечению признаков, основанный на применении вейвлет-анализа [49]. При этом используется вычислимое вейвлет-преобразование Хаара. Изображения отображаются из пространства пикселей в пространство признаков вейвлета Хаара, которое содержит достаточно полное (приемлемое) описание объектов.

Другим часто используемым подходом для извлечения признаков при обнаружении лиц является PCA. PCA позволяет спроектировать набор изображений из пространства высокой размерности пикселей в пространство более низкой размерности, содержащее набор изображений в качестве основных компонентов. Однако этот метод имеет ограничение, поскольку при его применении возникают проблемы с устранением шума.

В разделе 1.2.2. установлено, что среди методов снижения размерности данных наилучшим является PCA. Этот метод эффективно работает в больших базах изображений и он также подходит для высокоскоростного анализа видеоизображений, поступающих с веб-камеры. Таким образом, для извлечения признаков объекта целесообразно использовать вейвлет-преобразование, а затем применять PCA.

Для решения задач распознавания движущихся объектов на видеопоследовательности с высокой скоростью и точностью требуются алгоритмы детектирования и трекинга.

Процесс распознавания множества объектов на видеопоследовательности состоит из двух основных этапов:

- первый этап – извлечение и сохранение признаков известных объектов в БД;

- второй этап – поиск области движущихся объектов на видеопоследовательности и сравнение признаков найденных объектов с признаками, находящимися в БД.

В данной работе осуществляется разработка оригинальных алгоритмов для распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях в реальном режиме времени. Актуальность данной задачи обусловлена необходимостью создания новых высокоскоростных алгоритмов, способных осуществлять распознавание объектов на изображениях в присутствии шума и видеопоследовательностях в режиме реального времени с высокой точностью. Развитие указанного направления может найти применение при создании интеллектуальных систем безопасности, анализе биомедицинских изображений, в рамках организации и проведения видеоконференций и т.д.

Целью диссертационной работы является разработка алгоритма, основанного на эффективных способах выделения признаков объектов, позволяющего распознавать объекты в присутствии шума на статических изображениях и видеопоследовательностях в реальном режиме времени.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Осуществить анализ методов и алгоритмов распознавания лиц и жестов на изображениях и видеопоследовательностях и выбрать наиболее подходящие для решения данной задачи.

2. Исследовать способ выделения основных признаков объектов на изображениях, обеспечивающий высокую достоверность идентификации объекта поиска.

3. Разработать и реализовать алгоритмы и программные средства распознавания лиц и жестов на статических изображениях и видеопоследовательностях с использованием предложенного способа выделения признаков объектов.

4. Провести численные эксперименты по распознаванию разработанным программно-алгоритмическим комплексом лиц и жестов на изображениях и видеопоследовательностях в реальном режиме времени.

5. Сопоставить полученные результаты распознавания лиц и жестов на изображениях и видеопоследовательностях с результатами, полученными другими авторами.

1.6 Основные результаты и выводы по главе 1

Подводя итоги первой главы можно выделить следующие основные полученные результаты.

1. В данной главе проведен анализ методов и алгоритмов распознавания объектов на изображениях и видеопоследовательностях, включающий:

- метод главных компонент;
- линейный дискриминантный анализ;
- алгоритм *AdaBoost*;
- искусственные нейронные сети;
- метод опорных векторов.

2. Осуществлено рассмотрение возможности применения аппарата вейвлет-преобразований для обработки изображений и распознавания объектов. Показано, что дискретное вейвлет-преобразование целесообразно применять в процессе распознавания объектов на изображениях и видеопоследовательностях.

3. На основе проведенного анализа принято решение о целесообразности применения аппарата вейвлет-преобразований и РСА для извлечения основных признаков изображений при решении задачи распознавания лиц и жестов на статических изображениях в присутствии шума и видеопоследовательностях в реальном режиме времени. Указанный подход позволяет обеспечить приемлемую точность распознавания объектов и высокую скорость работы с видеопоследовательностями.

ГЛАВА 2. АЛГОРИТМЫ РАСПОЗНАВАНИЯ ЛИЦ И ЖЕСТОВ НА ОСНОВЕ ПРИМЕНЕНИЯ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ И МЕТОДА ГЛАВНЫХ КОМПОНЕНТ

В данной главе излагаются принципы решения задач распознавания статических и динамических объектов на изображениях и видеопоследовательностях. Приводится подробное описание оригинальных алгоритмов и способов, используемых для распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях в реальном режиме времени.

2.1 Способ выделения признаков объектов на изображениях

Представляется целесообразным для описания изображения применять вейвлет-преобразования, которые имеют свойства пространственно-частотной локализации и регулируемой детальности. В качестве типов вейвлетов были выбраны вейвлеты Хаара и Добеши.

В данной работе предложен способ выделения признаков объектов на изображениях, основанный на совместном применении вейвлет-преобразований Хаара и Добеши, позволяющий эффективнее распознавать объекты по сравнению с применением вейвлет-преобразования Хаара или Добеши в отдельности.

Процесс выделения признаков объектов на изображениях, основанный на совместном применении вейвлет-преобразований Хаара и Добеши, выполняется следующим образом (рис. 2.1).

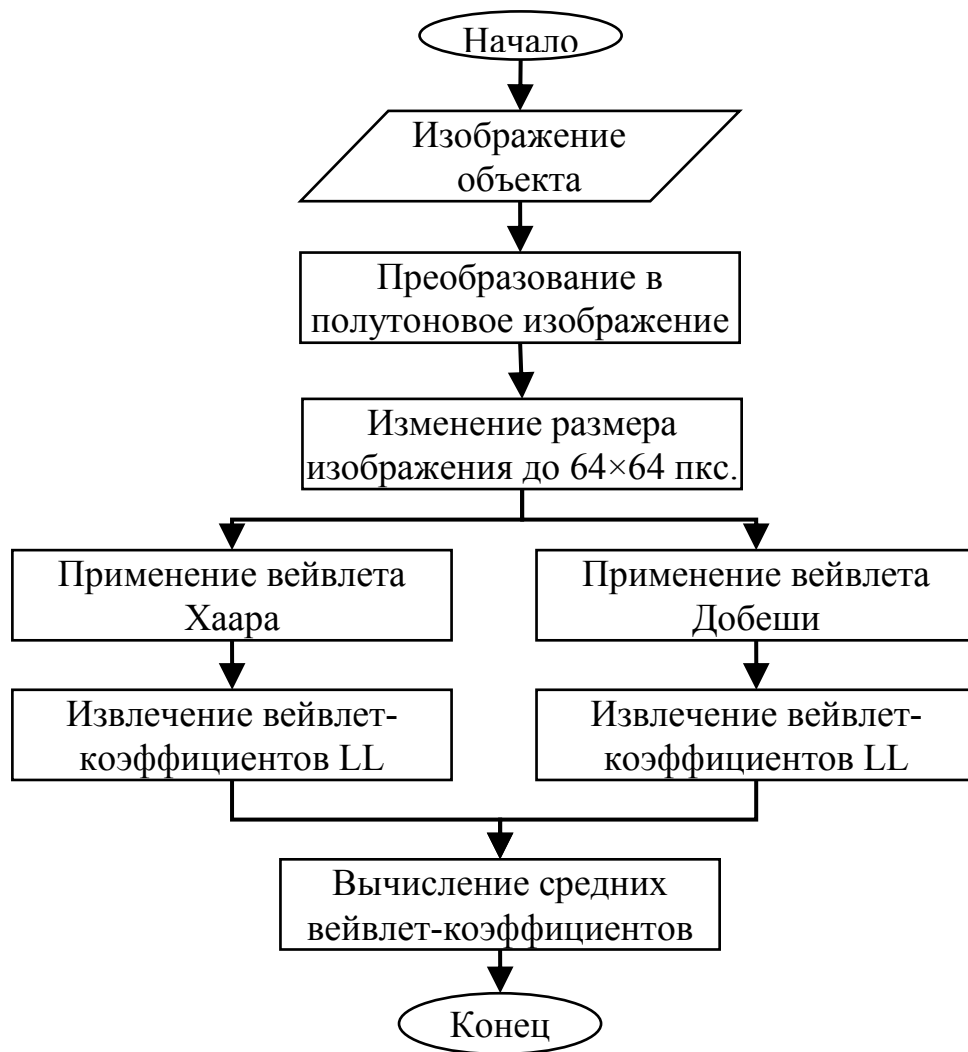


Рис. 2.1. Выделение признаков объектов на изображениях

Вначале происходит преобразование входного изображения из формата RGB в цветовое пространство YUV. Затем полутоновое изображение приводится к размеру 64×64 пикселя.

К полутоновому изображению применяются вейвлеты Хаара и Добеши первого уровня и извлекаются низкочастотные коэффициенты $X_{LL}(x, y)$ и $D_{LL}(x, y)$, где $x = 1, \dots, 32$; $y = 1, \dots, 32$ соответственно.

На рис. 2.2 представлены изображения низкочастотных коэффициентов и гистограммы их значений, которые получаются при применении вейвлетов Хаара и Добеши к полутоновому изображению.

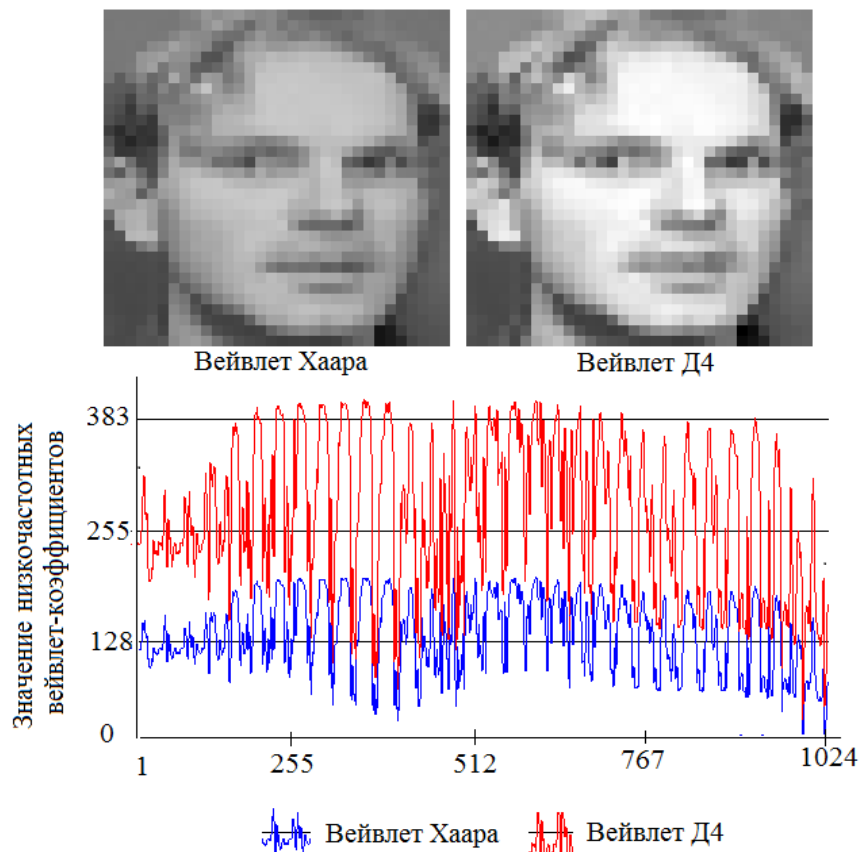


Рис. 2.2. Изображения низкочастотных вейвлет-коэффициентов и гистограммы их значений

Из рис. 2.2 следует, что средний уровень значений вейвлет-коэффициентов Хаара располагается ниже, чем средний уровень значений вейвлет-коэффициентов Добеши. Значения вейвлет-коэффициентов Хаара имеют меньший разброс по амплитуде по сравнению с соответствующими значениями для вейвлет-коэффициентов Добеши.

Для того чтобы уменьшить различие между значениями вейвлет-коэффициентов Добеши, было решено ограничивать эти значения следующим образом:

$$D_{LL}(x, y) = \begin{cases} D_{LL}(x, y), & \text{если } D_{LL}(x, y) < D \\ D, & \text{иначе} \end{cases}. \quad (2.1)$$

Здесь D – параметр для ограничения значений вейвлет-коэффициентов Добеши. Значение D определяется на основе эмпирических данных.

Для осуществления комплексного вейвлет-преобразования предлагается вычисление среднего арифметического вейвлет-коэффициентов Хаара и Добеши по формуле:

$$K_{cp}(x, y) = \frac{X_{LL}(x, y) + D_{LL}(x, y)}{2},$$

где $x = 1, \dots, 32$; $y = 1, \dots, 32$.

Характерная особенность преобразования Хаара заключается в том, что оно является разделимым и легко вычисляется. По сравнению с вейвлет-преобразованием Хаара, вейвлет-преобразование Добеши имеет более высокую вычислительную сложность. Вейвлеты Добеши являются функциями, вычисляемыми итерационным путем. В процессе выполнения преобразования происходит перекрытие между итерациями. Можно предположить, что именно за счет перекрытия применение вейвлетов Добеши позволяет учесть детали, которые пропущены вейвлетами Хаара.

2.2 Алгоритм распознавания лиц и жестов на статических изображениях

2.2.1 Алгоритм распознавания объектов при применении РСА

В результате проведенного анализа в главе 1 установлено, что РСА является одним из наиболее эффективных методов, который целесообразно применять для решения задач распознавания объектов.

Основная идея использования РСА для распознавания объектов состоит в том, чтобы отобразить большой одномерный вектор пикселей, построенный из двумерного изображения объекта, в компактные основные компоненты пространства признаков. Результат преобразования можно назвать *подпространством проекции*. Подпространство рассчитывается путем определения собственных векторов ковариационной матрицы, полученных из набора изображений объектов [30].

Алгоритм распознавания объектов при применении РСА имеет следующий вид [45].

Процесс создания собственных объектов

Шаг 1. Создать набор изображений $I_1, \dots, I_M$, отражающий вариации объектов. Представить их в виде векторов $\vec{I}_1, \dots, \vec{I}_M$.

Шаг 2. Вычислить среднее изображение по формуле:

$$\vec{I}_{cp} = \frac{1}{M} \sum_{i=1}^M \vec{I}_i.$$

Шаг 3. Вычесть среднее изображение из каждого изображения:

$$\vec{\Phi} = \vec{I}_i - \vec{I}_{cp}.$$

Шаг 4. Вычислить собственные объекты $\vec{u}_i$:

$$\vec{u}_i = \sum_{k=1}^M v_{ik} \vec{\Phi}_k, i = 1, \dots, M,$$

где $\vec{v}_i$ – собственные векторы матрицы $W^T W$, $W = \{\vec{\Phi}_1, \dots, \vec{\Phi}_M\}$.

Процесс распознавания объектов

Шаг 1. Поступление на вход нового изображения неизвестного объекта I_{ex} .

Шаг 2. Вычисление для каждого объекта в БД соответствующего ему вектора ω_i в пространстве собственных объектов V_p :

$$\omega_i = \vec{u}_i^T (\vec{I}_{ex} - \vec{I}_{cp}), i = 1, \dots, M.$$

Шаг 3. Вычисление проекции ω любого нового изображения I_{ex} в пространстве собственных объектов V_p .

Шаг 4. Если расстояние $d = |I_{ex} - \omega|$ больше некоторого предопределенного порога ε_1 , то классифицировать изображение как «иной объект».

Шаг 5. В случае если минимальное расстояние $d_k = |\omega - \omega_k|$ между проекцией нового изображения и известным представителем набора объектов меньше некоторого заданного порога ε_2 , то классифицировать изображение как «объект номер k ».

Шаг 6. В оставшемся случае ($d < \varepsilon_1$ и $d_k \geq \varepsilon_2$) классифицировать изображения как «неизвестный объект» и если это необходимо, то добавить новое изображение в БД и заново вычислить собственные объекты.

2.2.2 Принципы и схемы работы предложенного алгоритма распознавания объектов на изображениях

Для решения задачи распознавания лиц и жестов на статических изображениях предложен алгоритм, основанный на совместном применении вейвлет-преобразований Хаара, Добеши и РСА. В этом алгоритме лицо или жест рассматриваются как объект.

Процесс создания базы признаков объектов

Процесс создания базы признаков из M изображений объектов обучающей выборки, основанный на совместном применении вейвлет-преобразований Хаара, Добеши и РСА, состоит из следующих шагов (рис. 2.3).

1. Выделение признаков каждого i -ого изображения объекта из M изображений объектов обучающей выборки на основе совместного применения вейвлет-преобразований Хаара и Добеши.

2. Представление полученных вейвлет-коэффициентов в виде вектора $\vec{I}_i$, где $i = 1, \dots, M$.

3. Вычисление среднего изображения по формуле:

$$\vec{I}_{cp} = \frac{1}{M} \sum_{i=1}^M \vec{I}_i.$$

4. Вычитание среднего изображения из каждого изображения $\vec{\Phi}_i = \vec{I}_i - \vec{I}_{cp}$.

5. Вычисление собственных объектов:

$$\vec{u}_i = \sum_{k=1}^M v_{ik} \Phi_k, i = 1, \dots, M,$$

где $\vec{v}_i$ – собственные векторы матрицы $W^T W$, $W = \{\vec{\Phi}_1, \dots, \vec{\Phi}_M\}$.

6. Вычисление для каждого объекта соответствующего ему вектора $\vec{\Omega}_k^T = \{\omega_1, \dots, \omega_M\}$ в пространстве собственных объектов, где $\omega_i = \vec{u}_i^T (\vec{I}_i - \vec{I}_{cp})$, $i = 1, \dots, M$.

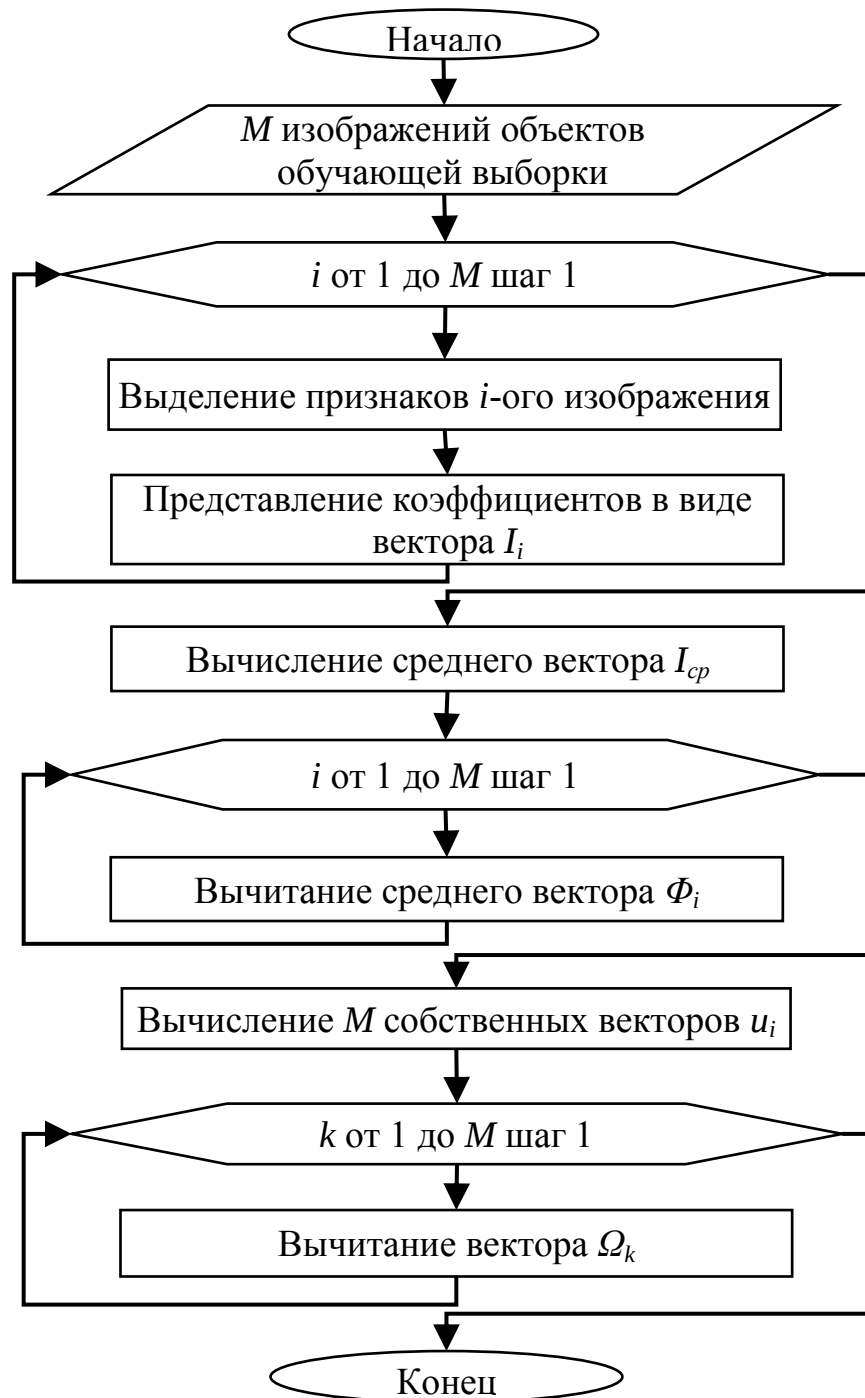


Рис. 2.3. Создание базы признаков объектов

Процесс распознавания объекта

После того, как создана база признаков, процесс распознавания нового входного объекта выполняется следующим образом (рис. 2.4).

1. Выделение признаков объекта входного изображения на основе совместного применения вейвлет-преобразований Хаара и Добеши.

2. Представление полученных вейвлет-коэффициентов в виде вектора

$\vec{I}_{вх}$.

3. Вычисление для входного объекта соответствующего ему вектора $\vec{\Omega}^T = \{\omega_1, \dots, \omega_M\}$ в пространстве собственных объектов, где $\omega_i = \vec{u}_i^T (\vec{I}_{ex} - \vec{I}_{cp})$, $i = 1, \dots, M$.

4. Сопоставление входного объекта с k -ым известным объектом из базы признаков на основе вычисления расстояния Евклида:

$$d_k^2 = \|\vec{\Omega} - \vec{\Omega}_k\|^2. \quad (2.2)$$

5. Определение принадлежности входного объекта к известному объекту осуществляется путем сравнения величины d_k со значением порога δ . Если величина расстояния Евклида d_k , определяемого по формуле (2.2), меньше значения порога δ для некоторого k , то такое изображение объекта принадлежит k -му классу. Иначе оно считается новым.

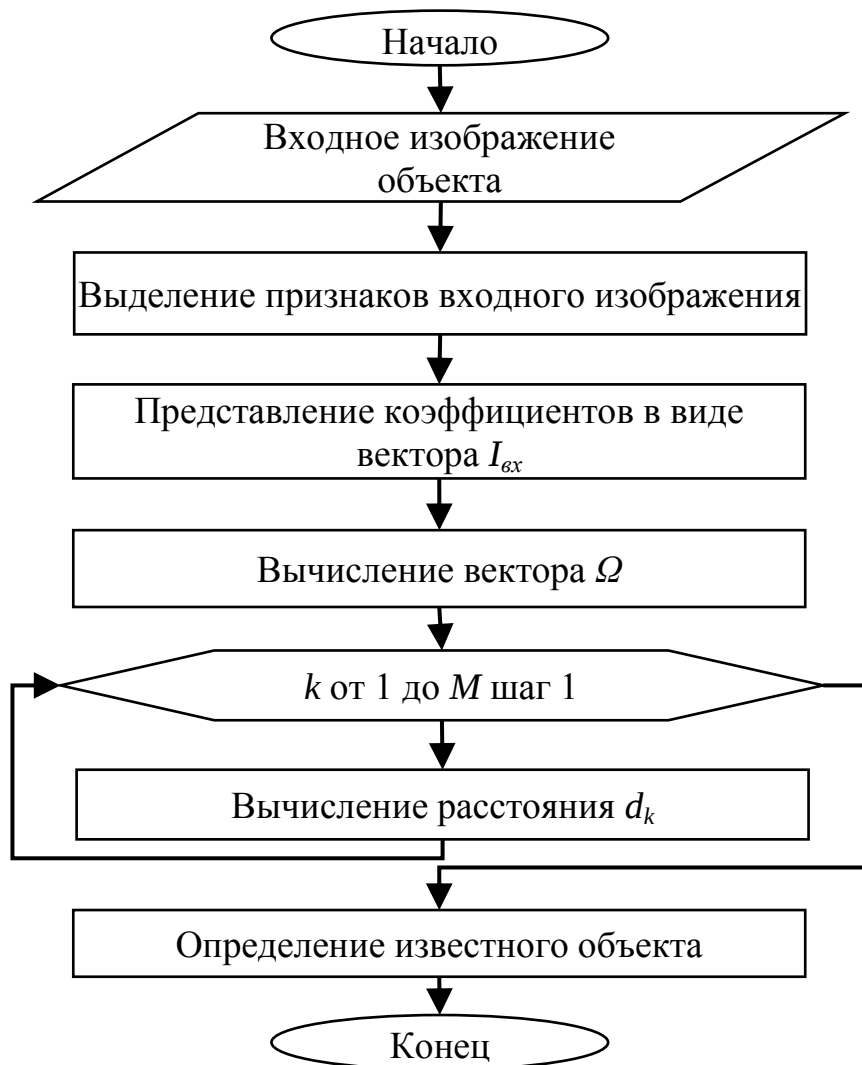


Рис. 2.4. Распознавание объекта

2.3 Экспериментальное оценивание точности распознавания объектов на основе применения вейвлет-преобразований и PCA

В этом разделе приводятся результаты трех экспериментов со 100 изображениями 10 человек (по 10 изображений каждого лица). Изображения, используемые для эксперимента, были выбраны из базы изображений лиц *The ORL Database of Faces* [95].

Первый эксперимент используется для исследования зависимости точности распознавания от количества изображений обучающей выборки и уровня разложения вейвлет-преобразований. Второй эксперимент выполняется для настройки значения параметра D (вейвлет-преобразование Добеши) и в нем исследуется влияние значения порога признания лица (метод главных компонент) на точность распознавания. Третий эксперимент проводится для исследования зависимости эффективности работы вейвлет-преобразований от их типа при решении задачи распознавания.

Распознавание лиц состоит из двух процессов.

- Первый процесс – выделение признаков изображений лиц обучающей выборки и сохранение их в базе признаков.
- Второй процесс – выделение признаков изображений лиц тестовой выборки и сравнение полученных признаков с признаками, хранящимися в базе признаков.

Для оценки эффективности процесса распознавания применяются вероятности ошибки первого (*False Acceptance Rate*, FAR) и второго рода (*False Reject Rate*, FRR). Уровнем равной ошибки (*Equal Error Rate*, EER) называется точка, в которой ошибка первого рода равна ошибке второго рода.

Исследование зависимости точности распознавания лиц от уровня разложения вейвлет-преобразований и количества изображений обучающей выборки

В этом эксперименте используются комбинация вейвлет-преобразования Хаара и PCA (Хаар+PCA) и комбинация вейвлет-преобразования Добеши и PCA (Д4+PCA) для выделения признаков изображений лиц.

Для исследования зависимости точности распознавания от количества изображений обучающей выборки были созданы разные обучающие выборки, которые содержат 10, 20, 30, 40 и 50 изображений (1, 2, 3, 4 и 5 изображений каждого лица). При этом изображения каждого лица обучающей выборки случайно выбираются из 10 изображений такого же лица.

После того, как случайно выбраны 10, 20, 30, 40 и 50 изображений для обучения из 100 изображений, остальные изображения используются для тестирования. Т.е. тестовые выборки соответственно содержат 90, 80, 70, 60 и 50 оставшихся изображений.

Процесс создания обучающей выборки для каждого количества изображений проводится 100 раз, т.е. имеются 100 обучающих выборок с 10 изображениями. И, аналогично, для каждого количества из 20, 30, 40 и 50 изображений создаются 100 обучающих выборок.

На рис. 2.5 представлена обучающая выборка, состоящая из 50 изображений, случайно выбранных из 100 изображений. На рис. 2.6 представлена тестовая выборка, состоящая из 50 остальных изображений.



Рис. 2.5. Пример обучающей выборки с 50 изображениями, случайно выбранными из 100 изображений.



Рис. 2.6. Пример тестовой выборки с 50 остальными изображениями

Для исследования зависимости точности распознавания от уровня разложения вейвлет-преобразований применяются вейвлет-преобразования Хаара и Добеши первого, второго, третьего и четвертого уровней.

Зависимости точности распознавания лиц на статических изображениях от количества изображений обучающей выборки и уровня

разложения вейвлет-преобразований представлены на рис. 2.7 (вейвлет-преобразование Хаара) и рис. 2.8 (вейвлет-преобразование Добеши).

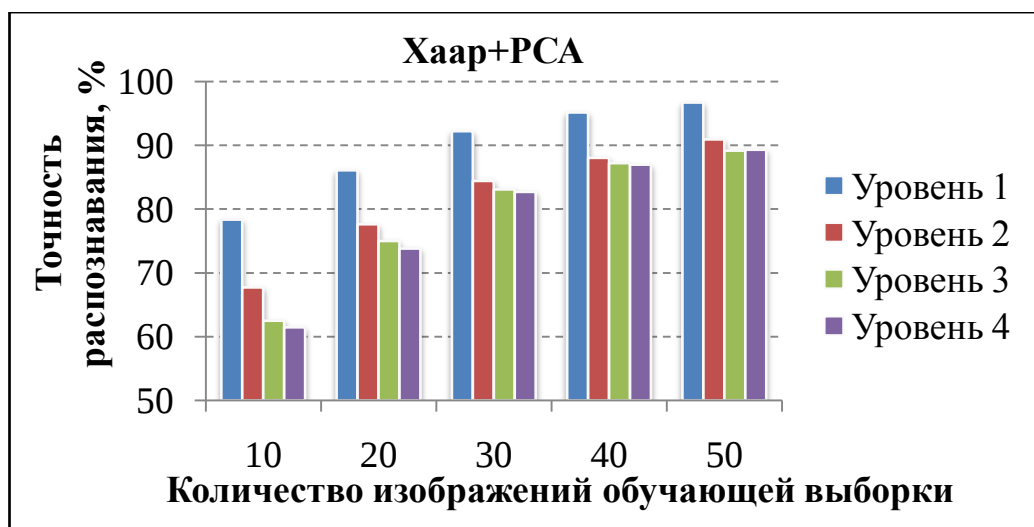


Рис. 2.7. Средние точности распознавания лиц в зависимости от уровня разложения вейвлет-преобразования Хаара

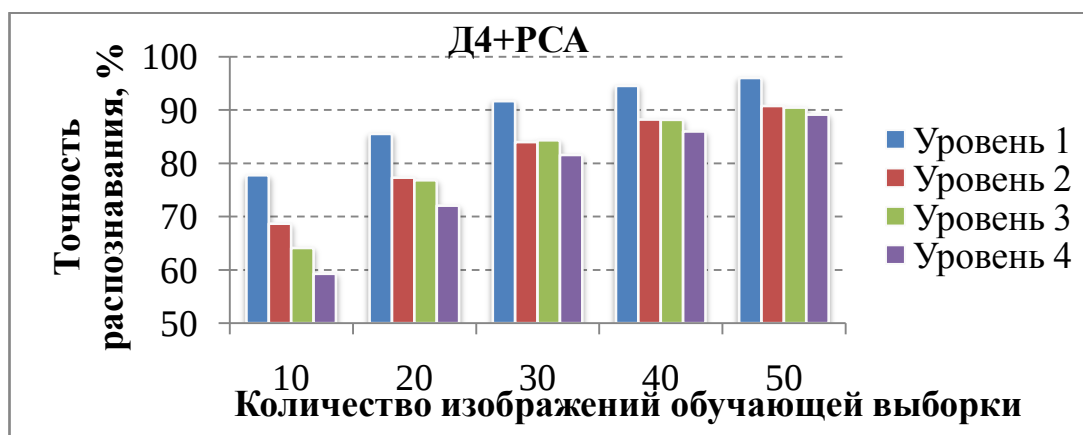


Рис. 2.8. Средние точности распознавания лиц в зависимости от уровня разложения вейвлет-преобразования Добеши

По результатам экспериментов можно сделать следующие выводы:

- При однократном применении двумерного вейвлет-преобразования Хаара и Добеши точность распознавания лиц на статических изображениях является наилучшей.
- Количество изображений обучающей выборки существенно влияет на точность распознавания. Чем больше количество изображений для обучения, тем лучше результат распознавания.

Настройка значения параметра D и исследование влияния значения порога признания лица δ на точность распознавания

Для настройки значения параметра D при ограничении значений вейвлет-коэффициентов в соответствии с (2.1) было проведено исследование его влияния на точность распознавания.

Рассматривались следующие комбинации:

- вейвлет-преобразование Хаара и PCA (Хаар+PCA);
- вейвлет-преобразование Добеши и PCA (Д4+PCA);
- вейвлет-преобразования Хаара, Добеши и PCA (Хаар+Д4+PCA), без введения ограничения на величину значений вейвлет-коэффициентов Добеши;

- вейвлет-преобразования Хаара, Добеши и PCA (Хаар+Д4(200)+PCA), (Хаар+Д4(250)+PCA), (Хаар+Д4(300)+PCA) и (Хаар+Д4(350)+PCA), при этом значение параметра D в формуле (2.1) принималось равным 200, 250, 300 и 350 соответственно.

Для обучения используются 100 обучающих выборок, содержащих 50 случайных изображений лиц (из 100). Для тестирования используются 100 тестовых выборок, содержащих 50 оставшихся из 100 изображений базы ORL. Затем к изображениям подготовленных тестовых выборок из базы ORL были добавлены 165 изображений базы *Yale Face Database* [122] для вычисления ошибки первого рода.

Зависимость ошибки первого и второго рода от значений порога признания лица δ при распознавании лиц приведена на рис. 2.9.

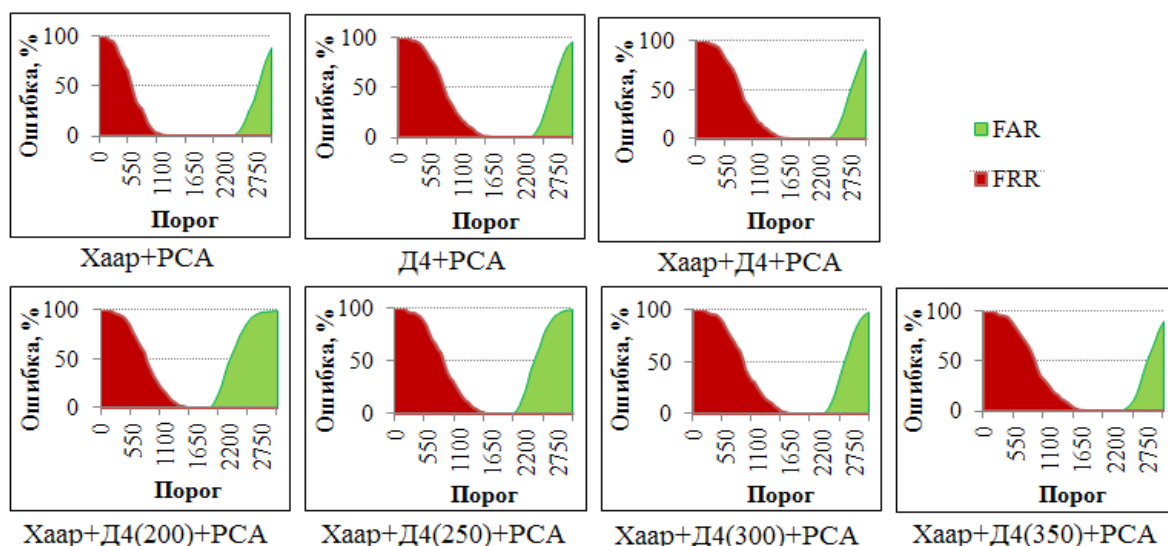


Рис. 2.9. Ошибки первого и второго рода при различных значениях порога признания лица

Из приведенного рисунка 2.9 следует, что уровень равной ошибки у каждой комбинации является различным. Значение порога признания лица для каждой комбинации находится в следующих диапазонах: 1550–1700 для комбинации Хаар+Д4(200)+РСА и 1700–2200 для комбинации Хаар+Д4+РСА. Для сравнения результатов распознавания лиц всеми комбинациями было выбрано общее значение порога признания лица равное 1700.

Результаты исследования зависимости точности распознавания лиц от значения параметра D , при использовании общего значения порога ($\delta=1700$) для комбинации Хаара, Добеши и РСА приведены на рис. 2.10.

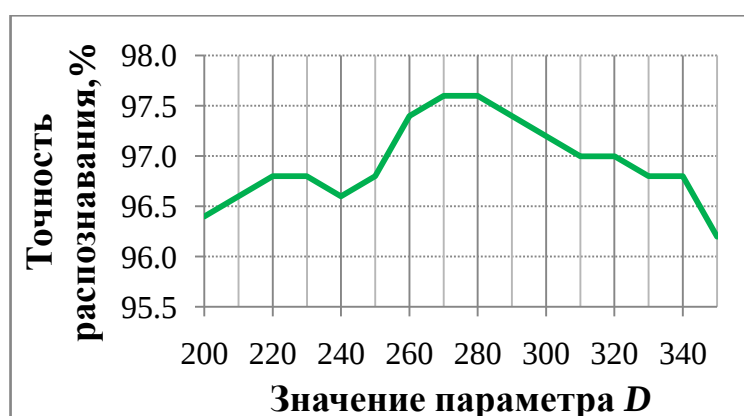


Рис. 2.10. Зависимость точности распознавания лиц при использовании общего значения порога ($\delta=1700$) от значения параметра D для комбинации Хаара, Добеши и РСА

Полученные результаты (рис. 2.10) показывают, что точность распознавания лиц изменяется с ростом значений параметра D . Наилучшая точность распознавания лиц составляет 97,6% при значении параметра D равном 280.

На рис. 2.11 представлены результаты распознавания лиц различными комбинациями при использовании общего значения порога ($\delta=1700$).

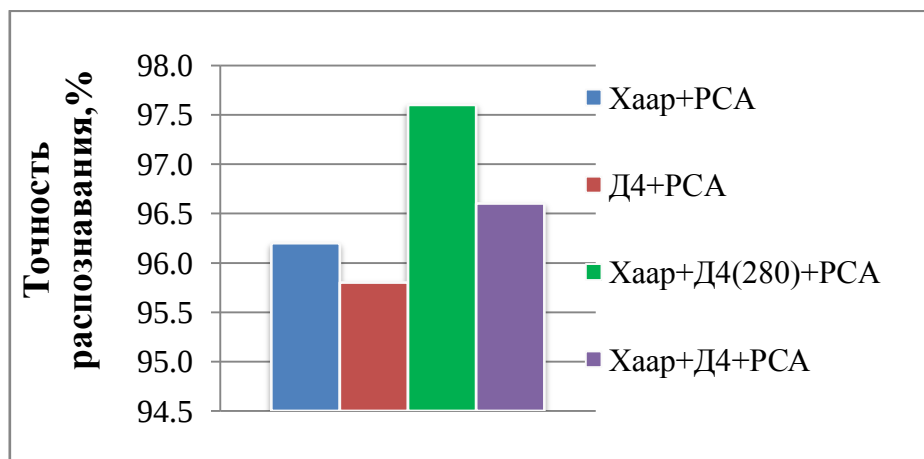


Рис. 2.11. Результаты распознавания лиц различными комбинациями при использовании общего значения порога признания лица

Показано, что наилучшая точность распознавания лиц комбинацией вейвлет-преобразований Хаара, Добеши и РСА без введения ограничения на величину значений вейвлет-коэффициентов Добеши составляет 96,6%. Введение ограничений значений вейвлет-коэффициентов Добеши (характеризуемых параметром D) позволило показать более высокую точность распознавания лиц (на 1,0%) на тестовых изображениях.

Таким образом, в дальнейшем при проведении численных экспериментов значение параметра D для комбинации вейвлет-преобразований Хаара, Добеши и РСА устанавливается равным 280, соответственно пороговое значение признания лица $\delta = 1700$.

Исследование эффективности работы комбинаций вейвлет-преобразований Хаара, Добеши и РСА для распознавания лиц в присутствии шума на изображениях

В данном эксперименте для выделения признаков изображений лиц применялись следующие комбинации: вейвлет-преобразования Хаара и РСА,

вейвлет-преобразования Добеши и PCA и вейвлет-преобразований Хаара, Добеши и PCA. В соответствии с приведенными выше результатами пороговое значение признания лица устанавливается равным 1700, а значение параметра D – 280.

Применяются вейвлет-преобразования Хаара и Добеши только первого уровня.

Используются те же, как и в предыдущем эксперименте, обучающие и тестовые выборки из базы ORL. Кроме этого, были созданы дополнительные тестовые выборки путем добавления 5 и 10% шума типа «соль и перец» к изображениям подготовленных тестовых выборок.

На рис. 2.12 и 2.13 представлены примеры изображений двух новых тестовых выборок с добавлением 5 и 10% шума типа «соль и перец» к изображениям тестовой выборки, приведенным на рис 2.6.



Рис. 2.12. Пример изображений с добавлением 5% шума типа «соль и перец» к изображениям тестовой выборки



Рис. 2.13. Пример изображений с добавлением 10% шума типа «соль и перец» к изображениям тестовой выборки

Результаты численных экспериментов по распознаванию лиц различными комбинациями представлены на рис. 2.14.

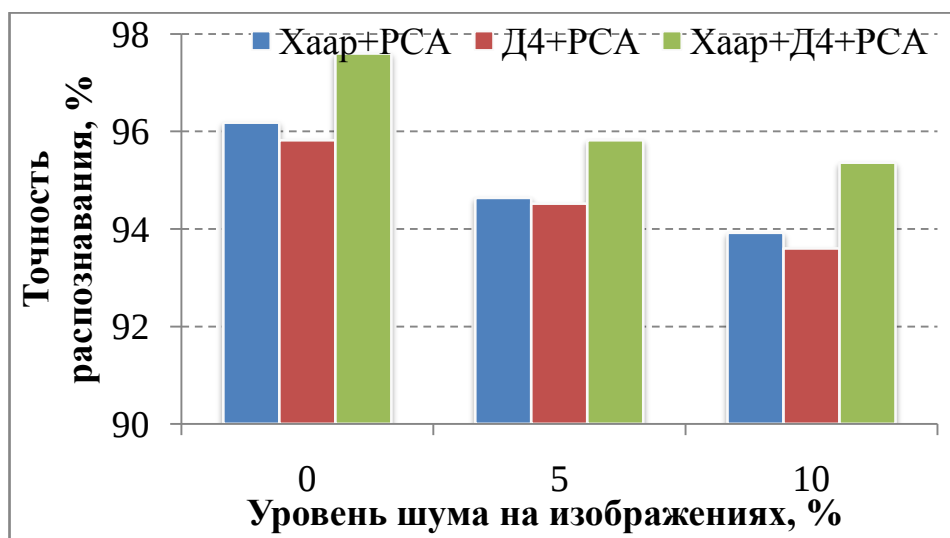


Рис. 2.14. Результаты распознавания лиц различными комбинациями

Показано, что комбинация вейвлет-преобразований Хаара, Добеши и РСА дает наилучшую точность распознавания лиц как в присутствии шума, так и в его отсутствии на изображениях тестовой выборки.

Как и следовало ожидать, точность распознавания лиц на статических изображениях при отсутствии шумов выше, чем в присутствии шума.

2.4 Алгоритм распознавания лиц на видеопоследовательности

В данной работе рассматривается задача высокоскоростного распознавания множества лиц на видеопоследовательности. Для решения этой задачи предложен алгоритм, основанный на совместном применении предложенного алгоритма распознавания лиц на статических изображениях и метода Виолы-Джонса.

Процесс создания базы признаков лиц на видеопоследовательности

Пусть имеется обучающая выборка, состоящая из M кадров, на каждом из которых расположено одно лицо. Процесс создания базы признаков происходит следующим образом (рис. 2.15).

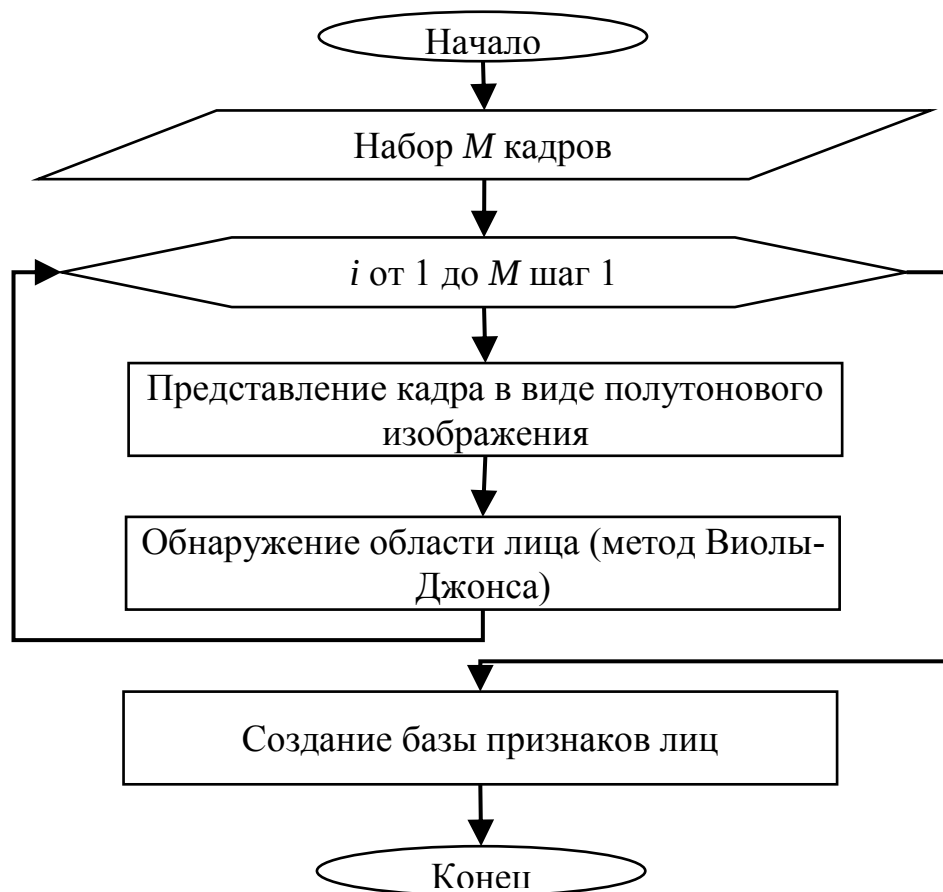


Рис. 2.15. Создание базы признаков лиц

1. Представление каждого из M кадров обучающей выборки в виде полутонового изображения.

2. Применение к каждому полутоновому изображению метода Виолы-Джонса для поиска области лица.

3. Выполнение процесса создания базы признаков из M обнаруженных областей лиц.

Процесс распознавания множества лиц на видеопоследовательности

Для распознавания множества лиц на видеопоследовательности к каждому очередному кадру применяются следующие шаги (рис. 2.16).

1. Представление очередного кадра в виде полутонового изображения.
2. Применение к полученному полутоновому изображению метода Виолы-Джонса для поиска N областей, содержащих лица.
3. Выполнение процесса распознавания каждого k -ого из N обнаруженных лиц.
4. Вывод результата распознавания на экран.

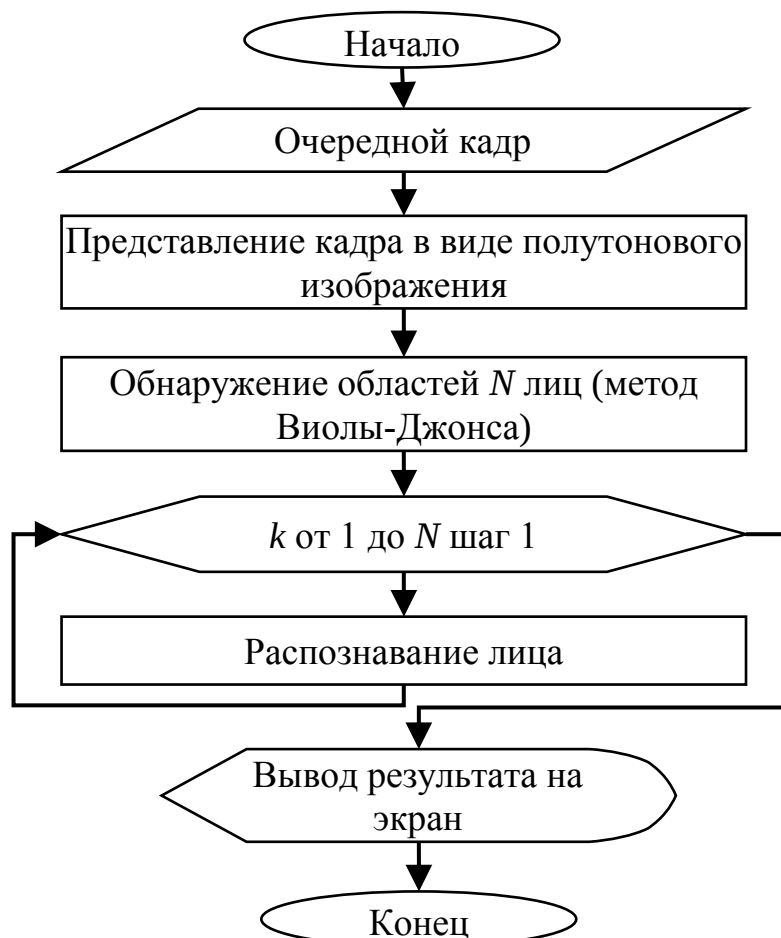


Рис. 2.16. Распознавание множества лиц

Для детектирования объектов на видеопоследовательности предложено применение метода Виолы-Джонса [117, 118]. Метод Виолы-Джонса является самым популярным методом для поиска области лица на изображении, из-за его высокой скорости и эффективности. Детектор лица Виолы-Джонса основан на трех основных идеях: интегральном представлении изображения, методе построения классификатора на основе алгоритма адаптивного бустинга (*AdaBoost*) и методе комбинирования классификаторов в каскадную структуру. Эти идеи позволяют построить надежный детектор лица, способный работать в режиме реального времени.

2.4.1 Интегральное представление изображений

Для того чтобы рассчитывать яркость прямоугольного участка изображения, применяется *интегральное представление* [19]. Интегральное представление (*integral image*) позволяет быстро рассчитывать суммарную яркость произвольного прямоугольника на данном изображении, причем время расчета не зависит от площади прямоугольника.

Интегральное представление изображения представляет собой матрицу, совпадающую по размерам с исходным изображением. В каждом её элементе хранится сумма интенсивностей всех пикселей, находящихся левее и выше данного элемента. Элементы матрицы рассчитываются по следующей формуле:

$$I(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'),$$

где $I(x, y)$ – значение точки (x, y) интегрального изображения; $i(x, y)$ – значение интенсивности исходного изображения.

На основе применения интегрального представления изображения вычисление признаков одинакового вида, но с разными геометрическими параметрами, происходит за одинаковое время.

Каждый элемент матрицы $I(x, y)$ представляет собой сумму пикселей в прямоугольнике от $i(0, 0)$ до $i(x, y)$, т.е. значение каждого элемента $I(x, y)$ равно сумме значений всех пикселей левее и выше данного пикселя $i(x, y)$. Расчет

матрицы занимает линейное время, пропорциональное числу пикселей в изображении и его можно производить по следующей формуле:

$$I(x, y) = i(x, y) - I(x-1, y-1) + I(x, y-1) + I(x-1, y).$$

Пример преобразования исходного изображения в интегральное представление имеет вид:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow{\text{Исходное изображение в интегральное представление}} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}.$$

2.4.2 Хаар-подобные признаки

Простые используемые признаки, напоминающие базисные функции Хаара, были использованы Папагеоргиу и другими при обнаружении объектов [96]. В частности, ими использовались три вида признаков (рис. 2.17).

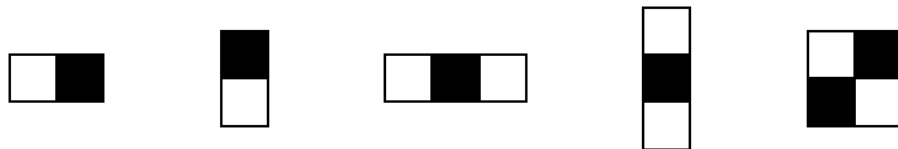


Рис. 2.17. Пример признаков прямоугольника, показанных относительно окрестности окна обнаружения.

Значением признака из двух прямоугольников является разница между суммой пикселей в этих прямоугольных областях. Области имеют одинаковый размер и форму и по горизонтали и по вертикали. Сумма пикселей, которые лежат в белых прямоугольниках вычитается из суммы пикселей в черных прямоугольниках. Признак из трёх прямоугольников вычисляет сумму в двух внешних прямоугольниках и вычитает её из суммы в центральном прямоугольнике. Признак из четырёх прямоугольников вычисляет разницу между диагональными парами прямоугольников.

Учитывая, что основное разрешение детектора 24×24 пикселей, исчерпывающий набор признаков прямоугольников довольно большой и

составляет 160000. В отличие от базиса Хаара, множество признаков прямоугольников является переполненным².

С точки зрения необходимости использования достаточно простых алгоритмов получения признаков, перспективным является использование *Хаар-подобных признаков (Haar wavelet-like features)*, представляющих собой результат сравнения яркостей в двух прямоугольных областях изображения.

В стандартном методе Виолы-Джонса используются прямоугольные признаки (рис. 2.18). Эти признаки называются примитивами Хаара. А в расширенном методе Виолы-Джонса, представленном в библиотеке *OpenCV*, используются дополнительные признаки (рис. 2.19).

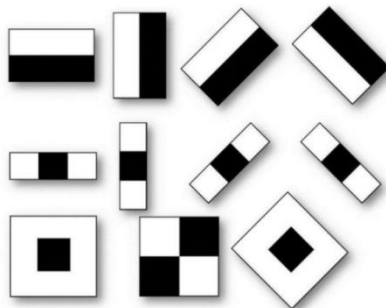


Рис. 2.18. Примитивы признаков Хаара

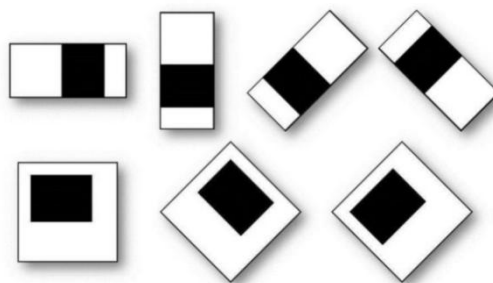


Рис. 2.19. Дополнительные признаки Хаара

Хаар-подобный признак вычисляется по формуле:

$$F_{Haar} = R_{white} - R_{Black} ,$$

где R_{white} – сумма значений яркостей точек, закрываемых светлой частью признака; R_{black} – сумма значений яркостей точек, закрываемых темной частью признака. Для их вычисления используется интегральное

²Следует отметить, что количество базисных элементов определяется размером изображения и составляет в данном случае $24 \times 24 = 576$ пикселей. Поэтому полный набор из 160000 признаков является переполненным.

представление изображения. Хаар-подобные признаки описывают значение перепада яркости по оси X и Y изображения соответственно.

На рис. 2.20 представлена схема получения Хаар-подобных признаков.

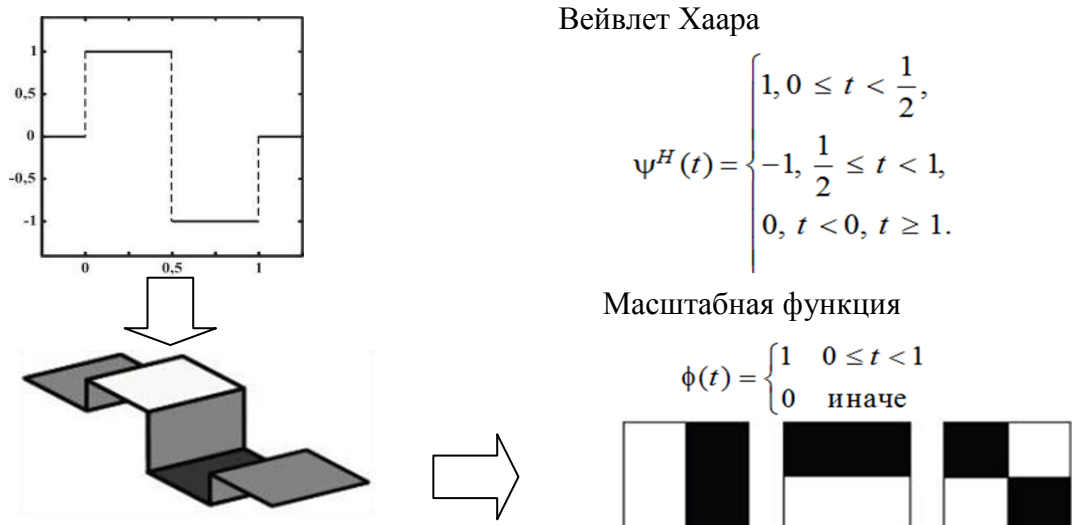


Рис. 2.20. Схема получения Хаар-подобных признаков

Метод Виолы-Джонса ищет лица и черты лица по принципу сканирующего окна (рис. 2.21).

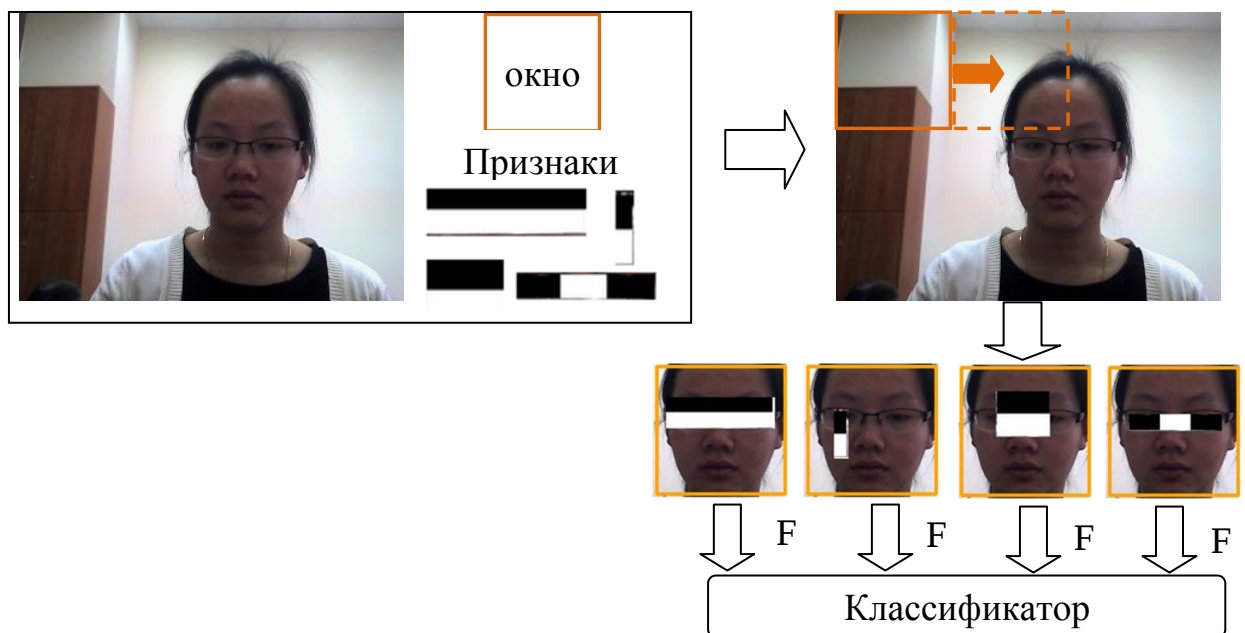


Рис. 2.21. Принцип поиска лиц методом Виолы-Джонса

Шаг 1. Сканирующее окно движется по изображению с шагом в одну ячейку окна (размер самого окна составляет 24×24 ячейки).

Шаг 2. Вычисляются признаки в каждом окне с различными формами, расположениями и размерами.

Шаг 3. Сканирование производится с различными масштабами окна.

Шаг 4. Все найденные признаки поступают на вход классификатора, который определяет, относится ли предъявляемый объект к классу «лиц».

2.4.3 Метод построения классификатора на основе алгоритма бустинга

Бустинг (boosting) – это комплекс методов, которые способствуют повышению точности аналитических моделей. Бустинг основан на усилении «слабых» моделей [105]. Процедура усиления построена в виде последовательной композиции алгоритмов машинного обучения, в которой каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов. При этом, ищется ответ на вопрос, каким образом, имея множество плохих (незначительно отличающихся от случайных) алгоритмов обучения, получить один хороший [65].

В результате работы алгоритма бустинга на каждой итерации формируется простой классификатор вида:

$$h_j(z) = \begin{cases} 1, & \text{если } p_j f_j(z) < p_j \theta_j \\ 0, & \text{иначе} \end{cases},$$

где p_j – показывает направление знака неравенства; θ_j – значение порога; $f_j(z)$ – вычисленное значение признака; z – окно изображения размером 20×20 пикселей.

В методе Виолы-Джонса, *AdaBoost* используется как для выбора признаков, так и для обучения классификатора [66]. Поскольку, как отмечено выше, количество признаков прямоугольников большое, то вычислить полное множество признаков чрезмерно трудоемко. Оказывается, что небольшое количество этих признаков может быть объединено для формирования эффективного классификатора. Нахождение этих признаков является главной задачей.

Для реализации этой цели разработан слабый алгоритм обучения, предназначенный для выбора единственного прямоугольного признака, который лучше всего отделяет положительные и отрицательные примеры. Для каждого признака, слабый классификатор определяет оптимальную пороговую функцию классификации, обеспечивающую минимальное число неправильно классифицированных примеров. Слабый классификатор $h(x)$ таким образом, состоит из значений признака объекта f_j , порога и паритета p , с указанием направления знака неравенства.

2.4.4 Метод комбинирования классификаторов в каскадную структуру

Каскадная структура позволяет повышать скорость обнаружения, фокусируя свою работу на наиболее информативных областях изображения.

Структура каскадного детектора приведена на рис. 2.22. Каскад состоит из слоев, которые представляют собой классификаторы, обученные с помощью процедуры бустинга [18].

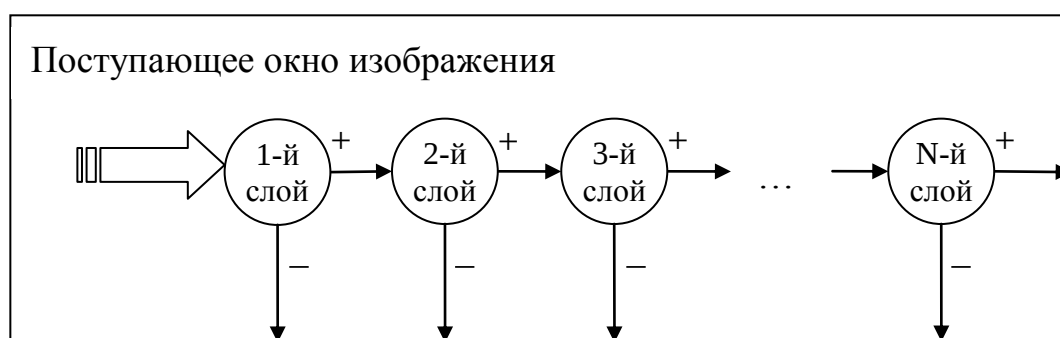


Рис. 2.22. Структура каскадного детектора

Общие принципы работы каскадного классификатора на изображении приведены на рис. 2.23.

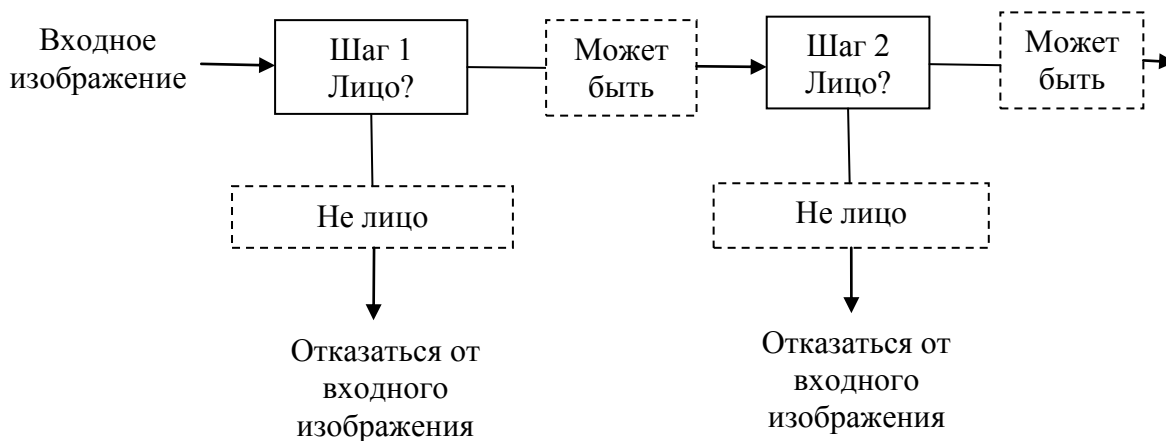


Рис. 2.23. Общие принципы работы каскадного классификатора

2.5 Алгоритм распознавания жестов на видеопоследовательности

Для решения задачи распознавания жестов на видеопоследовательности в режиме реального времени предложен оригинальный алгоритм, основанный на применении вейвлет-преобразований, PCA, метода Виолы-Джонса и алгоритма *CAMShift*.

Процесс создания базы признаков жестов на видеопоследовательности

Пусть имеется обучающая выборка, состоящая из M кадров, на каждом из которых расположен один жест. Процесс создания базы признаков происходит следующим образом (рис. 2.24).

1. Представление каждого из M кадров обучающей выборки в виде полутонового изображения.
2. Применение к каждому полутоновому изображению метода Виолы-Джонса для поиска области жеста.
3. Выполнение процесса создания базы признаков из M обнаруженных областей жестов.

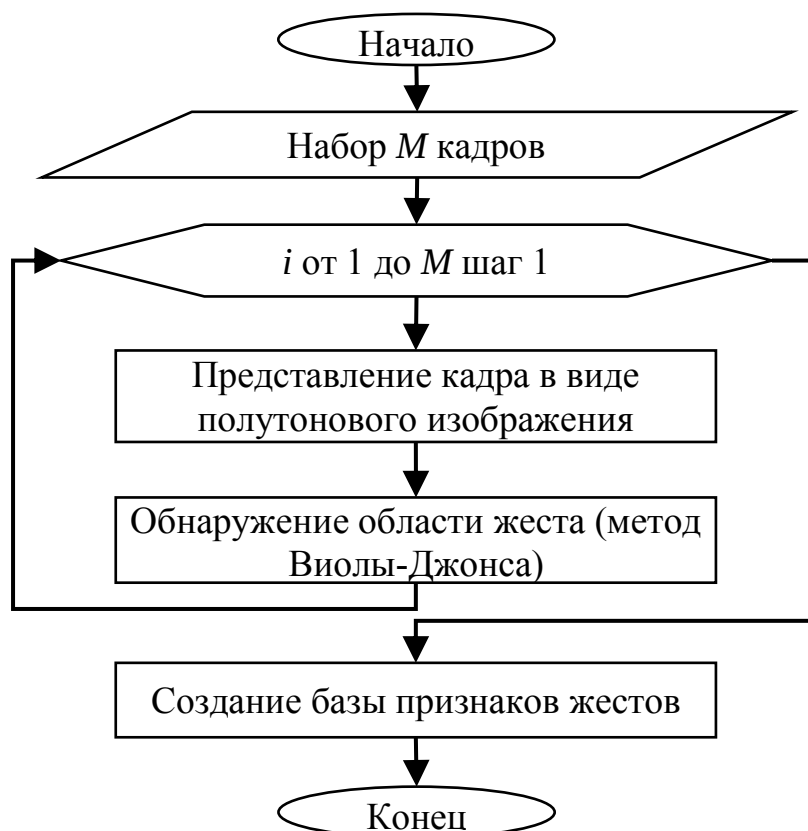


Рис. 2.24. Создание базы признаков жестов

Процесс распознавания жестов на видеопоследовательности происходит следующим образом (рис. 2.25).

Представляется очередной кадр в виде полутонового изображения. Затем к полутоновому изображению применяется метод Виолы-Джонса для поиска местоположения области жеста. Если область жеста не обнаружена, то метод Виолы-Джонса применяется для следующего очередного кадра. Если область жеста обнаружена, то к очередному кадру применяется алгоритм *CAMShift* для отслеживания области жеста. Если область жеста отслежена, то выполняется процесс распознавания жеста для этой области жеста.

Затем для следующего очередного кадра применяется алгоритм *CAMShift* и запускается процесс распознавания. Если область жеста не отслежена, применяется метод Виолы-Джонса для поиска нового местоположения области жеста.

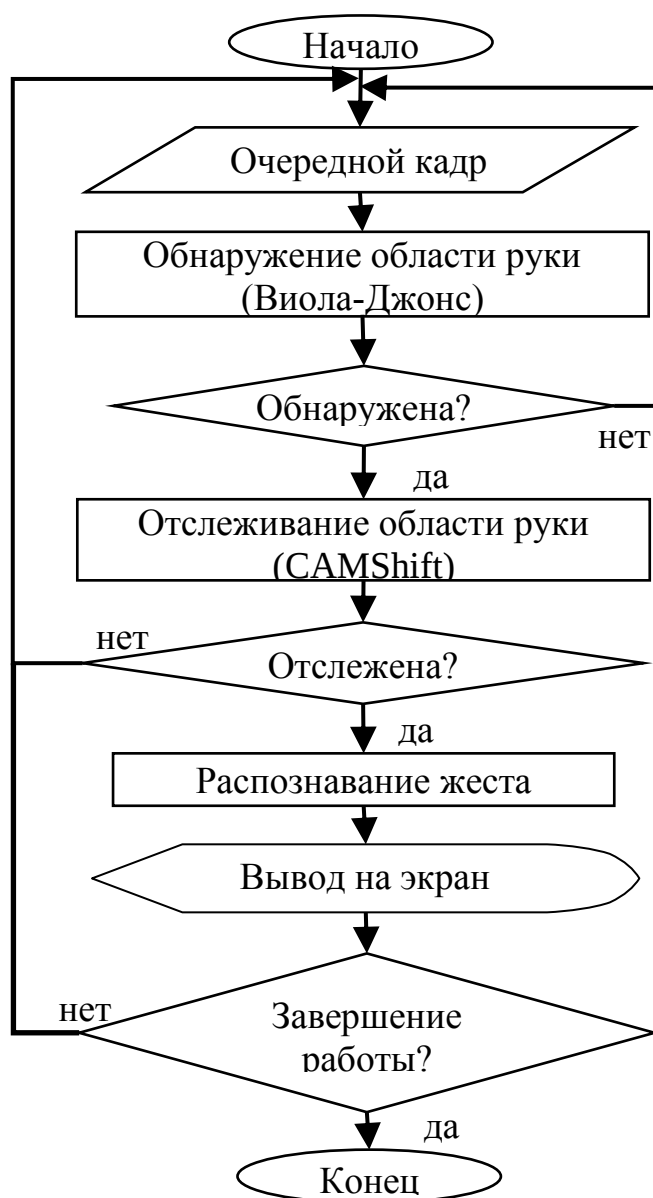


Рис. 2.25. Распознавание жестов

Предложенный алгоритм для распознавания жестов на видеопоследовательности в режиме реального времени работает только с областью жеста, благодаря этому он может обеспечить высокую скорость распознавания.

Для отслеживания жестов на видеопоследовательности предложено применение алгоритма *CAMShift*. Указанный алгоритм обладает следующими преимуществами по сравнению с другими алгоритмами: сравнительно высокой производительностью, способностью отслеживать вращение объектов в плоскости изображения и малой чувствительностью к изменениям освещенности. Благодаря указанным характеристикам алгоритм *CAMShift*

полностью соответствует задаче распознавания жестов на видеопоследовательности в режиме реального времени с высокой точностью.

Алгоритм *CAMShift* выполняет процедуру отслеживания объектов, которая состоит из следующих этапов [7].

Шаг 1. Выбор области поиска, внутри которой предполагается появление объекта. Выбирается начальное положение окна поиска.

Шаг 2. В области поиска строится изображение правдоподобия с помощью алгоритма обратной проекции гистограмм.

Итеративно выполняются следующие шаги:

Шаг 3. Рассчитываются моменты нулевого и первого порядка, которые определяют центр масс внутри окна поиска изображения.

Центр масс в окне поиска рассчитывается следующим образом:

1. Вычисление нулевого момента

$$M_{00} = \sum_x \sum_y I(x, y).$$

2. Вычисление первого момента для x и y

$$M_{10} = \sum_x \sum_y xI(x, y); M_{01} = \sum_x \sum_y yI(x, y).$$

3. Вычисление среднего смещения окна поиска

$$x_c = \frac{M_{10}}{M_{00}} \text{ и } y_c = \frac{M_{01}}{M_{00}}.$$

Шаг 4. Найденный центр масс определяет положение окна поиска на следующей итерации. Размеры окна не меняются.

Если различия в положении окна поиска в соседних итерациях малы (шаги 3–4), то вычисления прекращаются.

Шаг 5. Расширяется окно поиска. Подсчитываются моменты до второго порядка, которые определяют окончательное положение и размер отслеживаемого объекта внутри окна поиска изображения.

2.6 Основные результаты и выводы по главе 2

1. Предложен способ выделения признаков объектов на изображениях, основанный на совместном применении вейвлет-преобразований Хаара и Добеши, позволяющий эффективнее распознавать объекты по сравнению с применением вейвлет-преобразования Хаара или Добеши в отдельности.

2. Предложен алгоритм, основанный на применении вейвлет-преобразований и РСА для распознавания лиц и жестов на статических изображениях в присутствии шума.

3. На основе численных экспериментов проведен сравнительный анализ применимости вейвлет-преобразований Хаара, Добеши, РСА и их комбинации при решении задач распознавания лиц на статических изображениях в присутствии шума и в его отсутствии.

4. Разработан алгоритм и схема его применения для распознавания лиц на видеопоследовательности на основе применения предложенного алгоритма распознавания лиц на изображениях и метода Виолы-Джонса.

5. Создан алгоритм и схема его применения для распознавания жестов на видеопоследовательностях на основе применения предложенного алгоритма распознавания жестов и алгоритма *CAMShift*.

6. Перечисленные в пп. 2, 3 и 5 алгоритмы способны работать с высокой скоростью и точностью при распознавании объектов на изображениях в присутствии шума и на видеопоследовательностях в реальном режиме времени.

ГЛАВА 3. ТЕСТИРОВАНИЕ И ПРИМЕНЕНИЕ РАЗРАБОТАННЫХ АЛГОРИТМОВ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ И ЖЕСТОВ НА ИЗОБРАЖЕНИЯХ И ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ

В данной главе приводятся численные эксперименты по распознаванию объектов (лиц и жестов) на большом количестве статических изображений, видеороликов и видеопоследовательностей в реальном режиме времени. По результатам численных экспериментов делаются выводы об эффективности предложенного алгоритма распознавания объектов. Проводится сравнительный анализ предложенных алгоритмов с другими современными алгоритмами.

3.1 Постановка экспериментов

Для оценки эффективности предложенного алгоритма распознавания объектов требуется проведение численных экспериментов. В первой главе отмечалось, что изменение освещённости, внешности человека и т.п. существенно влияет на результаты распознавания объектов.

Процесс распознавания объектов состоит из двух частей: **обучения** и **распознавания**. В процессе обучения необходимо предъявить отдельный объект с указанием его принадлежности тому или другому образу [39]. В задаче распознавания лиц, классификатор определяет, является ли предъявляемый объект «лицом».

Процесс распознавания осуществляется следующим образом. Сначала определяется местонахождение объекта относительно гиперсфер. При этом выделяется прямоугольник для захватывания области трекинга объекта. В случае попадания объекта в гиперсферу, которая соответствует одному и только одному образу, процедура распознавания прекращается. Процесс продолжается до тех пор, пока принадлежность неизвестного объекта тому или иному образу не определится однозначно [17].

Сравнение качества распознавания объектов разнообразными методами осложнено многими причинами. Важнейшей из них является то, что в

большинстве случаев приходится полагаться только на данные испытаний, предоставляемые самими авторами. Не представляется возможным осуществления корректного сравнения между методами на едином наборе изображений из-за отсутствия: универсальных коллекций тестовых данных, общепринятой формулы для оценки вычислительной мощности и т.д.

3.2 Тестирование разработанного ПО

Для тестирования работы предложенных алгоритмов разработано ПО на языке объектно-ориентированного программирования C# (*Visual Studio* 2010) с использованием библиотек *OpenCV*, *AForge.Net* и *Emgu CV*. Созданы 3 пакета программ: «*PatternRecognition*», «*FaceRecognition*» и «*GestureRecognition*». Программы протестированы на ноутбуке с процессором *Intel Core™ 2 Duo* 2 ГГц, объемом оперативной памяти 2 Гб, видеокамерой 1,3 Мп, передающей 30 кадров в секунду с разрешением 320×240 пикселей.

3.2.1 Распознавание лиц на статических изображениях

Целью исследования является поиск изображений лиц, которые соответствуют заданным тестовым образам. Численные эксперименты по распознаванию лиц на статических изображениях выполнены на основе разработанного пакета «*PatternRecognition*» с использованием различных баз изображений лиц, таких как: *Collection of Facial Images*, *Yale Face Database* и *ORL Face Databases*.

База изображений лиц *Collection of Facial Images*

Эта база содержит 366 изображений лиц различных людей, по 20 изображений каждого лица [106]. При формировании БД размер изображений и условия съемки были одинаковыми. Применялся 24-битный формат JPEG. В базе содержатся снимки людей мужского и женского пола, разных национальностей и возрастов. В ней отражены изменения внешности человека (различные прически, присутствие бороды и очков).

Пример использованных изображений лиц *Collection of Facial Images* для численных экспериментов представлен на рис. 3.1.



Рис. 3.1. Пример изображений лиц БД *Collection of Facial Images*

В процессе подготовки к экспериментам были созданы две обучающие выборки. Первая из них содержит 5 изображений каждого человека (всего $5 \times 366 = 1830$ изображений). Вторая – 10 изображений лиц для обучения (всего $10 \times 366 = 3660$ изображений). Тестовая выборка состоит из 20-ти изображений лиц каждого человека (всего $20 \times 366 = 7320$ изображений).

При выполнении экспериментов используются два типа вейвлет-преобразования: Хаара и Добеши.

Результаты численных экспериментов распознавания лиц из БД *Collection of Facial Images* представлены в табл. 3.1.

Таблица 3.1. Точность распознавания лиц из БД *Collection of Facial Images*

Обучающая выборка	Тестовая выборка	Тип вейвлет-преобразования	
		Хаара, %	Добеши, %
1830	7320	94,57	94,69
3660	7320	98,40	98,40
Средняя точность		96,49	96,55

Как следует из табл. 3.1, результаты распознавания лиц показывают, что различие между двумя типами вейвлет-преобразования Хаара и Добеши составляет небольшую величину (0,06%).

База изображений лиц *Yale Face Database*

Тестирование работоспособности предложенного алгоритма проводилось на базе изображений лиц *Yale Face Database* [122], подготовленной в Йельском университете. В ней содержится 165 черно-белых изображений лиц 15 различных людей, по 11 изображений каждого лица при трех вариантах освещения, в очках и без очков, с шестью различными выражениями лица (радость, грусть, удивление и т.д.). Размер каждого изображения, представленного в формате GIF, составлял 320×243 пикселя.

Пример использованных изображений базы *Yale Face Database* для численных экспериментов показан на рис. 3.2.



Рис. 3.2. Пример изображений лиц базы *Yale Face Database*

При этом обучающая выборка содержит соответственно 15, 30, 45, 60 и 75 изображений (одно изображение одного лица, т.е. $1 \times 15 = 15$; 2 изображения одного лица, т.е. $2 \times 15 = 30$; 3 изображения одного лица, т.е. $3 \times 15 = 45$; 4 изображения одного лица, т.е. $4 \times 15 = 60$; 5 изображений одного лица, т.е. $5 \times 15 = 75$). Остальные 150, 135, 120, 105 и 90 изображений БД используются для тестирования.

При выполнении экспериментов использовалось вейвлет-преобразование Хаара.

Результаты численных экспериментов распознавания лиц из базы изображений лиц *Yale Face Database* представлены в табл. 3.2.

Таблица 3.2. Точность распознавания лиц из базы изображений лиц *Yale Face Database*

Обучающая выборка	Тестовая выборка	Точность, %
1 (1×15)	10 (10×15)	66,67
2 (2×15)	9 (9×15)	87,41
3 (3×15)	8 (8×15)	92,38
4 (4×15)	7 (7×15)	97,08
5 (5×15)	6 (6×15)	99,51

В случае, когда количество обучающих изображений сопоставимо с количеством изображений для тестирования, точность распознавания лиц является наилучшей.

База изображений лиц *ORL Face Databases*

Было проведено тестирование работоспособности предложенного алгоритма на БД *The ORL Database of Faces* [95], которая содержит 40 черно-белых изображений лиц различных людей, по 10 изображений каждого лица (всего 400 изображений). Эта база изображений лиц подготовлена в научно-исследовательской лаборатории компании *Olivetti*.

Все изображения в БД полутоновые и имеют 256 уровней яркости. Размер каждого изображения, представленного в формате PGM, составлял 92×112 пикселя. Снимки лиц сделаны в разное время, при различном освещении, для различной мимики (открытые/закрытые глаза, наличие улыбки или ее отсутствие) и лицевых деталей (наличие или отсутствие очков).

В процессе подготовки к экспериментам БД *The ORL Database of Faces* была разделена на две равные части. Каждая часть содержит 200 изображений лиц 40 человек, по 5 изображений каждого лица. Т.е. обучающая выборка содержит 200 изображений (5 изображений одного лица, т.е. $5 \times 40 = 200$). Остальные 200 изображений БД используются для тестирования.

В проводимом эксперименте используются два способа для выделения признаков изображений лиц:

- PCA;
- комбинация вейвлет-преобразований Хаара, Добеши и PCA (Хаар+Д4+PCA).

При этом на основе результатов экспериментов, приведенных в главе 2, применяются вейвлет-преобразования Хаара и Добеши первого уровня. Пороговое значение признания лица устанавливается равным 1700. Если величина расстояния Евклида d_k по формуле (2.2) меньше значения порога для некоторого k , то такое изображение лица принадлежит k -му классу. Иначе оно считается новым.

На рис. 3.3 показан пример 80 случайных изображений обучающей выборки, составленной из 200 случайных изображений БД ORL.



Рис. 3.3. Пример 80 случайных изображений из обучающей выборки

На рис. 3.4 показан пример 80 случайных изображений из тестовой выборки, составленной из 200 изображений БД ORL, оставшихся после формирования обучающей выборки.



Рис. 3.4. Пример 80 случайных изображений из тестовой выборки

На рис. 3.5 показан пример 80 случайных изображений, содержащих 5% шума типа «соль и перец». Указанные изображения выбирались из тестовой выборки, составленной, как и ранее, из 200 изображений БД ORL, оставшихся после формирования обучающей выборки.



Рис. 3.5. Пример 80 случайных изображений из тестовой выборки с 5% шума

На рис. 3.6 показан пример 80 случайных изображений, содержащих 15% шума типа «соль и перец». Указанные изображения выбирались из

тестовой выборки, составленной, как и ранее, из 200 изображений БД ORL, оставшихся после формирования обучающей выборки.



Рис. 3.6. Пример 80 случайных изображений тестовой выборки с 15% шума

На рис. 3.7 показан пример 80 случайных изображений, содержащих 30% шума типа «соль и перец». Указанные изображения выбирались из тестовой выборки, составленной, как и ранее, из 200 изображений БД ORL, оставшихся после формирования обучающей выборки.



Рис. 3.7. Пример 80 случайных изображений тестовой выборки с 30% шума

На рис. 3.8 представлены результаты применения PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания лиц, приведенных на рис. 3.3–3.7.

При этом обучающая выборка содержит 40 изображений (одно изображение одного лица, т.е. $1 \times 40 = 40$). Остальные 360 изображений БД используются для тестирования.

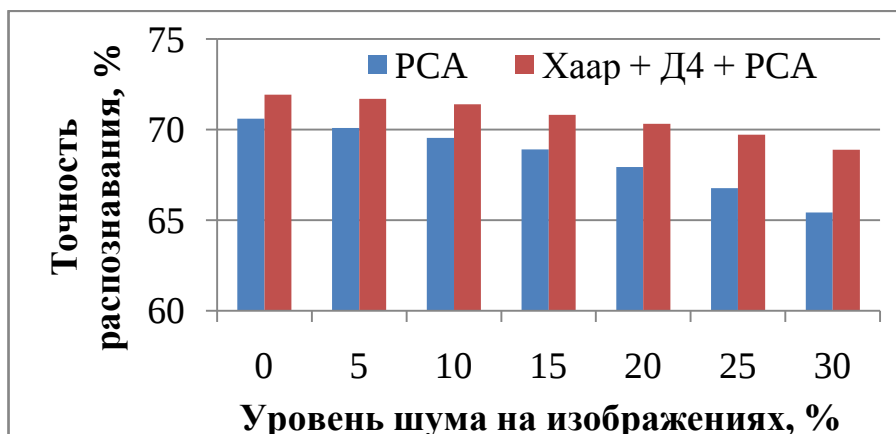


Рис. 3.8. Результаты распознавания лиц с 40 изображениями обучающей выборки и 360 изображениями тестовой выборки

На рис. 3.9 представлены результаты применения PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания лиц из БД ORL. Обучающая выборка содержит 80 изображений (2 изображения одного лица, т.е. $2 \times 40 = 80$). Остальные 320 изображений БД используются для тестирования.

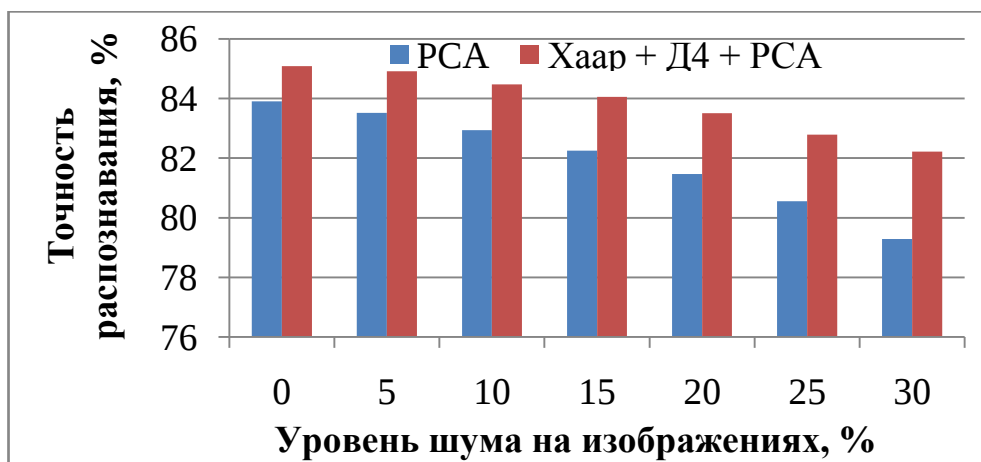


Рис. 3.9. Результаты распознавания лиц с 80 изображениями обучающей выборки и 320 изображениями тестовой выборки

На рис. 3.10 представлены результаты применения PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания лиц из БД ORL. Обучающая выборка содержит 120 изображений (3 изображения одного лица, т.е. $3 \times 40 = 120$). Остальные 280 изображений БД используются для тестирования.

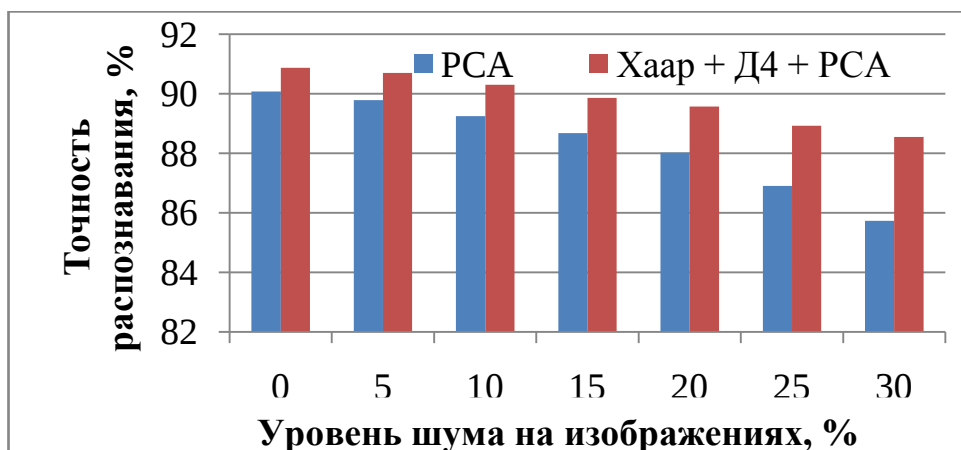


Рис. 3.10. Результаты распознавания лиц со 120 изображениями обучающей выборки и 280 изображениями тестовой выборки

На рис. 3.11 представлены результаты применения PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания лиц из БД ORL. Обучающая выборка содержит 160 изображений (4 изображения одного лица, т.е. $4 \times 40 = 160$). Остальные изображения БД используются для тестирования.

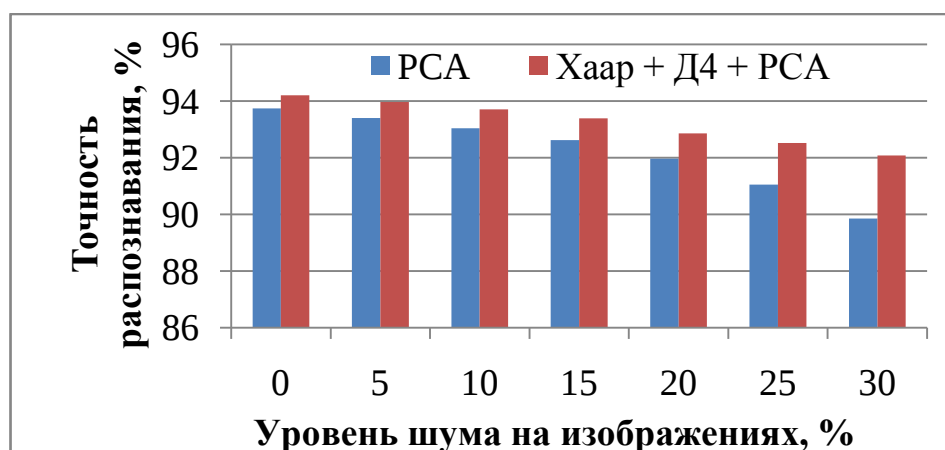


Рис. 3.11. Результаты распознавания лиц со 160 изображениями обучающей выборки и 240 изображениями тестовой выборки

На рис. 3.12 представлены результаты применения PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания лиц из БД

ORL. Обучающая выборка содержит 200 изображений (5 изображений одного лица, т.е. $5 \times 40 = 200$). Остальные 200 изображений БД используются для тестирования.

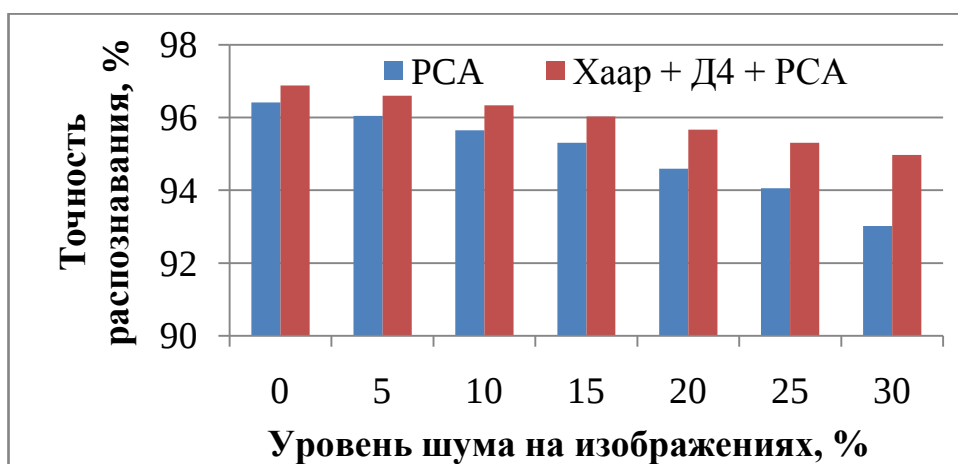


Рис. 3.12. Результаты распознавания лиц с 200 изображениями обучающей выборки и 200 изображениями тестовой выборки

На приведенных выше рисунках 3.8–3.12 показано, что комбинация вейвлет-преобразований Хаара, Добеши и PCA распознает лица эффективнее, чем PCA как в отсутствии, так и в присутствии шумов на изображениях тестовой выборки. Разница между PCA и комбинацией вейвлет-преобразований Хаара, Добеши и PCA становится более заметной в условиях зашумленных изображений тестовой выборки.

Результаты, полученные на основе предложенного алгоритма, сопоставлялись с данными, полученными на основе применения следующих алгоритмов: *Gabor* и PCA [85], PCA + *Moment Invariant* + *Face Colour* [89], PCA + *Moment Invariant* [97], PDBNN (*Probabilistic Decision-based Neural Network*) [84], *Point-matching* [80], *Pseudo 2-D HMM (two-dimensional Hidden Markov Models)* + DCT (*Discrete Cosine Transform*) [63], LVQ (*Learning Vector Quantization*) + RBF (*Radial Basic Function*) + FEC (*Forward Error Correction*) [112], PCA + RBF (*Radial Basic Function*) [119], UDT (*Uncorrelated Discriminant Transformation*) [73], *Wavelet* + RBF (*Radial Basic Function*) [124], RBF (*Radial Basic Function*) [87] (рис. 3.13).

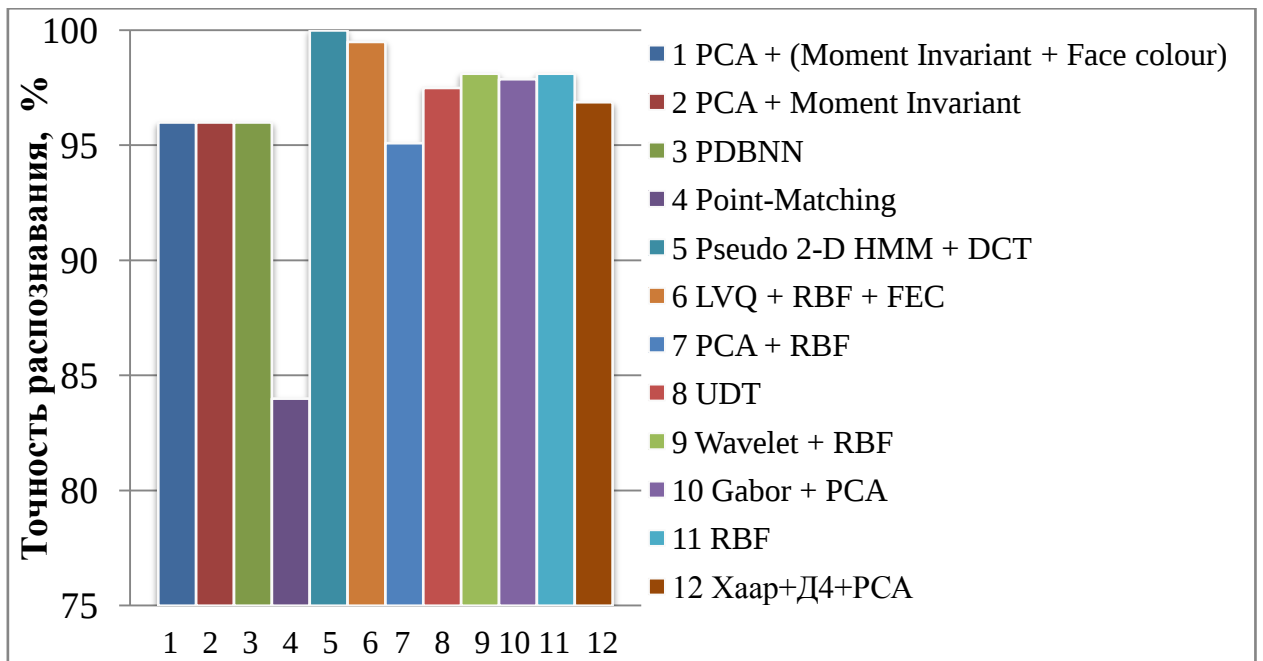


Рис. 3.13. Сравнение различных алгоритмов на задаче распознавания лиц из БД *ORL Face Databases*

На рис. 3.13 показаны результаты распознавания лиц из БД *ORL Face Databases* двенадцатью методами и алгоритмами, приведенными выше. Худшие результаты получились у алгоритма *Point-matching*. Остальные алгоритмы достигли точности распознавания лиц более 90%. Результаты работы предложенного алгоритма (№ 12) сопоставимы с методами: *PCA Moment Invariant Face Colour*, *PCA Moment Invariant*, *PDBNN*, *PCA RBF*, *UDT*, *Wavelet RBF* и *Gabor PCA*.

3.2.2 Распознавание жестов на статических изображениях

Для тестирования работы алгоритма распознавания жестов на статических изображениях используется часть БД *Cambridge Gesture Database* [76]. Эта база изображений жестов состоит из 5 различных частей, изображения в которых получены при различных условиях освещенности (рис. 3.14). В данной работе все жесты в БД делятся на 12 классов, представленных на рис. 3.15.



Рис. 3.14. Примеры изображений жестов каждой части БД *Cambridge Gesture Database* при различных условиях освещенности

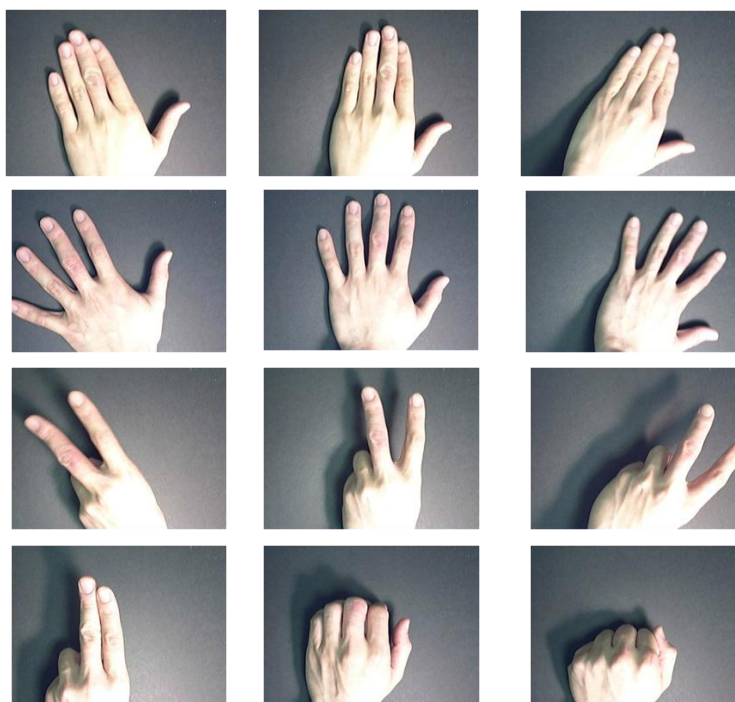


Рис. 3.15. Примеры 12 классов, использующихся при распознавании жестов

Для каждой из 5 выше приведенных частей создана база изображений жестов для тестирования, которая содержит 200 изображений каждого класса (всего $12 \times 200 = 2400$ изображений). Смещенная база изображений жестов из 5 частей для тестирования содержит 1000 изображений каждого класса (всего $12 \times 1000 = 12000$ изображений). Для каждой части также создана база изображений жестов для обучения, которая содержит 20 изображений каждого класса (всего $12 \times 20 = 240$ изображений). Смещенная база изображений жестов из 5 частей для обучения содержит 100 изображений каждого класса (всего $12 \times 100 = 1200$ изображений).

На рис. 3.16 показан пример изображений жеста первой части тестовой выборки, содержащих 5, 15 и 30% шума типа «соль и перец».

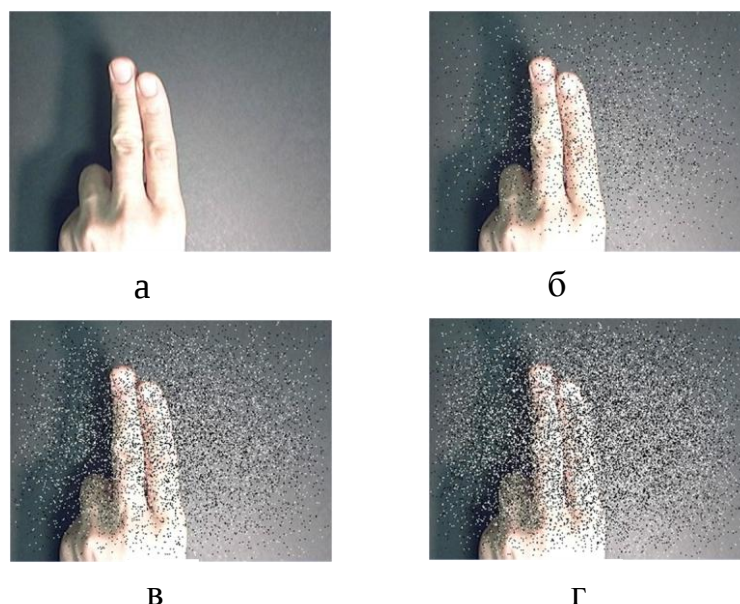


Рис. 3.16. Пример изображений жеста часть 1: а) изображение без шума, б–г) изображение с шумом 5, 15 и 30% соответственно

В проводимом эксперименте используются три способа для выделения признаков изображений жестов:

- комбинация вейвлет-преобразования Хаара и PCA (Хаар+PCA);
- комбинация вейвлет-преобразования Добеши и PCA (Д4+PCA);
- комбинация вейвлет-преобразований Хаара, Добеши и PCA (Хаар+Д4+PCA).

При этом на основе результатов экспериментов, приведенных в главе 2, применяются вейвлет-преобразования Хаара и Добеши первого уровня.

Для распознавания жестов на статических изображениях использована разработанная оригинальная программная реализация «*PatternRecognition*».

На рис. 3.17 приведены результаты применения комбинации вейвлет-преобразования Хаара и PCA, комбинации вейвлет-преобразования Добеши и PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания жестов из первой части базы *Cambridge Gesture Database*.

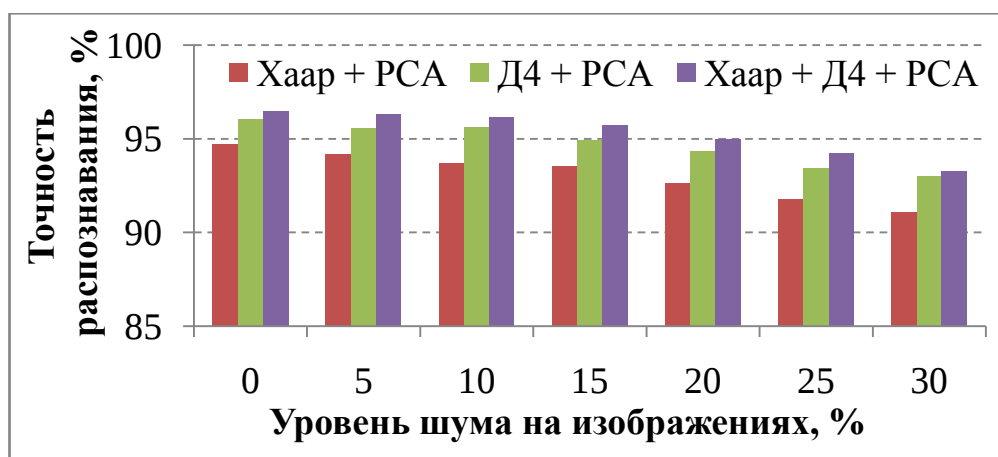


Рис. 3.17. Результаты распознавания жестов из первой части базы *Cambridge Gesture Database*

На рис. 3.18 приведены результаты применения комбинации вейвлет-преобразования Хаара и PCA, комбинации вейвлет-преобразования Добеши и PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания жестов из второй части базы *Cambridge Gesture Database*.

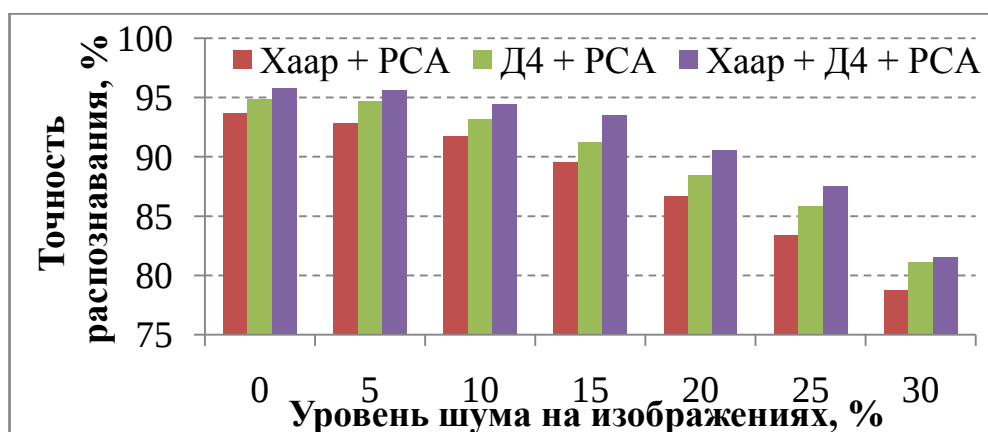


Рис. 3.18. Результаты распознавания жестов из второй части базы *Cambridge Gesture Database*

На рис. 3.19 приведены результаты применения комбинации вейвлет-преобразования Хаара и PCA, комбинации вейвлет-преобразования Добеши и PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания жестов из третьей части базы *Cambridge Gesture Database*.

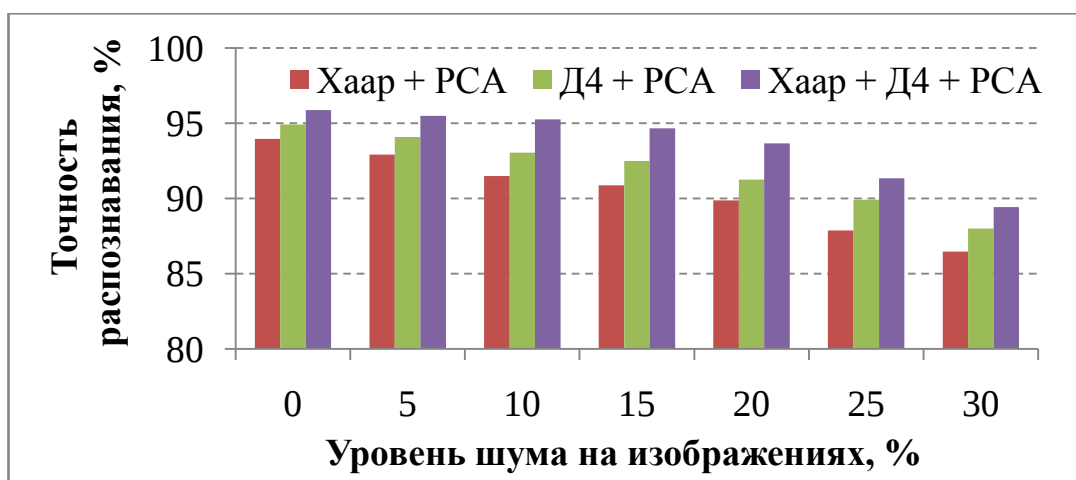


Рис. 3.19. Результаты распознавания жестов из третьей части базы *Cambridge Gesture Database*

На рис. 3.20 приведены результаты применения комбинации вейвлет-преобразования Хаара и PCA, комбинации вейвлет-преобразования Добеши и PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания жестов из четвертой части базы *Cambridge Gesture Database*.

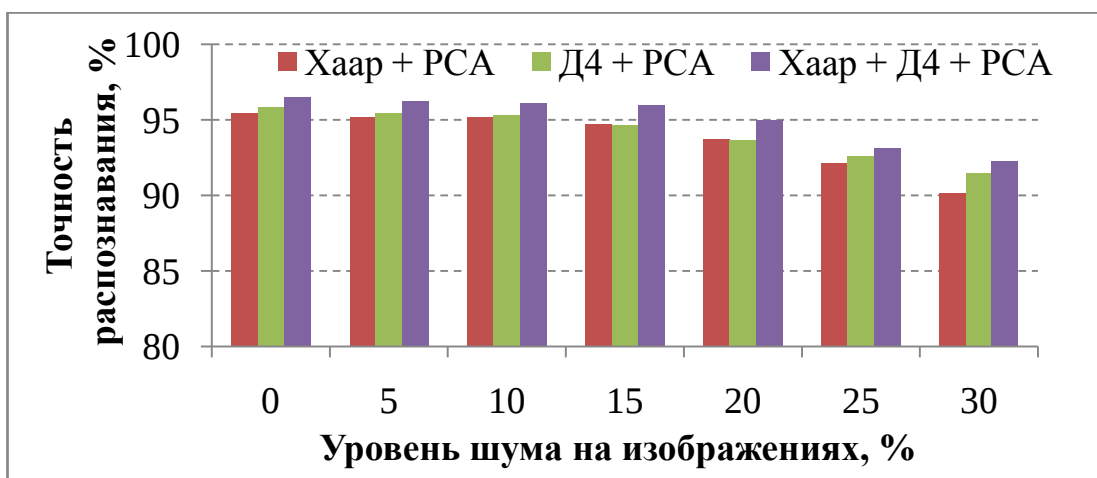


Рис. 3.20. Результаты распознавания жестов из четвертой части базы *Cambridge Gesture Database*

На рис. 3.21 приведены результаты применения комбинации вейвлет-преобразования Хаара и PCA, комбинации вейвлет-преобразования Добеши и PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания жестов из пятой части базы *Cambridge Gesture Database*.

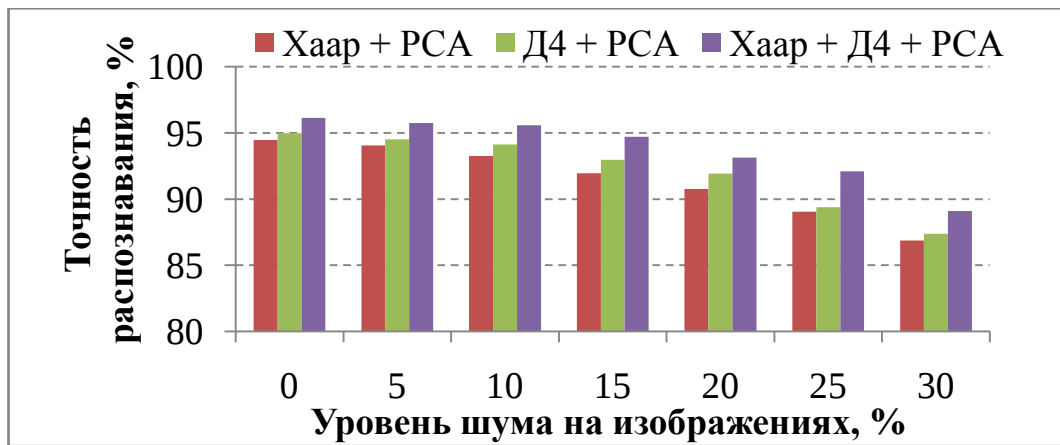


Рис. 3.21. Результаты распознавания жестов из пятой части базы *Cambridge Gesture Database*

На рис. 3.22 приведены результаты применения комбинации вейвлет-преобразования Хаара и PCA, комбинации вейвлет-преобразования Добеши и PCA и комбинации вейвлет-преобразований Хаара, Добеши и PCA для распознавания жестов из смещенной части базы *Cambridge Gesture Database*.

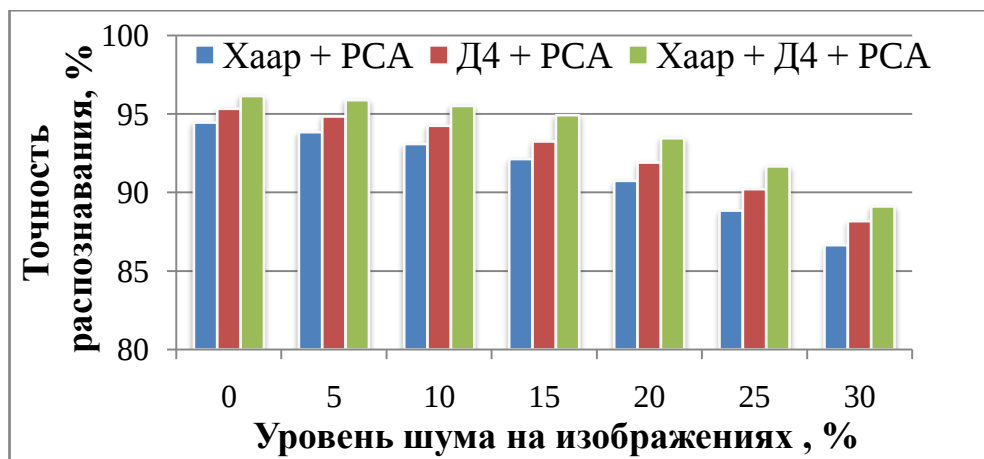


Рис. 3.22. Результаты распознавания жестов из смещенной части базы *Cambridge Gesture Database*

На приведенных выше рисунках 3.17–3.22 показано, что комбинация вейвлет-преобразований Хаара, Добеши и PCA распознает жесты эффективнее, чем комбинация вейвлет-преобразования Хаара и PCA и комбинация вейвлет-преобразования Добеши и PCA как в отсутствии, так и в присутствии шумов на изображениях тестовой выборки.

Результаты распознавания жестов на статических изображениях на основе предложенного алгоритма сопоставлялись с данными, полученными на основе применения следующих методов:

- *Support Vector Machine* (SVM) и *Relevance Vector Machine* (RVM) [121] (SVM и RVM);
- RVM на векторах 3D SIFT (3D SIFT RVM) [110];
- использование *Canonical Correlations* (CC) [53], *spatiotemporal canonical correlations* (ST-CC), *discriminative ST-CC* (ST-DCC);
- использование CC дескриптора SIFT (SIFT CC), *spatiotemporal canonical correlations of the SIFT vectors* (SIFT ST-CC) и SIFT ST-CC;
- модель иерархической временной памяти (модель ИВП) [5].

На рис 3.23 показано сравнение результатов распознавания жестов предложенным алгоритмом с результатами, полученными другими методами.

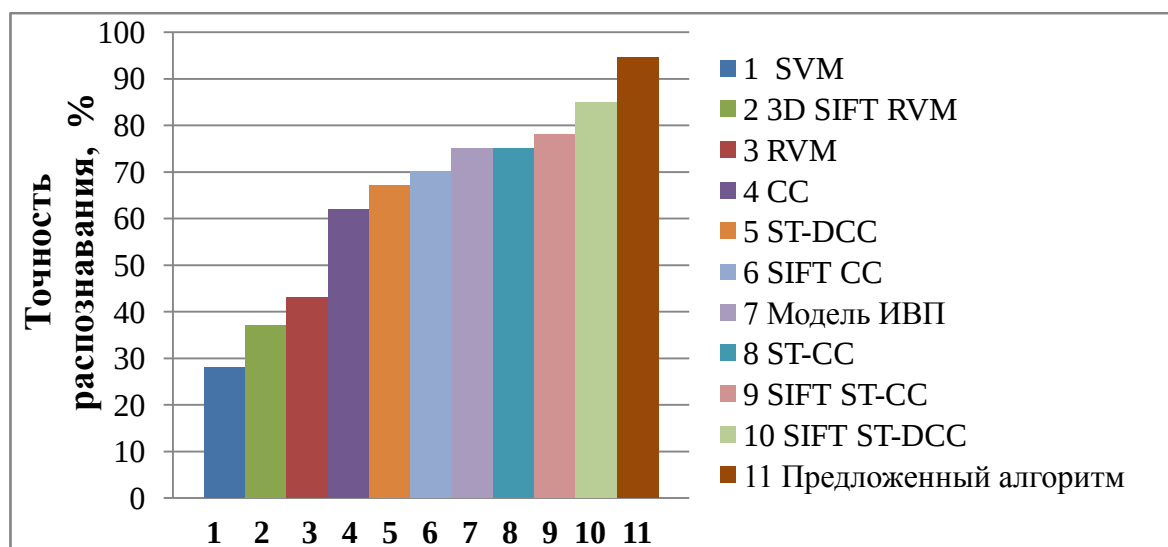


Рис. 3.23. Сравнение результатов распознавания жестов предложенным алгоритмом с результатами, полученными на основе других алгоритмов

На рис. 3.23 представлены результаты сопоставления созданного алгоритма (№11) с десятью известными алгоритмами для распознавания жестов, которые показывают, что предложенный алгоритм является наилучшим. Рассмотренные алгоритмы MGO SVM, 3D SIFT RVM, MGO RVM, CC, ST-DCC, SIFT CC, ST-CC и SIFT ST-CC дают точность распознавания жестов ниже 80%.

3.2.3 Распознавание множества лиц на видеопоследовательности

Для обработки последовательностей изображений и распознавания множества лиц в работе создана программная реализация «*FaceRecognition*», основанная на предложенном алгоритме.

На этапе обучения каждый человек находится перед камерой, совершая поступательные и вращательные движения головой. Вначале программа работает с детектором лиц. Она определяет появление лица в кадре. Если формируется ответ «Да», то программа захватывает и записывает 10 различных изображений одного человека, которые сохраняются в базе эталонных данных с именем этого человека. БД состоит из изображений, захваченных камерой системы.

Для проверки правильности работы программы этот процесс повторяется и программа в реальном времени определяет – кто в данный момент находится перед видеокамерой (процедура идентификации). В процессе идентификации захваченное лицо сравнивается с изображениями лиц в БД. При этом на экране появляются имена захваченных лиц.

Примеры работы программы распознавания лиц и двух лиц на видеопоследовательности в режиме реального времени приведены на рис. 3.24, 3.25.

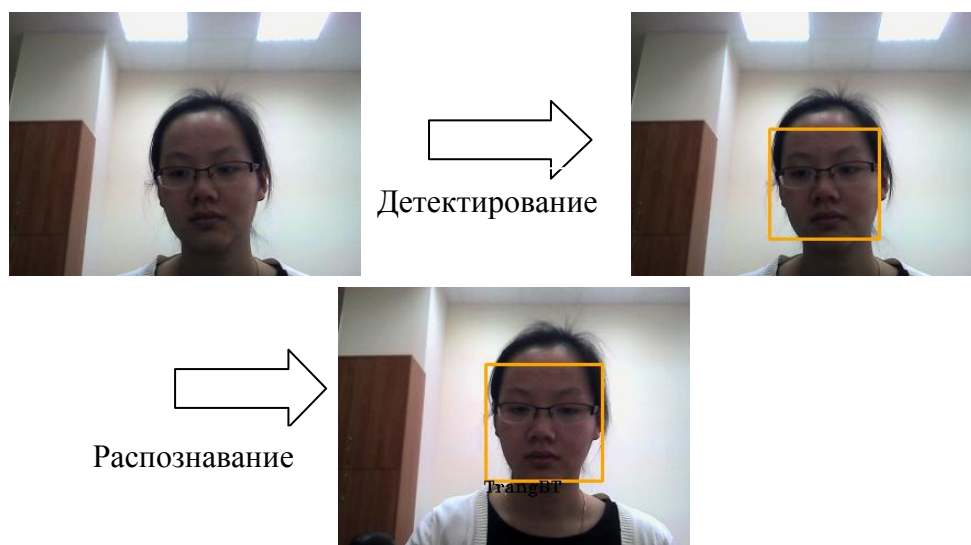


Рис. 3.24. Пример распознавания лица на видеопоследовательности в режиме реального времени

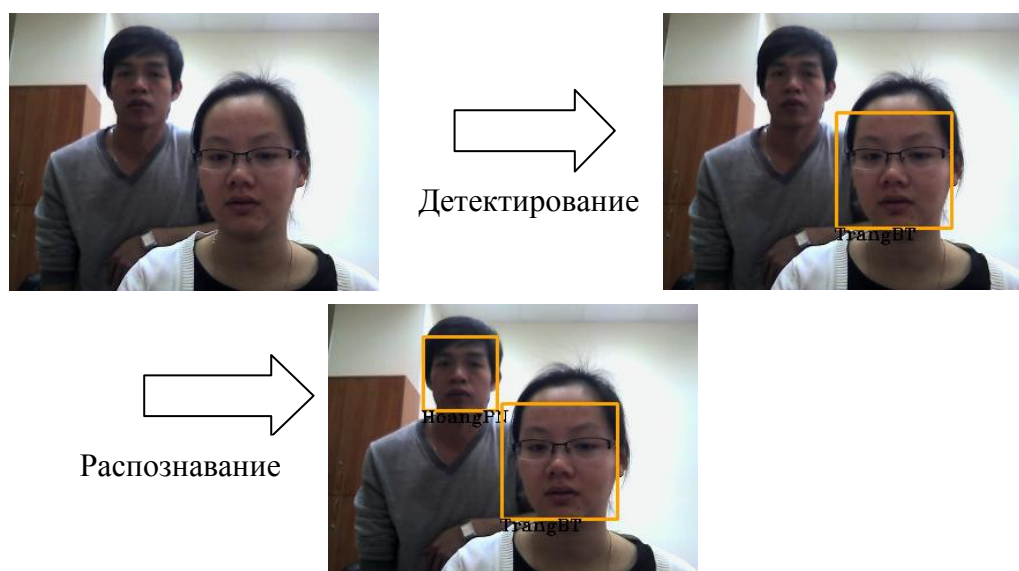


Рис. 3.25. Пример распознавания двух лиц на видеопоследовательности в режиме реального времени

База видеороликов NRC-IIT: *Facial Video Database*

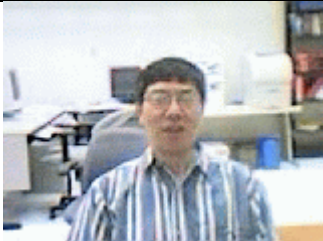
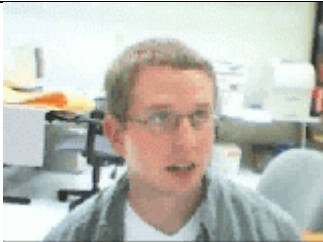
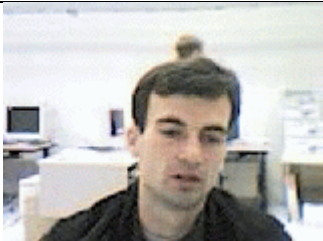
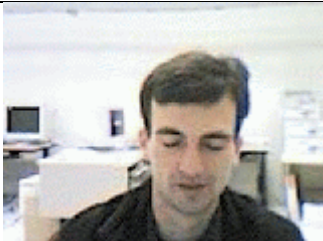
В процессе подготовки к экспериментам были взяты видеоролики из базы NRC-IIT: *Facial Video Database* [68]. Эта БД была создана для того, чтобы сформировать критерий оценки эффективности методов и алгоритмов распознавания лиц на видеопоследовательности. Создание указанной БД необходимо еще и для того, чтобы изучить влияние различных факторов и параметров на производительность распознавания.

База содержит 10 пар коротких видеороликов (один видеоролик для обучения, а второй – для тестирования), в которых показываются лица пользователей компьютера. Эти пользователи находятся перед монитором, совершая поступательные и вращательные движения головой. Лицо занимает от 1/4 до 1/8 части изображения.

В этой базе каждые два видеоролика одного лица снимались один за другим и условия съемки были одинаковыми. Эта база подходит для тестирования производительности алгоритма распознавания лиц на видеороликах. Исследуется влияние на производительность следующих факторов: низкого разрешения, размытости изображения, расположения лица вне фокуса (*out-of focus factor*), вариации выражения лица, изменения ориентации лица и окклюзии (*occlusions*).

Примеры видеороликов представлены в табл. 3.3.

Таблица 3.3. Примеры видеороликов для эксперимента

ID	Примеры видеороликов для обучения		Примеры видеороликов для тестирования	
	Видеоролики	Количество кадров	Видеоролики	Количество кадров
1		237		329
2		257		339
3		286		357
4		353		404
5		78		128

ID	Примеры видеороликов для обучения		Примеры видеороликов для тестирования	
	Видеоролики	Количество кадров	Видеоролики	Количество кадров
6		324		353
7		258		328
8		346		426
9		318		388
10		338		378

При проведении экспериментов по распознаванию лиц применялся алгоритм, основанный на комбинации вейвлет-преобразований Хаара, Добеши и PCA со значениями параметра $D=280$ и порога признания лиц $\delta=1700$. Сравнение результатов распознавания лиц на видеопоследовательностях предложенным алгоритмом с данными,

полученными на основе применения алгоритма «*Associative neural networks*» (ANN), приведено на рис. 3.26.

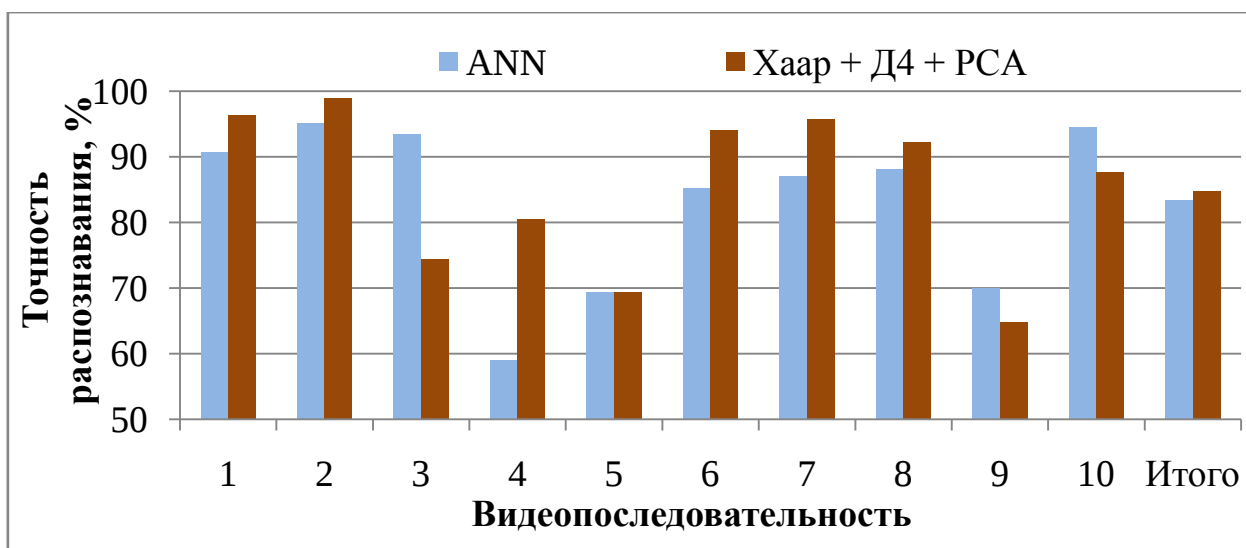


Рис. 3.26. Сравнение результатов распознавания лиц на видеопоследовательностях предложенным алгоритмом с алгоритмом ANN

Средняя точность распознавания лиц на видеопоследовательности предложенным алгоритмом составляет 85%, а алгоритмом ANN – 83%.

Результаты распознавания лиц на видеопоследовательностях предложенным алгоритмом показали в среднем более высокую точность по сравнению с данными, полученными на основе применения алгоритма ANN.

3.2.4 Распознавание жестов на видеопоследовательности

Для обработки последовательностей изображений и распознавания множества жестов в работе создана программная реализация «*GestureRecognition*».

На этапе обучения человек находится перед камерой, совершая движение рукой. Вначале программа работает с детектором области руки. Она определяет появление руки в кадре. Если формируется ответ «Да», то программа захватывает и записывает 10 различных изображений для каждого класса, которые сохраняются в базе эталонных данных. БД состоит из изображения жестов, захваченных камерой системы. В процессе работы на очередном видеокadre программа отслеживает область руки на основе применения алгоритма *CAMShift*. После этого происходит распознавание

жеста и на экране указывается, к какому классу относится представленный на кадре жест.

При тестировании работы алгоритма распознавания жестов на видеопоследовательности в реальном режиме времени используются 6 классов жестов: ноль, один, два, три, четыре и пять. Примеры результатов применения предложенного алгоритма для распознавания жестов на видеопоследовательности представлены на рис. 3.27. При выполнении тестирования программы распознавания жестов на видеопоследовательности наклон руки изменяется медленно по разным направлениям. Скорость обработки составляет 30–31 кадр в секунду.

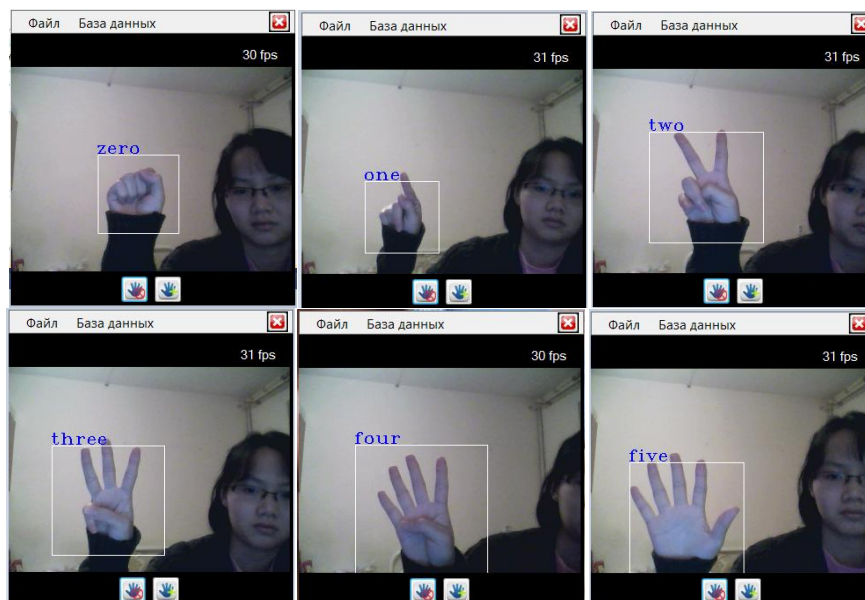


Рис. 3.27. Результаты работы алгоритма распознавания жестов на видеопоследовательности

3.3 Перспективы использования полученных результатов в жизни

На сегодняшний день вычислительная техника применяется в различных областях деятельности человека. Проблема безопасности играет важную роль в нашей жизни. После террористических актов 11 сентября 2001 г. в Америке, указанная проблема стала более острой. Лицо является самым распространенным в настоящее время биометрическим параметром, используемым для распознавания и идентификации человека.

Существует проблема распознавания замаскированного человека. Например, если человек отпустил или наоборот сбрил бороду, то у него есть много шансов не быть распознанным системой. Службы пограничного и паспортного контроля аэропортов, вокзалов, иммиграционные службы имеют все основания использовать преимущества систем распознавания лиц для автоматического сравнения фотографии на документе с «живым» лицом, поиска и предоставления информации из собственных БД, формирования видеоархива лиц людей, прибывающих в страну.

С 2012 г. в российском метро внедряют систему распознавания лиц для пассажиров [41]. На станциях метро устанавливается рентгеновское оборудование для досмотра пассажиров и их багажа, рамки металлодетекторов, газоанализаторы, средства радиационного контроля, оборудование подавления сигналов радиовзрывателей и другое оборудование.

Компьютерные игры помогают человеку развивать его интеллектуальные способности. Распознавание жестов на видеопоследовательностях уже стало привычным средством управления, например, в спортивной игре на приставке *Wii* [120], в развлечении на базе приставки *Xbox* и системы *360 Kinect* [77].

Система распознавания жестов может использоваться для удаленного управления затвором фотокамеры, для передачи информации, для управления компьютером или для переключения каналов телевизора. Перед началом работы в систему загружается определенная БД известных жестов, в которой представлены все возможные классы жестов. Затем, по изображению, полученному с камеры, происходит распознавание номера жеста. Этот номер затем может быть использован для выполнения какого-либо заранее заданного действия.

Система распознавания жестов позволяет взаимодействовать с компьютером путем совершения оператором специальных указаний руками в поле видимости видеокамеры. Такая система относится к системам класса

«Человеко-машинное взаимодействие» (*Human-Computer Interface*). Основными компонентами системы «Человек-компьютер» являются компьютер и человек-оператор системы, представляющий пользователя ЭВМ [22].

3.4 Основные результаты и выводы по главе 3

1. В данной главе приведено описание результатов апробации предложенных алгоритмов, основанных на совместном применении вейвлет-преобразований, PCA, метода Виолы-Джонса и алгоритма *CAMShift* для распознавания лиц и жестов в присутствии шума на статических изображениях и видеопоследовательностях в реальном режиме времени. Проведенные численные эксперименты показали, что предложенные алгоритмы являются эффективными для решения задач распознавания статических и динамических объектов. Осуществлено сопоставление предложенного алгоритма распознавания объектов с известными аналогами.

2. Анализ результатов сравнения данных проведенных численных экспериментов по распознаванию объектов на статических изображениях с результатами применения известных алгоритмов позволяет сделать вывод об успешной работе созданного алгоритма, основанного на применении вейвлет-преобразований и PCA.

3. Результаты проведенных компьютерных экспериментов показали, что предложенные оригинальные алгоритмы позволяют эффективно распознавать движущиеся объекты (лица и жесты) на видеопоследовательностях в режиме реального времени.

4. Проведено сопоставление результатов распознавания лиц на видеопоследовательностях с данными, полученными на основе применения алгоритма «*Associative neural networks*», которое показало в среднем более высокую точность распознавания лиц на видеопоследовательностях предложенным алгоритмом.

ГЛАВА 4. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ И ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ

Данная глава посвящена разработке на основе предложенных алгоритмов ПО для распознавания объектов (лиц и жестов) на статических изображениях и движущихся объектов на видеопоследовательностях в реальном режиме времени. Представлено описание архитектур разрабатываемых библиотек и программ распознавания объектов.

4.1 Выбор средств разработки

Для разработки ПО была выбрана визуальная среда программирования *Visual Studio C#*. Среда *Visual Studio C#* представляет собой интегрированную оболочку разработчика, в которую входит набор специализированных программ, ответственных за разные этапы создания готового приложения. Инструменты *Visual Studio C#* легко и удобно использовать для создания форм программы, к тому же *Visual Studio C#* поддерживает объектно-ориентированное программирование.

C# – один из языков программирования, который может использоваться для создания приложений, выполняемых в среде .NET CLR. Язык C# создан компанией *Microsoft* специально для использования на платформе .NET. Этот язык включает в себя самую полную поддержку структурного, компонентного ориентированного программирования, которую только можно ожидать от современного языка [2].

OpenCV

OpenCV (*Open Computer Vision Library*) обеспечивает выполнение множества операций по интерпретации изображений, ее функции включены в состав библиотеки *Intel Performance Primitives* (IPP). *OpenCV* реализована на языке C и содержит алгоритмы для сложных операций, таких как калибровка камеры по эталону, устранение оптических искажений, определение сходства, анализ перемещения объекта, определение формы объекта и

слежение за объектом, 3D-реконструкция, сегментация объекта, распознавание жестов и т.д. [94].

Open CV имеет следующие основные модули:

1. *Opencv_core* – основная функциональность. Включает в себя базовые структуры, вычисления, ввод и вывод для XML и т.д.
2. *Opencv_imgproc* – обработка изображений.
3. *Opencv_highgui* – модуль для создания пользовательского интерфейса.
4. *Opencv_feature2d* – распознавание и описание плоских примитивов (SURF, FASR и др.).
5. *Opencv_video* – анализ движения и отслеживание объектов.
6. *Opencv_objdetect* – обнаружение объектов на изображении.
7. *Opencv_calib3d* – калибровка камеры, поиск стерео-соответствия и элементы обработки трехмерных данных.
8. *Opencv_ml* – модели машинного обучения.
9. *Opencv_flann* – библиотека быстрого поиска ближайших соседей и обертка OpenCV.

AForge.NET – общедоступная структура C#, разработанная для разработчиков и исследователей в областях компьютерного зрения и искусственного интеллекта: обработки изображений, генетических алгоритмов, нечеткой логики, машинного обучения, робототехники и т.д. [50].

AForge представляет собой набор библиотек и примеров приложений, которые демонстрируют свои особенности:

1. *AForge.Imaging* – библиотека обработки изображений шаблонов и фильтров;
2. *AForge.Vision* – библиотека компьютерного зрения;
3. *Aforge.Video* – набор библиотек для обработки видео;
4. *AForge.Neuro* – библиотека для работы с нейронными сетями;

5. *AForge.Genetic* – библиотека для работы с генетическими алгоритмами;
6. *AForge.Fuzzy* – библиотека нечетких вычислений;
7. *AForge.Robotics* – библиотека, обеспечивающая поддержку некоторых наборов робототехники;
8. *AForge.MachineLearning* – библиотека машинного обучения.

Библиотека *Emgu CV* полностью написана на C#. Ее преимущества заключаются в том, что она может быть составлена в *Monodevelop* и, следовательно, может работать на любой платформе с поддержкой *Monodevelop*, в том числе *Windows*, *Linux*, *Mac OS X*, *iOS*, *iPhone*, *Ipad* и *Android* устройств.

Emgu CV также называют библиотекой машинного зрения [64]. *Emgu CV* используется для решения различных задач связанных с 2D графикой, распознаванием лиц и предметов на фотографиях и видеоизображениях.

4.1.1 Структура классов библиотеки

а. Классы для распознавания лиц и жестов на статических изображениях

Для реализации алгоритма распознавания лиц и жестов на статических изображениях создаются следующие классы: класс *Constant*, класс *Wavelet*, класс *WaveletImage*, класс *EigenObjects*, класс *EigenObjectsManager* и класс *ImageProcessing*.

Класс констант («Constant»)

Класс ***Constant*** предназначен для сохранения значений констант, таких как ссылки к файлам, в которых содержатся коэффициенты вейвлет-преобразования (Хаара и Добеши); размеры пикселей цветного и полутонового изображений; минимальное и максимальное значение пикселя полутонового изображения.

Объявление каждой переменной предваряется элементом *доступ*. Здесь элемент *доступ* означает *спецификатор доступа* (`public static`).

Структура класса **Constant** представлена на рис. 4.1. Основные переменные класса **Constant** приведены в табл. 4.1.

+ Constant
+ gray_min_value: byte + gray_max_value: byte + wavelet_haar: string + wavelet_daubechies_4: string

Рис. 4.1. Структура класса *Constant*

Таблица 4.1. Основные переменные класса *Constant*

Название переменной	Значение	Описание
byte GRAY_MIN_VALUE	0	Минимальное значение пикселя полутонового изображения.
byte GRAY_MAX_VALUE	255	Максимальное значение пикселя полутонового изображения.
string WAVELET_HAAR	@"Source\Wavelets\Haar.txt"	Ссылка к файлу, в котором содержатся коэффициенты для преобразования Хаара.
string WAVELET_DAUBECHIES_4	@"Source\Wavelets\Daubechies4.txt"	Ссылка к файлу, в котором содержатся коэффициенты для преобразования Добеши 4.

Класс вейвлетов («Wavelet»)

Класс **Wavelet** используется для загрузки из файла вейвлет-коэффициентов преобразований Хаара и Добеши. Класс **Wavelet** включает в себя локальные массивы коэффициентов, предназначенных для вычисления значений аппроксимации и различия.

В данном классе содержится конструктор, который загружает коэффициенты преобразования из файла в массивы. Класс содержит метод `get`, предназначенный для доступа к локальным массивам. Структура класса **Wavelet** представлена на рис. 4.2. Основные переменные и методы класса **Wavelet** приведены в табл. 4.2 и 4.3.

+ Wavelet
- highCoeffs: double[] - lowCoeffs: double[] - length: int
+ Wavelet(string) + getHighCoeffs(): double[] + getLowCoeffs(): double[] + getLength(): int

Рис. 4.2. Структура класса *Wavelet*

Таблица 4.2. Основные переменные класса *Wavelet*

Название переменной	Описание
private double[] highCoeffs	Массив вейвлет-коэффициентов для вычисления значения аппроксимации.
private double[] lowCoeffs	Массив вейвлет-коэффициентов для вычисления значения различия.
private int length	Количество коэффициентов вейвлет-преобразования (для Хаара – 1, Добеши – 4).

Таблица 4.3. Основные функции класса *Wavelet*

Название функции	Описание
Wavelet (string iFile)	Загрузка коэффициентов из файла iFile и сохранение этих признаков в массивах highCoeffs и lowCoeffs.
public double[] getHighCoeffs()	Доступ к массиву коэффициентов для вычисления значения аппроксимации.
public double[] getLowCoeffs()	Доступ к массиву коэффициентов для вычисления значения различия.
public int getLength()	Доступ к количеству коэффициентов вейвлет-преобразования.

Класс преобразования изображения в вейвлет-коэффициенты («*WaveletImage*»)

Класс *WaveletImage* преобразует полутоновое изображение в вейвлет-коэффициенты и сохраняет их в этом же классе для дальнейшей обработки. Класс *WaveletImage* включает в себя переменные, предназначенные для сохранения типа вейвлета, уровня разложения, ширины и высоты

изображения, временных преобразованных данных и преобразованных данных.

В классе содержатся методы, которые преобразуют исходное изображение по горизонтали и вертикали. Эти методы используются в конструкторе класса для преобразования исходного изображения по заданному типу вейвлета и заданному уровню разложения. Класс имеет методы, представляющие все коэффициенты вейвлета и коэффициенты вейвлета низкочастотного субдиапазона в виде изображения. Структура класса **WaveletImage** представлена на рис. 4.3. Основные переменные и методы класса **WaveletImage** приведены в табл. 4.4 и 4.5.

+ WaveletImge
- grayTemp2D: double[, ,] - grayData2D: double[, ,] - width: int - height: int - level: int - wavelet: Wavelet
+ WaveletImage(Imge<Gray, byte>, int, string) + Image(): Image<Gray, byte> + GetSubLL(): Image<Gray, byte> - ValidateIntensity(double): byte - HorizontalTransform2D(int): void - VerticalTransform2D(int): void

Рис. 4.3. Структура класса *WaveletImage*

Таблица 4.4. Основные переменные класса *WaveletImage*

Название переменной	Описание
private double[, ,] grayTemp2D	Матрица временных преобразованных данных полутонового изображения.
private double[, ,] grayData2D	Матрица преобразованных данных полутонового изображения.
private int width	Ширина изображения.
private int height	Высота изображения.
private int level	Уровень разложения вейвлет-преобразования.
private Wavelet wavelet	Тип вейвлета.

Таблица 4.5. Основные функции класса *WaveletImage*

Название функции	Описание
public WaveletImage(Image<Gray, byte> iSource, int iLevel, string iWavelet)	Преобразование полутонного изображения iSource по заданному типу вейвлета iWavelet и уровню разложения iLevel.
public Image<Gray, byte> Image()	Представление изображения после применения вейвлет-преобразования.
public Image<Gray, byte> GetSubLL()	Представление изображения низкочастотного субдиапазона после применения вейвлет-преобразования.
private static byte ValidateIntensity(double value)	Проверка значения интенсивности value.
private void HorizontalTransform2D(int iLevel)	Преобразование по горизонтали и заданному уровню разложения iLevel.
private void VerticalTransform2D(int iLevel)	Преобразование по вертикали и заданному уровню разложения iLevel.

Класс собственных объектов («*EigenObjects*»)

Класс *EigenObjects* предназначен для создания собственных известных объектов с использованием PCA. Класс включает в себя переменные, предназначенные для сохранения собственных изображений известных объектов, собственных значений этих собственных изображений, среднего изображения и названий известных объектов.

Класс имеет методы, которые позволяют вычислить собственные изображения, их среднее изображение и собственные значения, определить расстояния между входным изображением неизвестного объекта и изображениями известных объектов, найти известный объект, наиболее близкий к входному объекту. Структура класса *EigenObjects* представлена на рис. 4.4. Основные переменные и методы класса *EigenObjects* приведены в табл. 4.6 и 4.7.

+ EigenObjects
- eigenImages: Image<Gray, Single>[] - eigenValues: Matrix<float>[] - averageImage: Image<Gray, Single> - labels: string[]
- EigenObjects(Image<Gray, byte>[], string, MCvTermCriteria) - CalcEigenObjects(Image<Gray, byte>[], MCvTermCriteria, Image<Gray, Single>[], Image<Gray, Single>): void - GetEigenDecomposition(Image<Gray, Byte>, Image<Gray, Single>[], Image<Gray, Single>): float[] - GetEigenDistances(Image<Gray, Byte>): float[] - FindMostSimiliarObject(Image<Gray, Byte>, float, string): void + Recognize(Image<Gray, Byte>): string

Рис. 4.4. Структура класса *EigenObjects*

Таблица 4.6. Основные переменные класса *EigenObjects*

Название переменной	Описание
private Image<Gray, Single>[] eigenImages	Массив собственных изображений известных объектов.
private Matrix<float>[] eigenValues	Массив собственных значений указанных собственных изображений.
private Image<Gray, Single> averageImage	Среднее изображение указанных собственных изображений.
private String[] Labels	Массив названий известных объектов.

Таблица 4.7. Основные функции класса *EigenObjects*

Название функции	Описание
public EigenObjects(...)	Создание собственных известных объектов.
private void CalcEigenObjects(...)	Вычисление собственных изображений и их среднего изображения.
private float[] GetEigenDecomposite(...)	Вычисление собственных значений.
private float[] GetEigenDistances(Image<Gray, Byte> image)	Вычисление расстояний между входным изображением неизвестного объекта и изображениями известных объектов.
private void FindMostSimilarObject(...)	Определение известного объекта, на который наиболее похож входной неизвестный объект.
public String Recognize(Image<Gray, Byte> image)	Распознавание входного неизвестного объекта.

Класс управления собственными объектами («*EigenObjectsManager*»)

Класс *EigenObjectsManager* предназначен для управления собственными известными объектами. Класс включает в себя методы, предназначенные для сохранения собственных известных объектов в файле, и загрузки собственных известных объектов из файла. Структура класса *EigenObjectsManager* представлена на рис. 4.5. Основные функции класса *EigenObjectsManager* приведены в табл. 4.8.

+ <i>EigenObjectsManager</i>
+ createEigenObjects(Imge<Gray, Byte>[], string[], MCvTermCriteria): EigenObjects + Load(string): EigenObjects - Load(Stream): EigenObjects + Save(EigenObjects, string): void - Save(EigenObjects, Stream): void

Рис. 4.5. Структура класса *EigenObjectsManager*

Таблица 4.8. Основные функции класса *EigenObjectsManager*

Название функции	Описание
public EigenObjects createEigenObjects (...)	Создание собственных известных объектов.
public EigenObjects Load(string iFile)	Загрузка собственных известных объектов из файла.
EigenObjects Load(Stream iStream)	Загрузка собственных известных объектов из бинарного файла.
public void Save(EigenObjects iEigenObjects, string iFile)	Сохранение собственных известных объектов в файле.
void Save(EigenObjects iEigenObjects, Stream iStream)	Сохранение собственных известных объектов в бинарном файле.

Класс обработки изображений («*ImageProcessing*»)

Класс *ImageProcessing* предназначен для обработки изображений. Класс включает в себя методы, предназначенные для решения таких задач, как загрузка данных обучающей выборки и тестовой выборки из папки изображений, преобразование изображения в полутоновое, изменение размера изображения и извлечение признаков изображения. Структура

класса *ImageProcessing* представлена на рис. 4.6. Основные функции класса *ImageProcessing* приведены в табл. 4.9.

+ ImageProcessing
+ LoadData(string, Image<Gray, byte>[], string[]): void - ToGrayImage(Image<Bgr, byte>): Image<Gray, byte> - Resize(Size, Image<Gray, byte>): EigenObjects + FeatureExtract(Image<Gray, byte>): Image<Gray, byte> + FeatureExtract(Image<Gray, byte>[]): Image<Gray, byte>[] + Recognize(EigenObjects, Image<Gray, byte>): string

Рис. 4.6. Структура класса *ImageProcessing*

Таблица 4.9. Основные функции класса *ImageProcessing*

Название функции	Описание
public void LoadData(...)	Загрузка данных выборки в переменных oImages и oLabels из папки изображений iRootPath.
public Image<Gray, byte> ToGrayImage(Image<Bgr, byte> iImage)	Преобразование цветного изображения iImage в полутоновое.
public Image<Gray, byte> Resize(Size newSize, Image<Gray, byte> iImage)	Изменение размера изображения iImage до нового размера newSize.
public Image<Gray, byte> FeatureExtract(Image<Gray, byte> iImage)	Извлечение признаков изображения области объекта iImage.
public Image<Gray, byte>[] FeatureExtract(Image<Gray, byte>[] iImages)	Извлечение признаков изображений областей объектов iImages.
public string Recognize(EigenObjects iEigenObjects, Image<Gray, byte> iImage)	Распознавание изображения области объекта iImage с использованием собственных известных объектов EigenObjects.

б. Классы для распознавания множества лиц на видеопоследовательностях в режиме реального времени

Для реализации алгоритма распознавания множества лиц на видеопоследовательностях в режиме реального времени (раздел 2.4) используются классы раздела 4.1.1.а и создаются классы для обработки видеопоследовательности.

В классе **ImageProcessing** создается дополнительная функция, реализующая метод Виолы-Джонса. Также создаются классы **Database** и **VideoProcessing**, предназначенные для управления изображениями собственных известных лиц и обработки кадров видеопоследовательности.

Класс обработки изображений («ImageProcessing»)

Кроме методов, приведенных в табл. 4.9, класс **ImageProcessing** имеет дополнительную функцию, предназначенную для реализации метода обнаружения объектов Виолы-Джонса. Структура класса **ImageProcessing** представлена на рис. 4.7. Дополнительная функция класса **ImageProcessing** приведена в табл. 4.10.

+ ImageProcessing
+ LoadData(string, Image<Gray, byte>[], string[]): void - ToGrayImage(Image<Bgr, byte>): Image<Gray, byte> - Resize(Size, Image<Gray, byte>): Image<Gray, byte> + FeatureExtract(Image<Gray, byte>): Image<Gray, byte> + FeatureExtract(Image<Gray, byte>[]): Image<Gray, byte>[] + Recognize(EigenObjects, Image<Gray, byte>): string + DetectObjects(HaarCascade, Image<Bgr, byte>): Rectangle[]

Рис. 4.7. Структура класса *ImageProcessing*

Таблица 4.10. Дополнительная функция класса *ImageProcessing*

Название функция	Описание
public Rectangle[] DetectObjects (HaarCascade haarCascade, Image<Bgr, byte> iSource)	Обнаружение объектов на изображении iSource на основе использования метода Виолы-Джонса и Хаар-подобных признаков haarCascade.

Класс базы данных («Database»)

Класс **Database** предназначен для создания базы изображений и их соответствующих названий, использованных для создания собственных известных объектов с использованием PCA.

Класс включает в себя переменные, предназначенные для сохранения изображений известных объектов и их названий. В указанном классе содержатся методы, которые позволяют найти изображения объектов по их названию, добавить, удалить, заменить объекты и их изображения в базе.

Структура класса **Database** представлена на рис. 4.8. Основные переменные и методы класса **Database** приведены в табл. 4.11 и 4.12.

+ Database
- objectsImages: IDictionary<string, List<Image<Gray, byte>>>
+ Database(Image<Gray, byte>[], string[]) + SearchByName(string): List<Image<Gray, byte>> + Add(string, List<Image<Gray, byte>>): void + Delete(string, List<int>): void + Delete(string): void + Change(string, string)

Рис. 4.8. Структура класса *Database*

Таблица 4.11. Основные переменные класса *Database*

Название переменной	Описание
private IDictionary<string, List<Image<Gray, byte>>> objectsImages	Список изображений и названий объектов.

Таблица 4.12. Основные функции класса *Database*

Название функции	Описание
public Database(Image<Gray, Byte>[] images, String[] iLabels)	Создание базы изображений.
public List<Image<Gray, byte>> SearchByName(string iName)	Нахождение изображений объекта по заданному названию.
public void Add(string iName, List<Image<Gray, byte>> images)	Добавление изображения объекта.
public void Delete(string iName, List<int> imagesIndex)	Удаление изображения объекта.
public void Delete(string iName)	Удаление объекта.
public void Change(string iOldName, string iNewName)	Изменение названия объекта на новое название.

Класс управления базой данных («DatabaseManager»)

Класс **DatabaseManager** предназначен для управления базой изображений. Класс включает в себя методы, предназначенные для создания

базы изображений, сохранения базы изображений в файле и загрузки базы изображений из файла.

Класс также имеет функции, которые позволяют найти изображения объектов по их названию, добавить, удалить, заменить объекты и их изображения в базе. Структура класса ***DatabaseManager*** представлена на рис. 4.9. Основные переменные и методы класса ***DatabaseManager*** приведены в табл. 4.13.

+ DatabaseManager
+ createDataBase(Imge<Gray, Byte>[], string[]): Database + Load(string): Database - Load(Stream): Datatbase + Save(Database, string): void - Save(Database, Stream): void + SearchByName(Database, string): List<Image<Gray, byte>> + Add(Database, string, List<Image<Gray, byte>>): void + Delete(Database, string, List<int>): void + Delete(Database, string): void + Change(Database, string, string)

Рис. 4.9. Структура класса *DatabaseManager*

Таблица 4.13. Основные функции класса *DatabaseManager*

Название функции	Описание
public Database createDatabase(...)	Создание базы изображений объектов.
public Database Load(string iFile)	Загрузка базы изображений объектов из файла.
private Database Load(Stream iStream)	Загрузка базы изображений объектов из бинарного файла.
public void Save(Database iDatabase, string iFile)	Сохранение базы изображений объектов в файле.
private void Save(Database iDatabase, Stream iStream)	Сохранение базы изображений объектов в бинарном файле.
public List<Image<Gray, byte>> SearchByName(string iName)	Нахождение изображений объекта по заданному названию.
public void Add(string iName, List<Image<Gray, byte>> images)	Добавление изображения объекта.
public void Delete(string iName, List<int> imagesIndex)	Удаление изображения объекта.
public void Delete(string iName)	Удаление объекта.
public void Change(string iOldName, string iNewName)	Изменение названия объекта на новое название.

Класс обработки видео («*VideoProcessing*»)

Класс ***VideoProcessing*** предназначен для обработки видеопоследовательности. Класс включает в себя методы, предназначенные для обработки кадров видеопоследовательности, представления состояния и результатов распознавания на экране. Структура класса ***VideoProcessing*** представлена на рис. 4.10. Основные методы класса ***VideoProcessing*** приведены в табл. 4.14.

+ <i>VideoProcessing</i>
+ FramGrabber(Image<Bgr, byte>): void + ShowResultToFrame(Image<Bgr, byte>, Rectangle, string): void + ShowTextToFrame(Image<Bgr, byte>, string)

Рис. 4.10. Структура класса *VideoProcessing*

Таблица 4.14. Основные функции класса *VideoProcessing*

Название функции	Описание
<code>public void FramGrabber(Image<Bgr, byte iFrame>)</code>	Обработка кадров видеопоследовательности.
<code>public void ShowResultToFrame(...)</code>	Представление областей расположения лиц и результатов их распознавания на экране.
<code>public void ShowTextToFrame(Image<Bgr, byte iFrame, string iStatus>)</code>	Представление состояния процесса обнаружения и распознавания на экране.

с. Классы для распознавания жестов на видеопоследовательностях в режиме реального времени

Для реализации алгоритма распознавания жестов на видеопоследовательностях в режиме реального времени (раздел 2.5) используются классы раздела 4.1.1.а.

В классе ***ImageProcessing*** создаются дополнительные функции, предназначенные для реализации алгоритма отслеживания руки *CAMShift*.

Класс обработки изображений («*ImageProcessing*»)

Кроме функции, приведенных в табл. 4.9 и 4.10, класс *ImageProcessing* имеет дополнительные функции, предназначенные для реализации алгоритма отслеживания руки *CAMShift*. Структура класса *ImageProcessing* представлена на рис. 4.11. Дополнительные функции класса *ImageProcessing* приведены в табл. 4.15.

+ <i>ImageProcessing</i>
+ LoadData(string, Image<Gray, byte>[], string[]): void - ToGrayImage(Image<Bgr, byte>): Image<Gray, byte> - Resize(Size, Image<Gray, byte>): Image<Gray, byte> + FeatureExtract(Image<Gray, byte>): Image<Gray, byte> + FeatureExtract(Image<Gray, byte>[]): Image<Gray, byte>[] + Recognize(EigenObjects, Image<Gray, byte>): string + DetectObjects(HaarCascade, Image<Bgr, byte>): Rectangle[] + CamshiftInitialize(Imge<Bgr, byte>, Image<Gray, byte>, Image<Gray, byte>, DenseHistogram, Rectangle, Boolean): void + CamshiftUpdate(Imge<Bgr, byte>, Image<Gray, byte>, Image<Gray, byte>): void + CamshiftTrackObject(Image<Gray, byte>, Image<Gray, byte>, Image<Gray, byte>, DenseHistogram, Rectangle, Boolean): Rectangle

Рис. 4.11. Структура класса *ImageProcessing*

Таблица 4.15. Дополнительные функции класса *ImageProcessing*

Название функции	Описание
public void CamshiftInitialize(...)	Инициализация первоначальных данных для алгоритма <i>CAMShift</i> .
public void CamshiftUpdate(...)	Обновление данных алгоритма <i>CAMShift</i> при запросе нового кадра.
public Rectangle CamshiftTrackObject(...)	Слежение за объектами.

4.2 Программные средства для распознавания объектов на изображениях и видеопоследовательностях

В ходе реализации разработанных алгоритмов для распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях были созданы соответствующие программные средства. Программные реализации были построены с использованием библиотек *AForge.Net*, *Emgu CV* и *OpenCV*.

В связи с тем, что распознавание объектов в данной диссертационной работе заключается в распознавании лиц и жестов, были реализованы три программы. Первая реализация «*PatternRecognition*» представляет собой программу для распознавания лиц и жестов на статических изображениях. Эта программа позволяет производить автоматизированное распознавание на больших наборах статических изображений.

Вторая реализация «*FaceRecognition*» представляет собой программу для распознавания лиц и множества лиц на видеопоследовательностях в режиме реального времени. Эта программа позволяет производить автоматическое распознавание лиц в сложных сценах (полученных с видеоролика), на реальных практических задачах (изображениях, полученных с веб-камеры) и имеет специальный пользовательский режим.

Третья реализация «*GestureRecognition*» представляет собой программу для распознавания движущихся жестов на видеопоследовательности, полученной с веб-камеры.

4.2.1 Приложение для распознавания объектов на статических изображениях

Программа «*PatternRecognition*» предназначена для распознавания лиц и жестов на статических изображениях.

При запуске программы появляется главное окно формы (рис. 4.12).

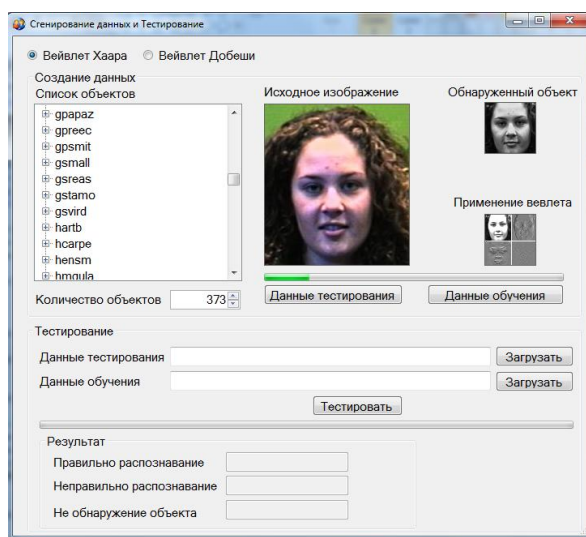


Рис. 4.12. Главное окно формы распознавания объектов

Для исследования разработанного алгоритма распознавания лиц и жестов на статических изображениях было реализовано приложение, позволяющее работать с большим количеством изображений. Оно позволяет осуществлять следующие операции.

1. Выбор типа вейвлет-преобразования Хаара или Добеши.
2. Процесс **«Создание данных»** состоит из следующих компонент: **«Список объектов»**, **«Исходное изображение»**, **«Обнаруженный объект»**, **«Применение вейвлета»**, **«Количество объектов»**, **«Данные тестирования»**, **«Данные обучения»**.
3. **«Тестирование»**: задание набора изображений для тестирования, задание набора изображений для обучения.
4. **«Результат»** распознавания объектов: процент правильного распознавания, процент неправильного распознавания, процент необнаруженного распознавания.

Для начала работы с программой необходимо создать данные, после этого загрузить папку с изображениями для тестирования и обучения. При нажатии на кнопку **«Тестировать»** производится запуск обработки.

Процесс **«Создание данных»** состоит из следующих этапов.

1. **«Список объектов»**: изображения каждого объекта находятся в папке **«Train»** (рис. 4.13). В каждой папке содержатся изображения соответствующей категории.

- Изображения лиц были взяты из БД: *Collection of Facial Images* [106], *ORL Face Databases* [95] и *Yale Face Database* [122].

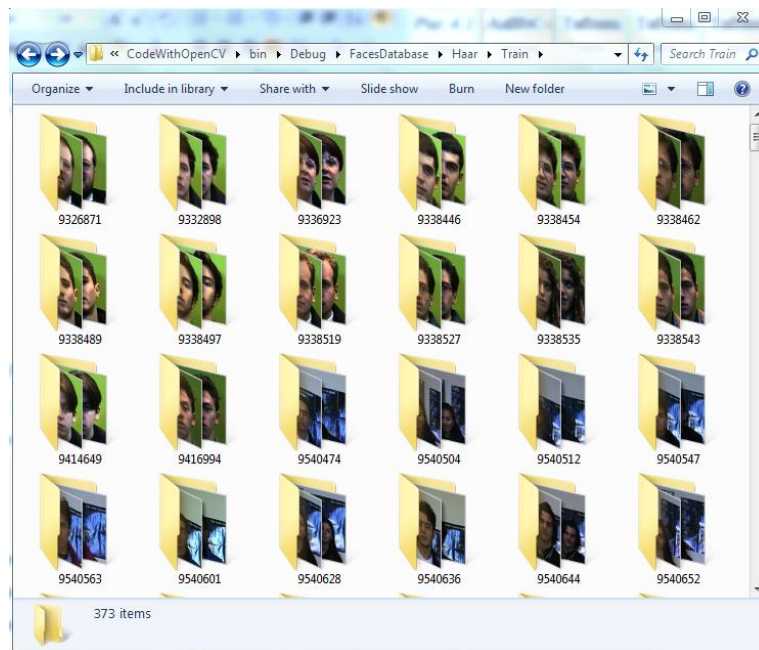


Рис. 4.13. Пример папки «Лица для обучения»

- Изображения жестов были взяты из БД *Cambridge Gesture Data Base* [76]. Эта база изображений жестов состоит из 5 различных частей, изображения в которых были получены при различных условиях освещенности. В данной работе, все жесты в БД делятся на 12 классов, рис. 4.14.

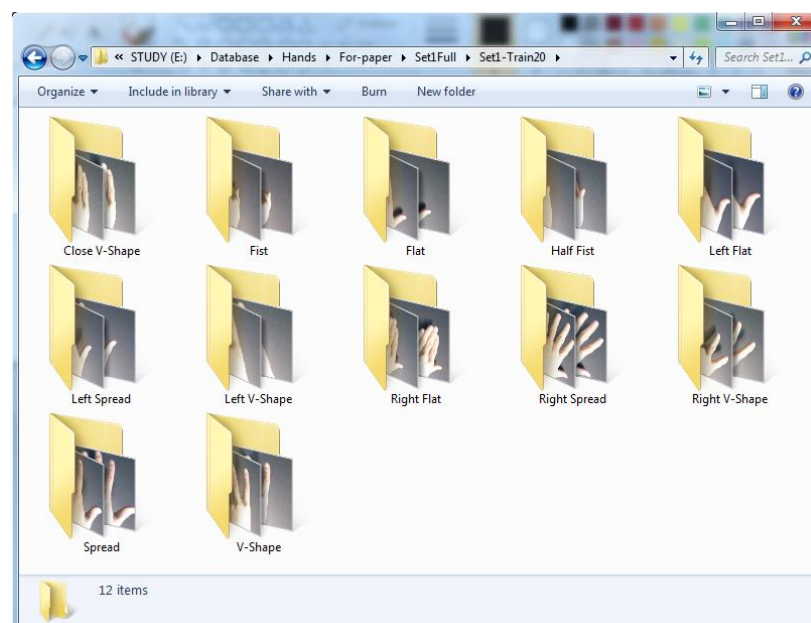


Рис. 4.14. Пример папки «Жесты для обучения»

2. «Исходное изображение», «Обнаруженный объект», «Применение вейвлета»: происходит показ изображений на экране.

3. **«Данные тестирования»:** при нажатии на эту кнопку запускается процесс создания данных для тестирования.

4. **«Данные обучения»:** при нажатии на эту кнопку запускается процесс создания данных для обучения.

После создания данных происходит загрузка соответствующих файлов. Результаты распознавания объектов появляются на экране через несколько минут в зависимости от количества изображений.

4.2.2 Приложение для распознавания лиц на видеопоследовательностях

Программа *«FaceRecognition»* предназначена для распознавания лиц и множества лиц на видеопоследовательностях в режиме реального времени.

При запуске программы появляется главное окно формы (рис. 4.15).

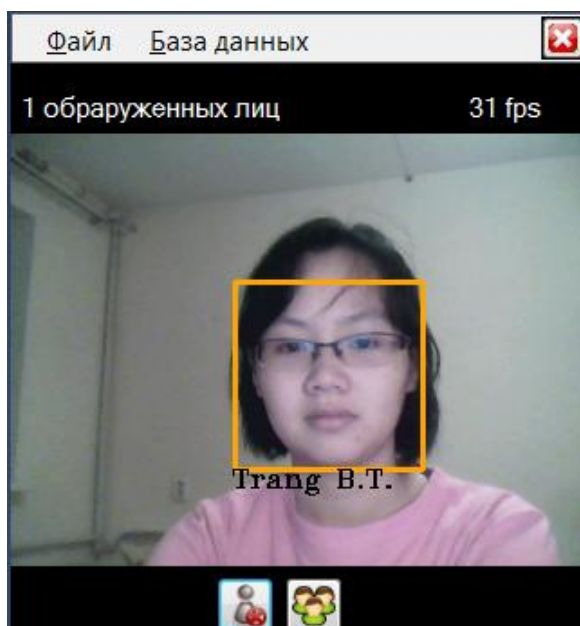


Рис. 4.15. Главное окно формы распознавания лиц на видеопоследовательности

Пользовательский интерфейс программы содержит 4 элемента.

1. **Главное меню («Файл», «База данных»)** расположено в верхней части окна. Оно предназначено для выбора функций и вызова вспомогательных окон.

2. Строка состояния находится внизу главного меню. Она показывает информацию о количестве обнаруженных объектов и скорости обработки (количество кадров в секунду).

3. **Рабочая область.** В рабочей области отображается видеопоследовательность, область обнаружения объектов, название обнаруженных объектов.

4. **Навигатор** расположен в нижней части окна. Он предназначен для более удобной работы с видеопоследовательностью. Навигатор содержит 2 элемента управления.

- Включение/отключение режима обнаружения объекта. При включении режима обнаружения объекта автоматически происходит запуск процесса распознавания.
- Добавление объектов в БД.

Меню «**Файл**» содержит следующие функции (рис. 4.16).

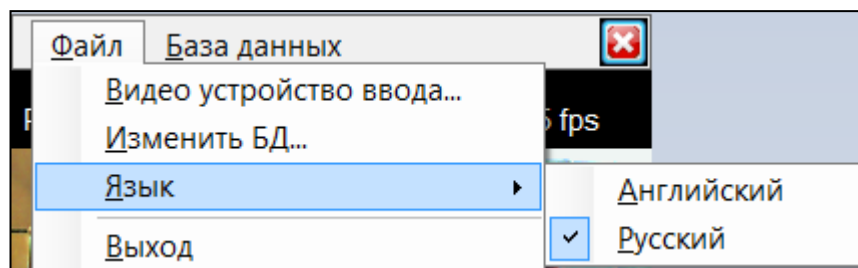


Рис. 4.16. Главное окно формы. Меню «Файл»

1. «**Видео устройство ввода...**». Открывает форму для устройства ввода видео.
2. «**Изменить БД**». Открывает форму для изменения БД.
3. «**Язык**». Позволяет выбрать язык пользовательского интерфейса: английский или русский.
4. «**Выход**». Позволяет осуществить выход из программы.

Меню «**База данных**» содержит следующие функции (рис. 4.17).

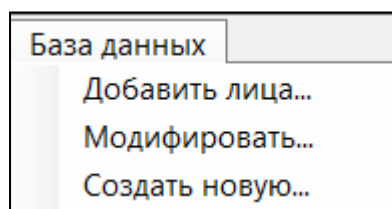


Рис. 4.17. Главное окно формы. Меню «База данных»

1. «Добавить лица...». Позволяет добавлять лица в БД (рис. 4.18). Функция навигатора «Поиск» обозначается 1, «Удаление» – 2, «Изменение» – 3 и «Добавление» – 4.

Сначала вводится имя в поле «Имя» на левой стороне, затем ставится галочка на наилучших изображениях лица или на всех изображениях. Завершается процесс при нажатии на кнопку «Добавление» – 4. Эти изображения будут сохраняться в БД.

Для поиска лица по имени вводится имя в поле «Имя» на правой стороне, результат поиска представлен рис. 4.19. Если необходимо посмотреть изображение, соответствующее найденному имени, то нужно нажать на название изображения лица в списке и тогда справа на экране появится изображение.

Для удаления изображения из БД необходимо войти в область «Лица в БД». Затем нужно найти название изображения в списке и нажать кнопку «Удаление» – 2 для удаления этого изображения.

Для изменения имени лица в БД, необходимо осуществить «Поиск» старого имени в поле «Имя» на левой стороне, затем нажать кнопку «Изменение» – 3 и ввести новое имя в поле «Имя» на правой стороне.

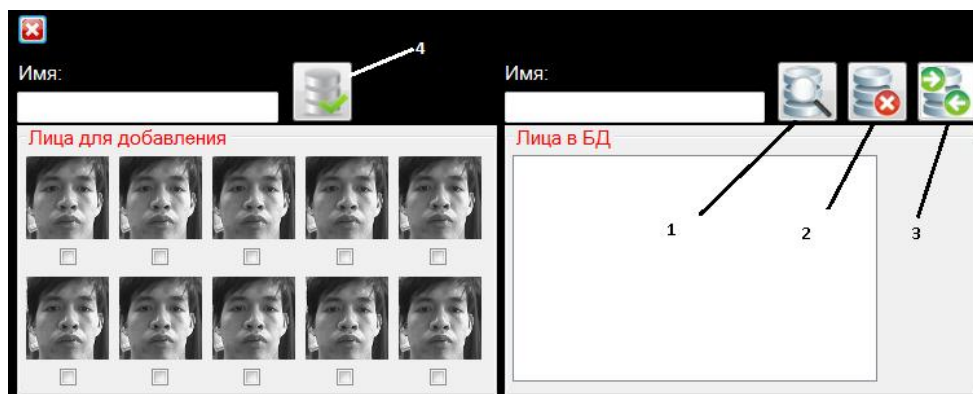


Рис. 4.18. Форма для добавления лиц в БД

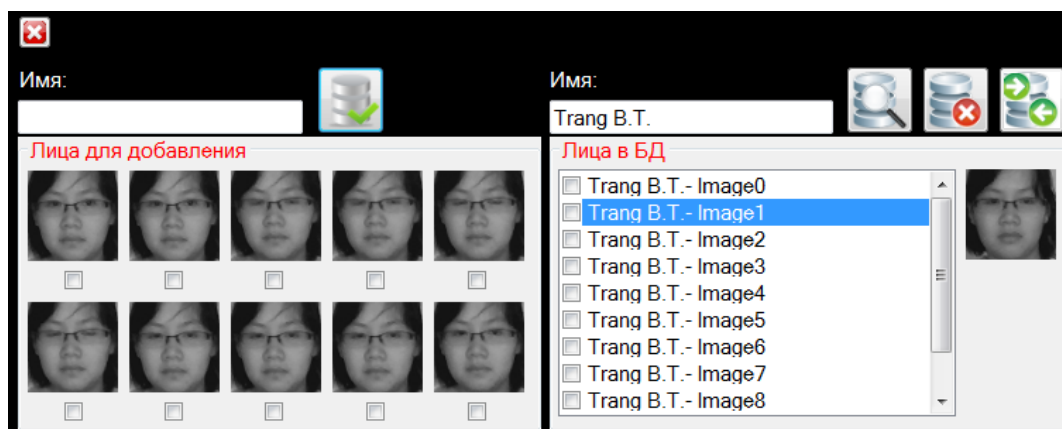


Рис. 4.19. Форма после добавления лиц в БД

2. «Модифицировать...». Позволяет модифицировать лица в БД.
3. «Создать новую...». Позволяет создать новую БД.

4.2.3 Приложение для распознавания жестов на видеопоследовательностях

Программа «*GestureRecognition*» предназначена для распознавания движущихся жестов на видеопоследовательностях в режиме реального времени.

При запуске программы появляется главное окно формы (рис. 4.20).



Рис. 4.20. Главное окно формы распознавания жестов на видеопоследовательностях

Алгоритм обучен на жестах 6 классов при отсутствии других объектов: ноль, один, два, три, четыре и пять (рис. 4.21). Эти классы описывают количество пальцев, представленных на изображениях.

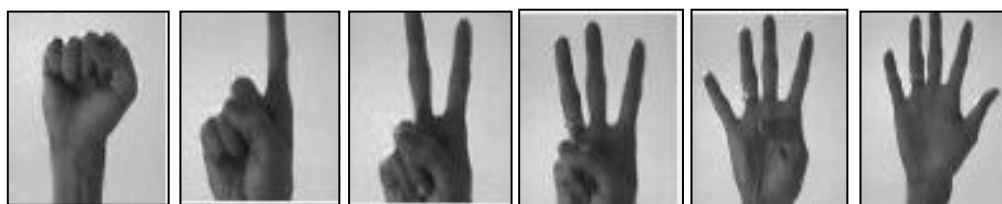


Рис. 4.21. Примеры 6 классов, использующихся при распознавании жестов на видеопоследовательностях

Сначала вводится название класса жеста в поле «**Тип жеста**» на левой стороне окна, затем отмечаются галочкой наилучшие изображения жеста или все изображения. Завершается этот процесс при нажатии на кнопку «**Добавление**» – 4. Эти изображения будут сохраняться в БД.

Для поиска типа жеста по названию необходимо ввести название в поле «**Тип жеста**» на правой стороне, результаты поиска представлены на рис. 4.22.

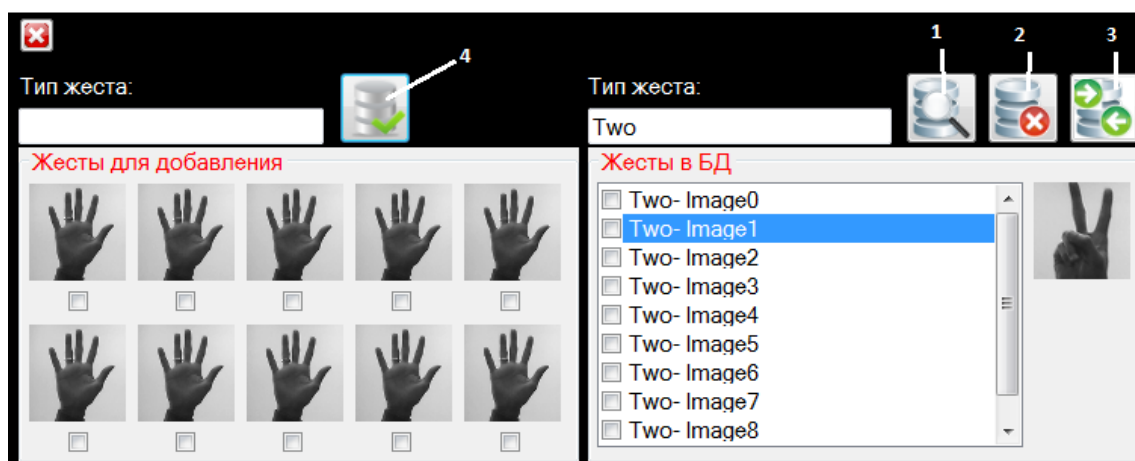


Рис. 4.22. Форма для добавления жестов в БД

Если нужно посмотреть изображения по найденному названию, то необходимо нажать на название изображения жеста в списке и тогда справа на экране появится изображение этого типа жеста.

Для удаления изображения из БД необходимо войти в область «**Жесты в БД**». Затем нужно найти название изображения в списке и нажать кнопку «**Удаление**» – 2 для удаления этого изображения жеста.

Для изменения названия типа жеста в БД необходимо осуществить «Поиск» старого названия в поле «Тип жеста» на левой стороне, затем нажать кнопку «Изменение» – 3 и ввести новое название в поле «Тип жеста» на правой стороне.

4.3 Основные результаты и выводы по главе 4

1. Приведено описание архитектуры программной реализации алгоритма на основе библиотек компьютерного зрения *OpenCV*, *AForge.Net* и *Emgu CV*.

2. Описана структура классов библиотеки программ, которая делится на 3 части: одна для распознавания лиц и жестов на статических изображениях, другая для распознавания лиц на видеопоследовательностях и третья для распознавания жестов на видеопоследовательностях в реальном режиме времени.

3. Разработаны программные средства, использующие полученную библиотеку для распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях в режиме реального времени.

ЗАКЛЮЧЕНИЕ

В диссертационной работе решена актуальная научно-техническая задача исследования и разработки алгоритмов распознавания объектов на статических изображениях и движущихся объектов на видеопоследовательностях.

В результате выполнения диссертационной работы были получены следующие основные научные и практические результаты и сделаны следующие выводы.

1. Предложен и разработан новый способ выделения признаков объектов на изображениях на основе совместного применения вейвлет-преобразований Хаара и Добеши.

2. Разработан оригинальный алгоритм распознавания лиц и жестов на статических изображениях на основе применения вейвлет-преобразований Хаара, Добеши и РСА.

3. Разработан новый алгоритм распознавания множества лиц на видеопоследовательности, основанный на совместном применении предложенного алгоритма распознавания лиц на изображениях и метода Виолы-Джонса.

4. Разработан оригинальный алгоритм распознавания жестов на видеопоследовательности в режиме реального времени, основанный на применении предложенного алгоритма распознавания жестов на изображениях и алгоритма *CAMShift*.

5. Созданы программные модули «*PatternRecognition*», «*FaceRecognition*» и «*GestureRecognition*», которые используются для распознавания объектов на изображениях и движущихся объектов на видеопоследовательностях в режиме реального времени.

ОБОЗНАЧЕНИЯ

$\vec{v}_i$	Собственные векторы матрицы $W^T W$.
μ_i	Собственные значения матрицы $W^T W$.
$\vec{I}$	Изображение лица.
λ_i	Собственные значения ковариационной матрицы.
$\psi(t)$	Материнский вейвлет.
$I(x,y)$	Значение точки (x,y) интегрального изображения.
$i(x,y)$	Значение интенсивности исходного изображения.
$I(x,y)$	Значение точки (x,y) интегрального изображения.
$i(x,y)$	Значение интенсивности исходного изображения.
M_{00}	Нулевой момент.
M_{10}	Первый момент.
x_c	Среднее смещение окна поиска.
R_{white}	Сумма значений яркостей точек, закрываемых светлой частью признака.
R_{black}	Сумма значений яркостей точек, закрываемых темной частью признака.
F_{Haar}	Хаар-подобный признак.
$f_j(z)$	Вычисленное значение признака.
z	Окно изображения размером 20×20 пикселей.
θ_j	Значение порога классификатора алгоритма бустинга.
$h_j(z)$	Классификатор алгоритма бустинга.

СПИСОК СОКРАЩЕНИЙ

CAMShift – Continuously Adaptive Mean Shift.

OpenCV – Open Source Computer Vision Library.

HSV – Hue Saturation Value.

IPP – Intel Performance Primitives.

PCA – Principal Component Analysis (Метод главных компонент).

LDA – Linear Discriminant Analysis (Линейный дискриминантный анализ).

SVM – Support Vector Machines (Метод опорных векторов).

ИНС – Искусственные нейронные сети.

ПО – Программное обеспечение.

ПС – Программная система.

пкс. – пиксель.

БД – База данных.

ПХ – Преобразование Хаара.

НВП – Непрерывное вейвлет-преобразование.

ДВП – Дискретное вейвлет-преобразование.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ И ЛИТЕРАТУРЫ

1. Алпатов, Б.А. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / Б.А. Алпатов, П.В. Бабаян, О.Е. Балашов, А.И. Степашкин. – М.: Радиотехника, 2008. – 176 с.: ил.
2. Архангельский, А.Я. Delphi 2006. Справочное пособие. Язык Delphi, классы, функции Win32 и .NET / А.Я. Архангельский. – 19 с.
3. Астафьева, Н.М. Вейвлет-анализ: основы теории и примеры применения / Н.М. Астафьева // Успехи физических наук. 1996. – Т. 166. – №11. – С. 1145–1170.
4. Белоусов, А.А. Применение генетических алгоритмов и вейвлет-преобразований для повышения качества изображений / А.А. Белоусов, В.Г. Спицын, Д.В. Сидоров // Известия Томского политехнического университета 2006. – Т. 309. – № 7. – С. 21–26.
5. Бовырин, А.В. Введение в разработку мультимедийных приложений с использованием библиотек OpenCV и IPP / А.В. Бовырин, А.В. Сысоев, А.Н. Половинкин, В.Д. Кустикова, В.Л. Ерухимов, В.Н. Писаревский, И.Б. Мееров, И.Д. Лысенков, Н.Ю. Золотых, П.Н. Дружков // Электронная книга.
6. Болотова, Ю.А. Алгоритмы обработки и анализа изображений иерархической временной сетью: Диссертация на соискание ученой степени кандидата технических наук / Ю.А. Болотова. – Томск, 2013.
7. Брадски, Г. Реализация некоторых алгоритмов обнаружения и слежения за объектом с использованием цветовых гистограмм / Г. Брадски, С.А. Молинов, С.А. Обломов, В.В. Черепенников // Научно-теоретический журнал «Искусственный интеллект» № 2, 2000.
8. Брилюк, Д. Распознавание человека по изображению лица и нейросетевые методы [Электронный ресурс] / Д. Брилюк, В. Старовойтов. – Режим доступа: <http://daily.sec.ru/publication.cfm?pid=4425>. Дата обращения: 10.10.2013.
9. Буй, Т.Т.Ч. Разложение цифровых изображений с помощью двумерного дискретного вейвлет-преобразования и быстрого преобразования / Т.Т.Ч.

- Буй, В.Г. Спицын // Известия Томского политехнического университета, 2011. – Т. 318. – № 5. – С. 73–76.
10. Буй, Т.Т.Ч. Алгоритмическое и программное обеспечение для классификации цифровых изображений с помощью вейвлет-преобразования Хаара и нейронных сетей / Т.Т.Ч. Буй, Н.Х. Фан, В.Г. Спицын // Известия Томского политехнического университета, 2011. – Т. 319. – № 5. – С. 103–106.
11. Буй, Т.Т.Ч. Распознавание лиц и жестов на основе применения вейвлет-преобразования и метода главных компонент / Т.Т.Ч. Буй, Н.Х. Фан, В.Г. Спицын // Нелинейный мир, 2012. – № 6. – Т. 10. – С. 371–379.
12. Буй, Т.Т.Ч. Анализ методов выделения краев на цифровых изображениях/ Т.Т.Ч. Буй, В.Г. Спицын // Журнал «Доклады ТУСУР». – Томск: Изд-во ТУСУР, декабрь 2010. – 2(22), часть 2. – С. 221–223.
13. Буй, Т.Т.Ч. Разложение цифровых изображений с помощью двумерного дискретного вейвлет-преобразования и быстрого преобразования / Т.Т.Ч. Буй, В.Г. Спицын // Известия Томского политехнического университета. – 2011. – Т. 318. – № 5. – С. 73–76.
14. Буй, Т.Т.Ч. Разложение цифровых изображений с помощью двумерного дискретного вейвлет-преобразования и быстрого преобразования Хаара / Т.Т.Ч. Буй, В.Г. Спицын // Проблемы информатики. – 2011. – № 2. – С. 11–15.
15. Вапник, В.Н. Теория распознавания образов / В.Н. Вапник, А.Я. Червоненкис. – М.: Наука, 1974. – 416 с.
16. Вежневек, В.П. Алгоритмы анализа изображения лица человека для построения интерфейса человек-компьютер: Автореферат диссертации на соискание ученой степени кандидата физико-математических наук / В.П. Вежневек. – Москва, 2004.
17. Волошин, Г.Я. Методы распознавания образов (конспект лекций). Автор: Г.Я. Волошин, редактор: А.А. Ильин.

18. Голубев, М.Н. Использование информации о цвете в алгоритме выделения лиц на базе бустинга / М.Н. Голубев // Нейроинформатика. – Москва, 2011. – Ч. 3. – С. 55–62.
19. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1072 с.
20. Горелик, А.Л. Методы распознавания / А.Л. Горелик, В.А. Скрипкин. – 4-е изд. – М.: Высшая школа, 1984, 2004. – 262 с.
21. Горелик, А.Л. Современное состояние проблемы распознавания. Некоторые аспекты / А.Л. Горелик, И.Б. Гуревич, В.А. Скрипкин. – М.: Радио и связь, 1985. – 162 с.
22. Гращенко, Л.А. Теоретические и практические основы человеко-компьютерного взаимодействия: базовые понятия человеко-компьютерных систем в информатике и информационной безопасности: Монография / Л.А. Гращенко, А.П. Фисун / Ред. А.П. Фисун. – Орел: ОГУ, 2004. – 169 с. – Деп. в ВИНТИ 15.10.2004 г. № 1624 – В2004.
23. Двойной, И.Р. Использование полярной цветовой модели для выделения лица на изображении / И.Р. Двойной, И.И. Сальников // Успехи современного естествознания. – 2012. – № 6. – С. 93–94.
24. Демьянович, Ю.К. Введение в теорию вэйвлетов. Курс лекций / Ю.К. Демьянович, В.А. Ходаковский. – Санкт-Петербург, 2007. – 49 с.
25. Желтов, С.Ю. Система распознавания и визуализации характерных черт человеческого лица в реальном времени на персональной ЭВМ с использованием web-камеры / С.Ю. Желтов, Ю.В. Визильтер, М.В. Ососков / // Материалы 12 Международной конференции по Компьютерной графике и машинному зрению. – Нижний Новгород, 2002. – С. 251–254.
26. Жизняков, А.Л. Теория и методы обработки многомасштабных последовательностей цифровых изображений в промышленных системах контроля качества: Автореферат диссертации на соискание ученой степени доктора технических наук / А.Л. Жизняков. – Москва, 2008.

27. Журавлёв, Ю.И. Распознавание образов и распознавание изображений / Ю.И. Журавлёв, И.Б. Гуревич // Распознавание, классификация, прогноз. – 1989. – Т. 2. – С. 5–73.
28. Журавлёв, Ю.И. Распознавание: математические методы. Программная система практические применения / Ю.И. Журавлёв. – М.: Фазис, 2006. – 147 с.
29. Загоруйко, Н.Г. Методы распознавания и их применение / Н.Г. Загоруйко. – М.: Советское радио, 1972. – 208 с.
30. Лукьяница, А.А. Цифровая обработка видеоизображений / А.А. Лукьяница, А.Г. Шишкин. – М.: Ай-Эс-Эс Пресс, 2009. – 511 с.
31. Местецкий, Л.М. Математические методы распознавания образов / Л.М. Местецкий. – М.: МГУ, ВМиК, 2002–2004. – 85 с.
32. Насонов, А.В. Применение метода суперразрешения для биометрических задач распознавания лиц в видеопотоке / А.В. Насонов, А.С. Крылов, О.С. Ушмаев // Системы высокой доступности, 2009. – № 1. – С. 26–34.
33. Некрасова, В.А. Введение в естественно-интуитивное взаимодействие с компьютером / В.А. Некрасова, К.А. Носов, О.А. Юфрякова, Ю.В. Березовская // Электронная книга.
34. Пахирка, А.И. Компьютерный метод локализации лиц на изображениях в сложных условиях освещения: Автореферат / А.И. Пахирка. – Красноярск, 2011.
35. Павлов, А.Н. Анализ структуры нестационарных, коротких и зашумленных сигналов на основе вейвлет-преобразования: Автореферат / А.Н. Павлов. – Саратов, 2009.
36. Прэтт, У. Цифровая обработка изображений / У. Прэтт. – М.: Мир, 1982.
37. Рыбак, Я.В. Применение метода анализа главных компонент в системах охранного видеонаблюдения для идентификации человека по изображению лица. Application of the method of analysis of the main components in security

video surveillance systems for human identification by face image (Y.V. Rybak, Bonch-Bruevich Saint-Petersburg State University of Telecommunications, Saint-Petersburg) / Я.В. Рыбак // 17-ая Международная научно-техническая конференция «Современное телевидение».

38. Сергеев, В.В. Применение методологии распознавания образов в задачах цифровой обработки изображений / В.В. Сергеев // Автометрия. – 1998. – № 2. – С. 63–67.

39. Сергей, Л.С. Курс лекций по предмету «Основы проектирования систем с искусственным интеллектом» / Л.С. Сергей. – г. Днепродзержинск, 1997–1999 г.

40. Симанков, В.С. Адаптивное управление сложными системами на основе распознавания образов / В.С. Симанков, Е.В. Луценко. – Краснодар: Техн. ун-т Кубан. гос. технол. ун-та, 1999. – 318 с.

41. Система распознавания лиц [Электронный ресурс]. – Режим доступа: <http://www.digit.ru/state/20120113/388557443.html>. Дата обращения: 07.09.2012.

42. Сойфер, В.А. Методы компьютерной обработки изображений / Сойфер В.А. – М.: Физматлит, 2003. – 784 с. (2-е издание)

43. Сойфер, В.А. Теоретические основы цифровой обработки изображений: Учебное пособие / Сойфер В.А., Сергеев В.В., Попов С.Б., Мясников В.В. – Самара: СГАУ, 2000. – 256 с.

44. Фан, Н.Х. Распознавание жестов на видеопоследовательности в режиме реального времени на основе применения метода Виолы-Джонса, алгоритма SAMShift, вейвлет-преобразования и метода главных компонент / Н.Х. Фан, Т.Т.Ч. Буй, В.Г. Спицын // Вестник Томского государственного университета. – Томск: ТГУ, 2013. – Т. 23. – № 2. – С. 102–111.

45. Форсайт, Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс. : Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 928 с.

46. Шерстобитов, А.И. Распознавание лиц на групповых фотографиях с использованием алгоритмов сегментации / А.И. Шерстобитов, В.П. Федосов,

- В.А. Приходченко, Д.В. Тимофеев // Известия ЮФУ: технические науки. – 2013. – № 5.
47. Шокуров, А.В. Оптимальное использование вейвлет-компонент / А.В. Шокуров, А.В. Михалев // Успехи математических наук. – 2007. – Т. 62. – № 4. – С. 171–172.
48. Ярышев, С.Н. Цифровые методы обработки видеoinформации и видеоаналитика: Учебное пособие / С.Н. Ярышев. – СПб.: СПбГУ ИТМО, 2011. – 83 с.
49. A Trainable System for Object Detection. Center of Biological and Computational Learning, Artificial Intelligence Laboratory, MIT, International Journal of Computer Vision, 2000. – 38(1). – P. 15–33.
50. Aforge [Электронный ресурс]. – Режим доступа: <http://code.google.com/p/aforge/>. Дата обращения: 17.09.2012.
51. Anuj, B. Image compression using modified fast Haar wavelet transform / B. Anuj, A. Rashid // World Applied Sciences Journal, 2009. – V. 7. – № 5. – P. 647–653.
52. Belhumeur, P.N. Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection / P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997. – Vol. 19. – P. 711–720.
53. Bowden, R. A linguistic feature vector for the visual interpretation of sign language / R. Bowden, D. Windridge, T. Kadir, A. Zisserman, M. Brady // In: ECCV, 2004. – P. 390–401.
54. Bradski, G.R. Computer Vision Face Tracking For Use in a Perceptual User Interface / G.R. Bradski // Intel Technology Journal. 1998, 2nd Quarter. – P. 1–15.
55. Bruce, V. Human Face Perception and Identification. Face Recognition: From Theory to Applications / V. Bruce, P.J.B. Hancock, A.M. Burton. H. Wechler, P.J. Phillips, F.F. Soulie, T. S. Huang (eds.) – Springer-Verlag, Berlin, 1998. – P. 51–72.

56. Chen, C.-F. Combination of PCA and Wavelet Transform for Face Recognition on 2.5D Images / C.-F. Chen, Y.-S. Tseng, C.-Y. Chen // Conf. of Image and Vision Computing. Palmerston North, November 2003. – P. 26–28.
57. Comaniciu, D. Mean Shift Analysis and Application / D. Comaniciu, P. Meer // In: Proc. IEEE International Conference on Computer Vision, Kerkyra, 1999. – P. 1197–1203.
58. Comaniciu, D. Real-Time Tracking of Non-rigid objects Using Mean Shift / D. Comaniciu, V. Ramesh, P. Meer // In: IEEE Computer Vision and Pattern Recognition, 2000. – Vol. 2. – P. 142–149.
59. Comaniciu, D. Mean Shift: A Robust Approach toward Feature Space Analysis / D. Comaniciu, P. Meer // In: IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002. – Vol. 24. – P. 603–619.
60. Comaniciu, D. Kernel-based Object Tracking / D. Comaniciu, V. Ramesh, P. Meer // In: IEEE Transactions Pattern Analysis and Machine Intelligence, 2003. – Vol. 25. – P. 564–577.
61. Dardas, N. Hand gesture recognition using bag-of-features and multi-class support vector machine / N. Dardas, Q. Chen, N. Georganas, E. Petriu // In Proc. 9th IEEE Int. Workshop HAVE, Phoenix, AZ, Oct. 16–17, 2010. – P. 1–5.
62. Daubechies, I. Ten lectures on wavelets / I. Daubechies. Philadelphia: S.I.A.M., 1992.
63. Eickeler, S. High quality face recognition in JPEG compressed images / S. Eickeler, S. Mueller, G. Rigoll // Proc. IEEE Int. Conf. Image Processing, 1999. – P. 672–676.
64. Emgu [Электронный ресурс]. – Режим доступа: http://www.emgu.com/wiki/index.php/Main_Page. Дата обращения: 17.09.2012.
65. Freund, Y. A Short Introduction to Boosting / Y. Freund, R.E. Schapire // Journal of Japanese Society for Artificial Intelligence, September 1999. – V.14. – № 5. – P. 771–780.

66. Freund, Y. A decision-theoretic generalization of online learning and an application to boosting / Y. Freund, R.E. Shapire // In: Computational Learning Theory: Eurocolt'95, Springer-Verlag, 1995. – P. 23–37.
67. Goldstein, A.J. Identification of Human Faces / A.J. Goldstein, L.D. Harmon, A.B. Lesk // In: Proceeding of the IEEE, 1971. – Vol. 59 – P. 748–760.
68. Gorodnichy, D.O. Video-based framework for face recognition in video / D.O. Gorodnichy // Second Workshop on Face Processing in Video (FPiV'05) in Proceedings of Second Canadian Conference on Computer and Robot Vision (CRV'05 Victoria, BC, Canada, 9–11 May, 2005. – P. 330–338.
69. Hallinan, P.L. Two- and Three-Dimensional Patterns of the Face / P.L. Hallinan, G.G. Gordon, A.L. Yuille, P. Giblin, D. Mumford // Natick: A.K. Peters Ltd. 1999. – 260 p.
70. Hutchison, D. Lecture Notes in Computer Science / D. Hutchison, T. Kanade, J. Kittler, J.M. Kleinberg, F. Mattern, J.C. Michell, M. Naor, O. Nierstrasz, C.P. Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M.Y. Vardi, G. Weikum. – Springer-Verlag, Berlin, 2011. – Vol. 7098. – 300 p.
71. Imran, Bajwa S. Feature Based Image Classification by using Principal Component Analysis / Bajwa S. Imran, M. Shahid Naweed, M. Nadim Asif, S. Irfan Hyder // ICGST-GVIP Journal, April 2009. –Vol. 9. – Issue II. – P. 11–17.
72. Jiang, Y. Towards optimal bag-of features for object categorization and semantic video retrieval / Y. Jiang, C. Ngo, J. Yang // In Proc. ACM Int. Conf. Image Video Retrieval, 2007. – P. 494–501.
73. Jin, J. Face recognition based on the uncorrelated discriminant transformation / J. Jin, J.-Y. Yang, Z.-S. Hu, Z. Lou // Pattern Recognition, 2001. – Vol. 34. – P. 1405–1416.
74. Joachims, T. Text Categorization with Support Vector Machines: Learning with Many Relevant Features / T. Joachims // Proceedings of the European Conference on Machine Learning, Springer, 1998.
75. Juell, P. A hierarchical neural network for human face detection / P. Juell, R. Marsh // Journal Pattern Recognition, 1996. – Vol. 29. – Issue 5. – P. 781–787.

76. Kim T.K., Wong S.F., Cipolla R. Cambrige Hand Gesture Data set [Электронный ресурс]. – Режим доступа: http://www.iis.ee.ic.ac.uk/~tkkim/ges_db.htm. Дата обращения 10.10.2013.
77. Kinect [Электронный ресурс]. – Режим доступа: <http://www.xbox.com/en-US/kinect/games>. Дата обращения: 24.09.2013.
78. Knight, B. The Role of Movement in Face Recognition / B. Knight, A. Johnston // Visual Cognition, 1997. – Vol. 4. – No. 3. – P. 265–274.
79. Kyungnam, K. Face Recognition using Principle Component Analysis / K. Kyungnam // Department of Computer Science University of Maryland, College Park, USA, June 2000.
80. Lam, K.-M. An analytic-to-holistic approach for face recognition based on a single frontal view / K.-M. Lam, H. Yan // IEEE Trans. Pattern Anal. Machine Intell., 1998. – Vol. 20. – P. 673–686.
81. Lazebnik, S. Beyond bags of features: Spatial pyramid matching of recognizing natural scene categories / S. Lazebnik, C. Schmid, J. Ponce // In Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2006. – P. 2169–2178.
82. Lepetit, V. Monocular Model-Based 3D Tracking of Rigid Objects: A Survey / V. Lepetit, P. Fua // Foundations and Trends In Comp. Graphics and Vision, 2005. – Vol. 1. – No. 1. – P. 1–89.
83. Li, Z. Bayesian Face Recognition using Support Vector Machines and Face Clustering / Z. Li, X. Tang // Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04) 2004.
84. Lin, S.-H. Face recognition/detection by probabilistic decision-based neural network / S.-H. Lin, S.-Y. Kung, L.-J. Lin // IEEE Trans. Neural Networks 8, 1997. – P.114–132.
85. Magesh, K.C. Gabor featured statistical modeling in face recognition with chaotic database / K.C. Magesh, G. Ragul, R. Thiyagarajan // International Journal of Emerging Technology and Advanced Engineering, January 2013. – Vol. 3. – Special Issue 1. – P.278–284.

86. Masoud, M. Face Recognition using Wavelet, PCA, and Neural Networks / M. Masoud, S. Kasaei // Proceeding of the First International Conference on Modeling, Simulation and Applied Optimization, Sharjah, U.A.E., 2005. – P. 1–3.
87. Meng, J.E. Face Recognition with Radial Basis Function (RBF) Neural Networks / J.E. Meng, W. Shiquian, L. Juwei, L.T. Hock // IEEE Transactions on Neural Networks, May 2002. – Vol. 13. – No. 3– P. 697–710.
88. Moghaddam, B. Probabilistic Visual Learning for Object Representation/ B. Moghaddam, A. Pentland // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997. – Vol. 19. – P. 696–710.
89. Mohammad, S.E.-B. Face recognition using multi-classifier/ S.E.-B. Mohammad // Applied Mathematical Sciences, 2012. – Vol. 6. – No. 45. – P. 2235–2244.
90. Mohammed, A. Face Recognition based on Haar Wavelet Transform and Principal Component Analysis via Levenberg-Marquardt Back propagation Neural Network / A. Mohammed, Z. Shaaban // European Journal of Scientific Research ISSN 1450-216X, 2010. – Vol. 42. – No. 1. – P. 25–31.
91. Mohammod, A.K. Face Recognition System Based on Principal Component Analysis (PCA) with Back Propagation Neural Networks (BPNN) / A.K. Mohammod, Md. Nasim Akhter, Shamim Ahmed, Md. Mahbub Alam // Canadian Journal on Image Processing and Computer Vision, 2011. – Vol. 2. – Issue 6. – P. 1–10.
92. Nasser, H.D. Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques / H.D. Nasser, N.D. Georganas // IEEE Transactions on Instrumentation and Measurement, November 2011. – Vol. 60. – No. 11. – P. 3592–3607.
93. Ole, H.J. Implementing the Viola-Jones face detection algorithm: Master's thesis / H.J. Ole. – DTU, 2008.
94. OpenCV [Электронный ресурс]. – Режим доступа: <http://old.kv.by/index2001513402.htm>. Дата обращения: 19.04.2012.

95. ORL Face Databases AT&T (Olivetti) Research Laboratories, Cambridge [Электронный ресурс]. – Режим доступа: http://www.uk.research.att.com/pub/data/orl_faces.zip. Дата обращения: 17.09.2012.
96. Papageorgiou, C. A general framework for object detection / C. Papageorgiou, M. Oren, T. Poggio // In: International Conference on Computer Vision, 1998.
97. Phiasai, T. Face recognition system with PCA and moment invariant method / T. Phiasai, S. Arunrungrusmi, K. Chamnongthai // IEEE Trans. – P. 165–168.
98. Rowley, H.A. Neural network-based face detection / H.A. Rowley, S. Baluja, T. Kanade. // IEEE Trans. Pattern Anal. Mach. Intell. 20, January 1998. – P. 23–38.
99. Schneiderman, H. A Statistical Method for 3D Object Detection Applied to Faces and Cars: Technical report / H. Schneiderman, T. Kanade. – Carnegie Mellon University, 2000.
100. Security solutions [Электронный ресурс]. – Режим доступа: http://securitysolutions.com/mag/security_looking_faces_super/. Дата обращения 26.08.2012.
101. Senthilkumaran, N.A. Study on Edge Detection Methods for Image Segmentation / N.A. Senthilkumaran, R. Rajesh // Proceedings of the International Conference on Mathematics and Computer Science (ICMCS-2009). – 2009. – Vol. 1. – P. 255–259.
102. Shumeet, B. Neural Network-Based Face Detection: Technical report / Shumeet B., H.A. Rowley, T. Kanade. – 1998.
103. Shweta, K.Y. Artificial neural network approach for hand gesture recognition / K.Y. Shweta, P.K. Bharne // International Journal of Engineering Science and Technology (IJEST), April 2011. – Vol. 3. – No. 4. – P. 2603–2608.
104. Sirovich, L. A Low-Dimensional Procedure for the Characterization of Human Faces / L. Sirovich, M. Kirby // In: J. Optical Soc. Am. A, 1987. – Vol. 4. – No. 3. – P. 519–524.

105. Sochman, J. AdaBoost with totally corrective updates for fast face detection [Электронный ресурс]. / J. Sochman, J. Matas. – Center for Machine Perception, Czech Technical University, Prague, 2004. – Режим доступа: <http://cmp.felk.cvut.cz/~sochmj1/>. Дата обращения: 10.10.2013.
106. Spacek L. Description of the Collection of Facial Images [Электронный ресурс]. / L. Spacek. – 2008. – Режим доступа: <http://cswwww.essex.ac.uk/mv/allfaces/index.html>. Дата обращения: 10.10.2013.
107. Stergiopoulou, E. Hand gesture recognition using a neural network shape fitting technique / E. Stergiopoulou, N. Papamarkos // Engineering Applications of Artificial Intelligence 22, 2009. – P.1141–1158.
108. Steinberg, I. Hand Gesture Recognition in Images and Video: CCIT Report #763 / I. Steinberg, Tomer M. London, Dotan Di Castro // Department of Electrical Engineering, Technion-Israel Institute of Technology, March 2010, 21 p.
109. Su, Y. Multiple Fisher Classifiers Combination for Face Recognition based on Grouping AdaBoosted Gabor Features / Y. Su, S. Shan, B. Cao, X. Chen, W. Gao // In: Proceedings of British Machine Vision Conference 2005, Oxford, UK.
110. Kim, T.-K. Gesture recognition Under Small Sample Size / T.-K. Kim, R. Cipolla // In: ACCV 2007. – Part 1. – LNCS 4843. – P. 335–344.
111. TinHninnHninn, M. Real-Time Hand Tracking and Gesture Recognition System Using Neural Networks / M. TinHninnHninn // World Academy of Science, Engineering and Technology 50, 2009. – P. 466–470.
112. Tolba, A.S. Combined classifiers for invariant face recognition / A.S. Tolba, A.N. Abu-Rezq // Proc. Int. Conf. Inform. Intell. Sys., 1999. – P. 350–359.
113. Turk, M.A. Face Recognition Using Eigenfaces / M.A. Turk, A.P. Pentland // In: Proc. IEEE, 1991. – P. 586–591.
114. Vapnik, V. Support-vector networks / V. Vapnik, C. Cortes // Machine Learning, 1995. – Vol. 20. – Issue 3. – P. 273–297.
115. Vasin, Yu.G. Invariant methods of binary image recognition / Yu.G. Vasin, L.I. Lebedev // Pattern Recognition and Image Analysis, 1998. – V. 8. – No. 2. – P. 172.

116. Viola, P. Fast and Robust Classification Using Symmetric AdaBoost and a Detector Cascade / P. Viola, M. Jones. // In: Proc. Neural Information Processing Systems, Vancouver, 2001. – P. 1311–1318.
117. Viola, P. Rapid object detection using a boosted cascade of simple features / P. Viola, M.J. Jones // IEEE Conf. On Computer Vision and Pattern Recognition, Kauai, Hawaii, USA, 2001. – V. 1. – P. 511–518.
118. Viola, P. Robust real-time face detection / P. Viola, M.J. Jones // International Journal of Computer Vision, 2004. – V. 57. – No. 2. – P. 137–154.
119. Virginia, E.-D. Biometric identification system using a radial basis network / E.-D. Virginia // Proc 34-th Annu. IEEE Int. Carnahan Conf. Security Technol., 2000. – P. 47–51.
120. Wii [Электронный ресурс]. – Режим доступа: <http://wii.com/>. Дата обращения: 14.05.2013.
121. Wong, S. Real-time interpretation of hand motions using a sparse Bayesian classifier on motion gradient orientation images / S. Wong, R. Cipolla / In: Proc. of BMVC 2005. – P. 379–388.
122. Yale Face Database [Электронный ресурс]. – Режим доступа: <http://vision.ucsd.edu/content/yale-face-database>. Дата обращения: 18.04.2013.
123. Yang, M.-H. Face Detection Using Multimodal Density Models / M.-H. Yang, N. Ahuja, D. Kriegman // Computer Vision and Image Understanding (CVIU), 2001. – Vol. 84. – No. 2. – P. 264–284.
124. Zhang, B.-L. Face recognition by wavelet domain associative memory / B.-L. Zhang, Y. Guo // Proc. Int. Synt. Intell. Multimedia, Video, Speech Processing, 2001. – P. 481–485.

СПИСОК ПУБЛИКАЦИЙ АВТОРА

Статьи в журналах, которые включены в перечень российских рецензируемых научных журналов и изданий для опубликования основных научных результатов диссертаций:

1. Буй, Т.Т.Ч. Анализ методов выделения краев на цифровых изображениях / Т.Т.Ч. Буй, В.Г. Спицын // Доклады Томского государственного университета систем управления и радиоэлектроники. – Томск: Изд-во ТУСУР, 2010. – № 2 (22). – Ч. 2. – С. 221–223. – 0,21 / 0,11 п.л.
2. Буй, Т.Т.Ч. Разложение цифровых изображений с помощью двумерного дискретного вейвлет-преобразования Хаара / Т.Т.Ч. Буй, В.Г. Спицын // Известия Томского политехнического университета. – Томск: Изд-во ТПУ, 2011. – Т. 318. – № 5. – С. 73–76. – 0,3 / 0,15 п.л.
3. Буй, Т.Т.Ч. Алгоритмическое и программное обеспечение для классификации цифровых изображений с помощью вейвлет-преобразования Хаара и нейронных сетей / Т.Т.Ч. Буй, Н.Х. Фан, В.Г. Спицын // Известия Томского политехнического университета. – Томск: Изд-во ТПУ, 2011. – Т. 319. – № 5. – С. 103–106. – 0,37 / 0,15 п.л.
4. Буй, Т.Т.Ч. Распознавание лиц на основе применения метода Виолы-Джонса, вейвлет-преобразования и метода главных компонент / Т.Т.Ч. Буй, Н.Х. Фан, В.Г. Спицын // Известия Томского политехнического университета. – Томск: Изд-во ТПУ, 2011. – Т. 320. – № 5. – С. 54–59. – 0,45 / 0,2 п.л.
5. Буй, Т.Т.Ч. Распознавание лиц и жестов на основе применения вейвлет-преобразования и метода главных компонент / Т.Т.Ч. Буй, Н.Х. Фан, В.Г. Спицын // Нелинейный мир. – Москва: Радиотехника, 2012. – Т. 10. – № 6. – С. 371–379. – 0,79 / 0,3 п.л.
6. Фан, Н.Х. Распознавание жестов на видеопоследовательности в режиме реального времени на основе применения метода Виолы-Джонса, алгоритма SAMShift, вейвлет-преобразования и метода главных компонент / Н.Х. Фан, Т.Т.Ч. Буй, В.Г. Спицын // Вестник ТГУ. Управление,

вычислительная техника и информатика. – Томск: ООО «Издательство научно-технической литературы», 2013. – Т. 23. – № 2. – С. 102–111. – 0,65 / 0,2 п.л.

Публикации в других научных изданиях:

7. Буй, Т.Т.Ч. Проблема подавления шума на изображениях и различные подходы к ее решению на основе применения вейвлет-преобразования / Т.Т.Ч. Буй, В.Г. Спицын // Сборник трудов VIII Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии». – Томск: Изд-во СПБ Графикс, 2010. – Ч. 2. – С. 225–226. – 0,19 / 0,1 п.л.

8. Буй, Т.Т.Ч. Подавление шума на изображениях и различные подходы к ее решению / Т.Т.Ч. Буй, В.Г. Спицын // Сборник трудов XVI Международной научно-практической конференции студентов, аспирантов и молодых ученых «Современные техника и технологии». – Томск: Изд-во ТПУ, 2010. – Т. 2. – С. 296–298. – 0,18 / 0,1 п.л.

9. Буй, Т.Т.Ч. Подавление шума на изображениях / Т.Т.Ч. Буй, В.Г. Спицын // Сборник докладов III Всероссийской научно-практической конференции «Научная инициатива иностранных студентов и аспирантов российских вузов». – Томск: Изд-во ТПУ, 2010. – С. 34–38. – 0,17 / 0,1 п.л.

10. Буй, Т.Т.Ч. Подавление шумов и реконструкция изображений на основе применения ядра регрессии / Т.Т.Ч. Буй, Н.Х. Фан // Сборник трудов XVII Международной научно-практической конференции студентов, аспирантов и молодых ученых «Современные техника и технологии». – Томск: Изд-во ТПУ, 2011. – Т. 2. – С. 299–300. – 0,26 / 0,13 п.л.

11. Буй, Т.Т.Ч. Разложение изображений с помощью двумерного дискретного вейвлет-преобразования и быстрого преобразования Хаара / Т.Т.Ч. Буй, В.Г. Спицын // Сборник трудов IX Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии». – Томск: Изд-во СПБ Графикс, 2011. – Ч. 1. – С. 198–199. – 0,2 / 0,1 п.л.

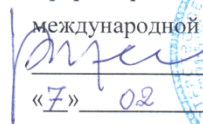
12. Буй, Т.Т.Ч. Разложение изображений с помощью двумерного дискретного вейвлет-преобразования и быстрого преобразования Хаара / Т.Т.Ч. Буй, В.Г. Спицын // Проблемы информатики. – Новосибирск: НГТУ, 2011. – № 2(10). – С. 4–10. – 0,29 / 0,2 п.л.

13. Буй, Т.Т.Ч. Классификация изображений на основе применения цветовой информации, вейвлет-преобразования Хаара и многослойной нейронной сети / Т.Т.Ч. Буй, Н.Х. Фан, В.Г. Спицын // Проблемы информатики. – Новосибирск: НГТУ, 2011. – Спецвыпуск. – С. 81–86. – 0,38 / 0,15 п.л.

14. Буй, Т.Т.Ч. Способ классификации изображений на основе применения вейвлет-преобразования Хаара и нейронных сетей / Т.Т.Ч. Буй, Н.Х. Фан, В.Г. Спицын // Материалы XIX Всероссийского семинара «Нейроинформатика, ее приложения и анализ данных». – Красноярск: Изд-во Сиб. федер. ун-та, 2011. – С. 159–164. – 0,31 / 0,12 п.л.

15. Bui, T.T.T. Face and Hand Gesture Recognition based on Wavelet Transforms and Principal Component Analysis / T.T.T. Bui, N.H. Phan, V.G. Spitsyn // 7th International Forum on Strategic Technology IFOST: Proceedings of IFOST. – Tomsk: TPU Press, 2012. – V. 1. – P. 588–591. – 0,4 / 0,15 п.л.

ПРИЛОЖЕНИЕ

УТВЕРЖДАЮ
Проректор по образовательной и
международной деятельности ТПУ

А.А. Чучалин
«7» 02 2014 г.


АКТ ВНЕДРЕНИЯ

результатов диссертационной работы Буй Тхи Тху Чанг
«Алгоритмы распознавания лиц и жестов на основе вейвлет-преобразований и
метода главных компонент»

Данный документ удостоверяет успешное применение в учебном процессе Института кибернетики ТПУ результатов диссертационной работы Буй Тхи Тху Чанг.

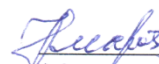
Начиная с 2012 г. студенты гр. 8BM11 и 8BM2A, ИК ТПУ, обучающиеся по направлению 230100 «Информатика и вычислительная техника» магистерской программы «Компьютерный анализ и интерпретация данных», при изучении спецкурса «Методы интеллектуальной обработки и анализа изображений» выполняют подготовленные Буй Тхи Тху Чанг лабораторные работы. Лабораторные работы содержат теоретическую часть, описывающую применяемые для обработки изображений методы и алгоритмы, а также задания для выполнения лабораторных работ. В лабораторных работах рассматривается применение вейвлет-преобразований Хаара и Добеши для разложения цифровых изображений, проводится изучение методов выделения краев цифровых изображений (Собеля, Лапласа, Робертса, Канни, Превитта и математической морфологии), использованных при подготовке материалов диссертации.

Изданы соответствующие методические указания по проведению лабораторных работ: Ю.А. Болотова, Т.Т.Ч. Буй, К.А. Кермани, В.Г. Спицын, Н.Х. Фан. Алгоритмы интеллектуальной обработки цифровых изображений. Часть 1: Методические указания к лабораторным работам по курсу «Методы интеллектуальной обработки и анализа изображений». Томск: Издательство ТПУ, 2012. – 58 с.

Директор ИК ТПУ,
д.т.н.


А.А. Захарова
«7» 02 2014 г.

Зав. кафедрой ВТ
профессор, д.т.н.


Н.Г. Марков
«7» 02 2014 г.