

ГОСУДАРСТВЕННОЕ НАУЧНОЕ УЧРЕЖДЕНИЕ
«ИНСТИТУТ ТЕХНИЧЕСКОЙ КИБЕРНЕТИКИ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ»

На правах рукописи

УДК 681.327.12.001.362

САМАЛЬ Дмитрий Иванович 

**АЛГОРИТМЫ ИДЕНТИФИКАЦИИ ЧЕЛОВЕКА ПО ФОТОПОРТРЕТУ
НА ОСНОВЕ ГЕОМЕТРИЧЕСКИХ ПРЕОБРАЗОВАНИЙ**

Специальность 05.13.01 – Системный анализ, управление и обработка информации

Диссертация на соискание учёной степени
кандидата технических наук

Научные руководители:

д.т.н., профессор АБЛАМЕЙКО С.В.
д.т.н. СТАРОВОЙТОВ В.В.

МИНСК 2002

*Посвящается моим родителям –
САМАЛЮ ИВАНУ МИХАЙЛОВИЧУ И
МУЛЁВОЙ НИНЕ ВЛАДИМИРОВНЕ*

ВВЕДЕНИЕ	4
ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ	8
Актуальность	8
Связь работы с крупными научными программами, темами	8
Цель и задачи исследования.....	9
Методы исследования.....	9
Научная новизна полученных результатов.....	10
Практическая значимость полученных результатов	10
Основные положения диссертации, выносимые на защиту	11
Личный вклад соискателя.....	11
Апробация результатов диссертации	11
Опубликованность результатов	12
Структура и объём диссертации	12
ГЛАВА 1 МЕТОДЫ АНАЛИЗА ФОТОПОРТРЕТОВ	13
1.1 КЛАССИФИКАЦИЯ ЗАДАЧ	13
1.1.1 Контроль удостоверений личности	13
1.1.2 Автоматический контроль доступа	14
1.1.3 Поиск фотопортретов заданного человека в базе данных	16
1.2 Подходы к распознаванию по фотопортретам	17
1.2.1 Антропометрический метод распознавания	17
1.2.2 Метод главных компонент	23
1.2.3 Статистическое распознавание изображений.....	27
1.2.4 Сопоставление изображения с эталоном	28
1.2.5 Распознавание с помощью нейронных сетей	29
1.2.6 Анализ оптических потоков изображений.....	32
1.2.7 Распознавание путём эластичного сопоставления графов.....	33
1.2.8 Применение фильтров Гabora для распознавания.....	36
1.3 АППАРАТ, ПРИМЕНЯЕМЫЙ ПРИ НОРМАЛИЗАЦИИ ФОТОПОРТРЕТОВ	38
1.3.1 Методы поиска области лица на фотопортрете	38
1.3.2 Методы поиска координат центров зрачков на фотопортрете	40
1.3.3 Определение условий получения изображения - ракурса съёмки.....	42
1.4 АНАЛИЗ СИСТЕМ ИДЕНТИФИКАЦИИ ЧЕЛОВЕКА ПО ФОТОПОРТРЕТУ	43
1.4.1 Достоинства и недостатки перечисленных методов.....	46
1.4.2 Выбор подхода для решения задач контроля удостоверений личности и поиска фотопортретов заданного человека в базе данных	47
1.5 Выводы	47
ГЛАВА 2 ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ИСХОДНЫХ ИЗОБРАЖЕНИЙ И ВЫДЕЛЕНИЕ ПРИЗНАКОВ	49
2.1 ТРЕБОВАНИЯ К ИСХОДНЫМ ИЗОБРАЖЕНИЯМ.....	49
2.1.1 Размер лица на изображении.....	49
2.1.2 Ракурс	50
2.1.3 Освещённость	50
2.1.4 Яркостные характеристики цифрового изображения.....	50
2.1.5 Эмоциональное выражение лица.....	51
2.2 ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ИСХОДНЫХ ИЗОБРАЖЕНИЙ	51
2.2.1 Алгоритм поиска области лица на фотопортрете	53

2.2.2 Алгоритм поиска координат центров зрачков.....	56
2.2.3 Выбор алгоритма поворота цифровых изображений	61
2.2.4 Масштабирование и кадрирование исходных изображений	69
2.2.5 Фильтрация и улучшение яркостных характеристик исходных изображений	72
2.3 АНАЛИЗ АНТРОПОМЕТРИЧЕСКИХ ТОЧЕК ЛИЦА	76
2.3.1 Антропометрические признаки в криминалистике	76
2.3.2 Отбор антропометрических точек лица для автоматического распознавания	78
2.4 ПРИМЕНЕНИЕ ФИЛЬТРОВ ГАБОРА ДЛЯ ОБНАРУЖЕНИЯ АНТРОПОМЕТРИЧЕСКИХ ТОЧЕК ЛИЦА	82
2.4.1 Исходная гипотеза.....	82
2.4.2 Методика выполнения экспериментов и условные обозначения	84
2.4.3 Результаты исследований	88
2.5 Выводы	92

ГЛАВА 3 АЛГОРИТМЫ ИДЕНТИФИКАЦИИ ЧЕЛОВЕКА И СРАВНЕНИЯ ФОТОПОРТРЕТОВ.....94

3.1 Поиск фотопортретов в базе данных	94
3.1.1 Заполнение базы данных	95
3.1.2 Первичный отбор портретов	97
3.1.3 Окончательный отбор портретов.....	100
3.2 Формирование признакового пространства	104
3.2.1 Оценка признаков на основе статистических данных	106
3.2.2 Экспериментальная проверка подхода	109
3.3 Сравнение фотопортретов методом эластичных экспоненциальных деформаций	114
3.4 Выводы	129

ГЛАВА 4 ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ КОМПЛЕКСА АЛГОРИТМОВ ИДЕНТИФИКАЦИИ ЧЕЛОВЕКА ПО ФОТОПОРТРЕТУ131

4.1 Экспериментальный комплекс программ	131
4.2 Программная реализация комплекса алгоритмов в виде динамически подключаемой библиотеки	134
4.3 Экспериментальная проверка разработанных алгоритмов и сравнение с аналогами	139
4.4 Использование результатов диссертационной работы	143
4.5 Выводы	146

ЗАКЛЮЧЕНИЕ147

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....149

ПРИЛОЖЕНИЕ 1162

Дополнительные иллюстрации к главам 2, 3	162
---	------------

ПРИЛОЖЕНИЕ 2167

Документы об использовании результатов диссертационной работы	167
--	------------

ВВЕДЕНИЕ

Постоянное увеличение объёмов обрабатываемой информации приводит к тому, что всё большее её количество передаётся только электронным путём (т.е. без получения копии на твёрдом носителе). Увеличивается количество сделок, совершаемых с помощью электронных средств связи, – от покупок в Интернет-магазинах и аукционах до игр на биржах. Возникла потребность введения так называемой «электронной подписи». В некоторых странах, в том числе и в Республике Беларусь, принятые законы, подтверждающие юридическую силу электронной подписи. Однако проблема надёжного распознавания пользователя ресурсов с ограниченным доступом (в том числе и компьютерных сетей) остаётся открытой. Отчасти поэтому в области обработки информации в последние годы наблюдается бум разработки систем идентификации человека по биометрическим признакам.

Под термином «биометрический признак человека» подразумевается некоторая присущая конкретному индивиду автоматически измеряемая характеристика. Биометрические признаки делятся на два класса – *физические* и *поведенческие*. К первому классу относятся характеристики, выделяемые на основе отпечатков пальцев и губ, геометрии кисти, изображений радужной оболочки глаза, лица, ушей и иные. Ко вторым, – манера работы на клавиатуре компьютера, динамика воспроизведения подписи или рукописного текста, артикуляция губ и т.п [15].

По прогнозам фирмы Frost & Sullivan Inc., занимающейся маркетинговыми исследованиями, к концу 2002 г. объём мирового рынка устройств и систем, измеряющих либо использующих биометрические признаки для решения различных задач, возрастёт до USD \$4 млрд. [70]. По данным Международной ассоциации по компьютерной безопасности, ожидается, что идентификация человека по изображению его лица будет наиболее используемой биометрической технологией, поскольку она может быть достаточно точной, быстрой, бесконтактной, а также прогнозируется, что мировой рынок биометрических систем достигнет оборота более чем \$2,3 млрд. к концу 2002 г. с ежегодным приростом в 31%.

Одной из причин повышенного внимания к биометрическим технологиям является существование обширного круга коммерческих и социальных приложений, где возможные решения названной проблемы будут восприняты весьма успешно. Так, например, идентификация людей по лицам может применяться в системах контроля удостоверений личности (паспортов, водительских прав, иммиграционных карт), информационной

безопасности (доступ к ЭВМ и отдельным программам, базам данных, криптографическим приложениям, медицинским сведениям, Интернету, системам электронной торговли и т.д.), наблюдения и расследования криминальных событий (автоматическое наблюдение за подозреваемыми, розыск людей, активное видеонаблюдение при чрезвычайных происшествиях и т.д.), а также в банковской сфере (банкоматах, системах удалённого управления счётом) [15, 66, 88, 120, 157, 173, 182].

Из всех биометрических подходов, таких, как распознавание по отпечаткам пальцев, губ, геометрии кисти, изображению радужной оболочки глаза, голосу и т.п., идентификация людей по изображениям лиц наиболее привычна, так как основывается на естественной способности человека узнавать окружающих. И несмотря на то, что идентификация человека человеком может осуществляться по комплексу признаков, включающему голос, запах, походку или одежду, именно лицо наиболее важно для опознания индивида. Существенным преимуществом распознавания по лицу перед большинством из указанных подходов, является отсутствие физического контакта проходящего процедуру идентификации человека с устройством или, иными словами, возможность идентификации на расстоянии. Аналогичным свойством на данный момент обладает лишь идентификация по радужке глаза, но расстояние между видеокамерой и человеком значительно меньше, и что ещё хуже, регулярное воздействие направленного пучка света на глаз может вызывать неприятные ощущения у человека, проходящего данную процедуру.

С точки зрения обмана системы, современные методы идентификации по лицу пока проигрывают в надёжности по сравнению с идентификацией по радужной оболочке глаза, но уже считаются более надёжными, чем распознавание по отпечаткам пальцев или геометрии кисти.

Проблема формализации и автоматизации процесса идентификации человека по изображению лица была затронута ещё на самых ранних стадиях развития систем распознавания образов, причём впервые – в СССР (монография Файна – 1970 г.) [40], и чуть позже – за рубежом (кандидатская диссертация Канада – 1973 г.) [96], но остаётся не разрешённой до сих пор [88, 121]. Это можно объяснить тем, что проблема идентификации человека по фотографии относится к разряду легко формулируемых словесно, но плохо формализуемых и трудно разрешимых.

В течение последних 10 лет количество научных исследований и публикаций по данной тематике за рубежом постоянно растёт, что также свидетельствует об актуальности этой проблемы [54, 72, 88, 100, 103, 105, 107, 110, 114, 118, 120, 121, 128, 135, 140]. В русскоязычной литературе (не

учитывая упомянутую монографию Файна) интерес к ней лишь начинает проявляться [15, 17, 35]. При этом возникают естественные затруднения в терминологии, так как понятие «распознавание лиц», являющееся дословным переводом английского термина “face recognition”, за рубежом трактуется очень широко – от непосредственно идентификации человека по изображению его лица, до таких задач, как анализ его эмоционального состояния по выражению лица, прослеживание положения головы в видеопоследовательности, «чтение» по губам и других связанных с анализом изображений лица человека задач. Проблема, рассматриваемая в данной диссертационной работе, – это автоматизация процесса «узнавания» человеком другого человека по его фотографическому изображению его лица.

Само понятие «распознавание» может быть трактовано как «отнесение исследуемого объекта, задаваемого в виде совокупности наблюдений, к одному из взаимоисключающих классов» [42, стр.4]. В таком смысле «распознавание образов является одной из разновидностей классификации» [6, стр.7], а «в тех случаях, когда каждый класс содержит только один объект, классификация эквивалентна идентификации» [6, стр.8]. Из этого следует, что, распознавание человеком лиц других людей правильно называть *идентификацией*, под которой в данном случае понимается «присвоение рассматриваемому объекту ... однозначного названия» [6, стр.7]. Понятие *верификации*, применительно к случаю распознавания человека, ещё уже – оно соответствует частному случаю идентификации при двух классах – «этот человек» и «не этот человек» (например, требуется заключить – на паре сравниваемых фотографий изображён один человек или разные люди).

Первые коммерческие приложения с функциями идентификации человека по оцифрованным фото появились в конце 90-х годов. В основном они представляли варианты систем автоматического контроля доступа и в контролируемых условиях получения изображений гарантировали до 95% корректного распознавания [120, 128, 134]. Однако, несмотря на повышенный интерес к проблеме, за последующее время не было найдено надёжных решений для других прикладных задач. Так, по данным систематически проводимых с 1996 г. тестов коммерческих и экспериментальных систем распознавания лиц – FERET, коэффициент правильного распознавания лучших систем резко падает с 98% до 60%, если им на вход поступают снимки одного человека, период времени между датами съёмки у которых составляет более года [128].

В то же время появилось много новых теоретических работ посвящённых данной проблеме. Так, например, активно применяется распознавание на основе нейронных сетей [80, 111, 155, 162], фильтров

Габора [85, 94, 108], метода главных компонент [15, 80, 136], нечёткой логики [181], геометрических преобразований [94, 103, 105, 178], локального дискриминантного анализа [85] и других [118, 122, 172, 175]. Основной трудностью для успешной реализации указанных подходов в виде прикладных систем идентификации человека остаётся вопрос автоматического определения инвариантных (к условиям съёмки, мимике, возрастным изменениям) признаков лица в реальном времени.

В настоящей работе используются отдельные методы статистического анализа данных, однако большая её часть посвящена исследованию геометрических признаков, функций расстояния, применению геометрических преобразований, что и отражено в названии работы. Под «геометрическим преобразованием плоскости» подразумевается «закон, согласно которому каждой точке плоскости ставится в соответствие определённая точка той же плоскости» [24, стр. 361]. На стадии предварительной обработки и нормализации исходных изображений в работе использовались аффинные преобразования (в частности, при сдвиге, повороте, масштабировании, выравнивании контраста, растяжении гистограммы), кроме того применялись топологические преобразования (к ним относится метод эластичных деформаций).

Поскольку существует много задач, связанных с анализом изображений лица (например, поиск лиц на групповой фотографии, прослеживание движений губ/головы на видеопоследовательности, определение эмоционального состояния и иных характеристик человека по его изображению, диагностика болезней, распознавание по профилю лица и т.д.) следует определить тип данных, поступающих на вход алгоритмов идентификации. Под термином «фотопортрет» в настоящей работе будет подразумеваться фотографическое изображение лица человека в фас (например, фото для документов) без головных уборов, украшений, солнечных очков и прочих предметов, которые могут закрывать или искажать части лица.

Анализируя тенденции развития и применения биометрических технологий, можно заключить, что поиск решений таких мало исследованных задач, как верификация человека по фото с удостоверений личности и идентификация человека по его фотопортрету путём поиска в базе данных, очень актуален в настоящее время. Исследованию алгоритмов для построения систем, решающих эти задачи, посвящена данная диссертационная работа.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность

С каждым годом растёт объём обрабатываемой информации, что требует формализации и последующей алгоритмизации процессов, ранее выполнявшихся вручную. Одним из ключевых понятий в автоматической обработке информации является “распознавание” объектов определённого класса. Когда алгоритмы выполняют распознавание на уровне эксперта-человека, автоматизация ведёт к ускорению работы систем обработки данных и повышению их эффективности.

В настоящее время активно развиваются биометрические технологии – «методы и технические средства получения и использования биометрических данных человека в целях его идентификации» [15, стр. 230]. Их целью является разработка автоматизированных систем идентификации человека на основе биометрических признаков: отпечатков пальцев, характеристик голоса и речи, рисунка радужной оболочки глаза, изображения лица. Это связано с тем фактом, что на данный момент основным препятствием для дальнейшего развития информационных сред, различного рода виртуальных сервисов и т.п. является проблема надёжной идентификации (верификации) пользователя. Ожидается, что применение подобных систем существенно уменьшит количество преступлений, связанных с несанкционированным доступом, в том числе и в компьютерных сетях, причём именно технология идентификации человека на основе изображений лиц признана наиболее приемлемой для массового применения, так как она не требует физического контакта с устройством, ненавязчива, естественна и, в потенциале, может обладать высокой надёжностью и скоростью.

Поиск в базах данных по фотопортретам человека, автоматизированный контроль удостоверений личности особенно актуальны для правоохранительных органов большинства стран в контексте увеличения количества террористических актов и уголовных преступлений на фоне повышения общей мобильности населения.

Связь работы с крупными научными программами, темами

Основой диссертационной работы являются исследования, проведённые автором в 1996-2001 гг. в Институте технической кибернетики НАН Беларуси в рамках научно-исследовательских программ: «Интеллект-17» (тема «Создание теоретических основ разработки, анализа и распознавания изображений с использованием методов искусственного интеллекта», 1996 –

2000 гг.), «Инфотех-12» (тема «Разработка методов и компьютерных технологий распознавания изображений», 2001 – 2005 гг.), «Информационные технологии» (тема «Разработка методов, инструментальных средств и интеллектуальных технологий формирования цифровых моделей местности на основе автоматизированного дешифрирования аэро- и космоснимков», 2001 – 2003 гг.), а также проектов Белорусского республиканского фонда фундаментальных исследований "Проблемно-ориентированное распознавание объектов полутоновых изображений" (грант БРФФИ №Т98-226 1999 – 2001 гг.) и "Предварительная обработка и анализ фотопортретов" (грант БРФФИ для молодых учёных №Т98М-102, 1999 – 2000 гг.).

Цель и задачи исследования

Целью диссертационной работы является разработка комплекса алгоритмов для автоматической идентификации человека по фотопортрету.

Достижение поставленной цели потребовало решения следующих основных **задач**:

- разработки алгоритма автоматической локализации центров зрачков человека, изображённого на фотопортрете;
- разработки алгоритма предварительной обработки и нормализации цифровых фотопортретов;
- разработки алгоритмов поиска антропометрических точек лица на фотопортретах;
- определения множества информативных признаков для распознавания человека по изображению лица;
- разработки алгоритмов сравнения фотопортретов и идентификации человека по фотопортрету.

Методы исследования

Теоретические методы исследования основывались на методах цифровой обработки изображений, распознавания образов, дискретных преобразований и системного анализа. Экспериментальная часть исследования базировалась на обработке и анализе цифровых изображений с помощью ЭВМ с последующей численной и визуальной оценкой результатов. Для программной реализации разработанных алгоритмов использовались методы создания программных систем и программирование на языках высокого уровня, моделирование с помощью специализированных пакетов обработки данных и изображений.

Научная новизна полученных результатов

- Разработаны алгоритмы поиска области лица на фотопортрете и локализации центров зрачков, отличающиеся способностью находить в режиме реального времени центры зрачков на слабоконтрастных изображениях, полученных сканированием фотографии с удостоверения личности.
- Разработан алгоритм повышения контраста и нормализации цифровых фотопортретов, позволяющий представить анализируемые фотопортреты в едином масштабе и ориентации с близкими яркостно-контрастными характеристиками. Он отличается возможностью выравнивания яркостных характеристик как слабоконтрастных (полученных сканированием удостоверения личности) фотопортретов, так и контрастных изображений человека, поступивших с видеокамеры.
- Разработаны алгоритмы поиска восьми основных антропометрических точек лица на цифровых фотопортретах. Впервые экспериментально определены параметры фильтра Гaborа, применяемого для обнаружения этих точек, и параметры нормализации фотопортретов.
- Определены информативные признаки изображения лица человека на основе антропометрических точек. Особенностью набора признаков является робастность к возрастным изменениям внешности человека, освещённости, макияжу и незначительным вариациям ракурса съёмки, инвариантность к масштабу, яркости и контрасту.
- Разработаны алгоритмы идентификации человека по цифровым фотопортретам на основе геометрических признаков и эластичных преобразований. Их основными отличиями от существующих алгоритмов являются: работа в условиях отсутствия априорной информации о людях, изображенных на сравниваемой паре фотографий, отсутствие необходимости в предварительном обучении системы, а также устойчивость к возрастным изменениям людей, изображённых на обрабатываемых фотопортретах.

Практическая значимость полученных результатов

Разработанные алгоритмы могут быть использованы в автоматизированных системах проведения криминалистической фотопортретной экспертизы, автоматических контрольно-пропускных устройствах. Они могут также применяться в системах информационной безопасности (доступ к ЭВМ и отдельным программам, базам данных, криптографическим приложениям, медицинским сведениям, Интернету, системам электронной торговли и т.д.), наблюдения и расследования криминальных происшествий (автоматическое наблюдение за подозреваемыми, розыск людей, активное

видеонаблюдение при чрезвычайных происшествиях и т.д.), а также в банковских целях (банкоматах, системах удалённого управления счётом).

Результаты диссертационной работы в виде законченной динамически подключаемой библиотеки программ внедрены в НПП “Регула”, производящее системы автоматизированного контроля паспортов для пунктов пограничного контроля РБ и РФ.

Результаты диссертационной работы были использованы в учебных процессах Белорусского государственного технологического университета и Белорусского государственного университета информатики и радиоэлектроники – курсах по обработке информации и распознаванию образов.

Основные положения диссертации, выносимые на защиту

- Алгоритм локализации центров зрачков на фотопортретах с разным контрастом и фоном.
- Алгоритм предварительной обработки и нормализации цифровых фотопортретов.
- Алгоритмы поиска антропометрических точек лица на цифровых фотопортретах с помощью фильтров Габора.
- Алгоритм сравнения двух фотопортретов без использования априорной информации.
- Технология поиска портретов людей, похожих на заданного человека в базах данных.

Личный вклад соискателя

Основные результаты и положения, выносимые на защиту, получены лично автором. Все алгоритмы, обсуждаемые в работе, были разработаны и экспериментально исследованы автором самостоятельно. Научные руководители принимали участие в постановке цели и задач исследования, их предварительном анализе, планировании экспериментов, а также в обсуждении полученных результатов.

Апробация результатов диссертации

Основные положения и результаты диссертационной работы докладывались и обсуждались на международных конференциях:

- “Computer Graphics and Image Processing” (г. Борки, Польша, 1998 г.);
- “Systems and Signals in Intelligent Technologies” (г. Минск, 1998 г.);
- “UkrOBRAZ'98” (г. Киев, Украина, 1998 г.);

- “Распознавание образов и обработка информации” (г. Минск, 1999 г., 2001 г.);
- “Nonlinear Signal Image Processing” (г. Анталья, Турция, 1999 г.);
- ”Acoustic, Speech, and Signal Processing” (г. Стамбул, Турция, 2000 г.);
- на международном симпозиуме “Optical Science, Engineering, and Instrumentation” (г. Денвер, США, 1999 г.)
- на Республиканском научном семинаре “Математическое моделирование сложных систем и анализ данных” (БГУ, 1999 г.);
- на семинарах Пражского Технического Университета (1997 г.), Щецинского Политехнического Университета (2000 г.) и Института технической кибернетики НАН Беларуси (1997 – 2001 гг.).

Опубликованность результатов

По теме диссертационной работы опубликовано 14 печатных работ. Из них 1 статья в международном журнале [146], 7 статей в трудах международных конференций [139, 143, 144, 145, 148, 149, 150], 1 препринт [30], 1 тезисы доклада на международной конференции [147], 4 статьи в сборниках ИТК НАН Беларуси [26, 27, 28, 29]. Прикладные результаты отражены в 6 отчётах о НИР по указанным выше темам [20, 22, 25, 31, 33, 34].

Структура и объём диссертации

Диссертация изложена на 170 страницах. Она состоит из введения (4 стр.), общей характеристики работы (5 стр.), четырёх глав (87 стр.), заключения (2 стр.). Работа содержит 68 иллюстраций (43 стр.) и 10 таблиц (6 стр.), список использованных источников, состоящий из 182 наименований (13 стр.), приложение 1 с экспериментальными данными и иллюстрациями (5 стр.), приложение 2 с документами об использовании результатов диссертационной работы (4 стр.).

ГЛАВА 1 МЕТОДЫ АНАЛИЗА ФОТОПОРТРЕТОВ

1.1 Классификация задач

Понятие «автоматическое распознавание человека по изображению лица» может быть рассмотрено с позиции трёх наиболее актуальных на данный момент задач:

1. *Автоматический контроль удостоверений личности/ верификация человека по предъявлому удостоверению личности.*
2. *Автоматический контроль доступа.*
3. *Идентификация человека по фотографии – поиск в базе данных изображений людей, похожих на заданного человека.*

1.1.1 Контроль удостоверений личности

С момента изобретения фотографии и до настоящего времени подавляющее большинство удостоверений личности обеспечиваются фотопортретом человека, которому данный документ был выдан. На протяжении времени действия документа внешность человека может претерпевать значительные изменения и органам контроля часто бывает сложно установить визуально, является ли предъявитель документа тем лицом, которое удостоверяет документ.

Основная сложность автоматизации данной задачи состоит в полном отсутствии какой-либо априорной информации при сравнении изображений предъявителя, полученных с видеокамеры, и фото, отсканированного с документа. Существенное затруднение вызывает также возможная разница в возрасте человека, изображенного на документе и стоящего перед камерой (срок действия паспорта может быть от 5 лет до 55).

Кроме того, проблемой является и получение более-менее одинаковых (по яркостным характеристикам) сравниваемых цифровых изображений. Если процесс получения качественного изображения предъявителя документа в настоящее время особых проблем не вызывает, то оцифровка фотопортрета, приклеенного на документ, осложняется несколькими факторами. Прежде всего, качество будет теряться из-за того, что для получения цифрового изображения используется не сам объект (оригинал), а его двумерное фотоизображение (т.е. копия), а также из-за невозможности контролировать первоначальные условия получения данной копии.

Во-вторых, искажение цифровых изображений будет происходить из-за присутствующих на современных документах средств защиты от подделки. Нанесение различной степени сложности узоров и печатей на документы используется во многих государствах и может послужить

дополнительным затруднением при анализе изображения и выделении из него признаков для распознавания.

Использование ламинанта – прозрачного пластикового покрытия – для увеличения срока службы документов и затруднения подделки является общепринятым средством, особенно при изготовлении паспортов и водительских прав. Он затрудняет получение контрастных изображений стандартными способами оцифровки. Изображения одного и того же человека, полученные с камеры и путём сканирования паспорта, а также их гистограммы и яркостные характеристики (средняя яркость пикселей и её стандартное отклонение) приведены на рис 1.1.

Следует отметить, что с повышением надёжности электронных удостоверяющих документов (идентификационных карт) можно прогнозировать выравнивание качества сравниваемых изображений, поскольку информация на таких документах представляется в цифровом виде и, соответственно, процессы формирования цифрового изображения предъявителя карты и записанного на ней будут одинаковыми (с допуском на разницу в характеристиках видеокамер и условий съёмки).

Подводя итог вышесказанному, можно утверждать, что основной отличительной особенностью данной задачи является сравнение двух изображений в режиме реального времени (т.е. в течение 2 – 20 с) в условиях отсутствия априорной информации об анализируемых фотопортретах.

1.1.2 Автоматический контроль доступа

В современных условиях всё чаще возникает необходимость ограничения доступа к различным помещениям, компьютерам, системам. Это предполагает разделение людей на две группы – тех, кто имеет право доступа к ресурсу, и тех, кто не имеет. Как правило, первая группа, а это 1-5-10-50 человек, намного меньше второй. Без изменения специфики задачи возможна и обратная ситуация, когда класс тех, кому доступ воспрещен, в несколько раз меньше другого класса. Например, проверка человека на его присутствие в некотором «чёрном» списке.

Чаще всего, настраивая систему, автоматически решающую данную задачу, можно контролировать условия получения изображений, которые будут храниться в базе, и добиться их соответствия тем условиям, в которых будет производиться идентификация человека. Под понятием «условия» в данном случае можно подразумевать как освещённость человека при съёмке, так и его положение перед камерой (ракурс, расстояние до объектива), мимика и пр. Для уменьшения вероятности неправильной идентификации, при создании классификатора можно предусмотреть использование

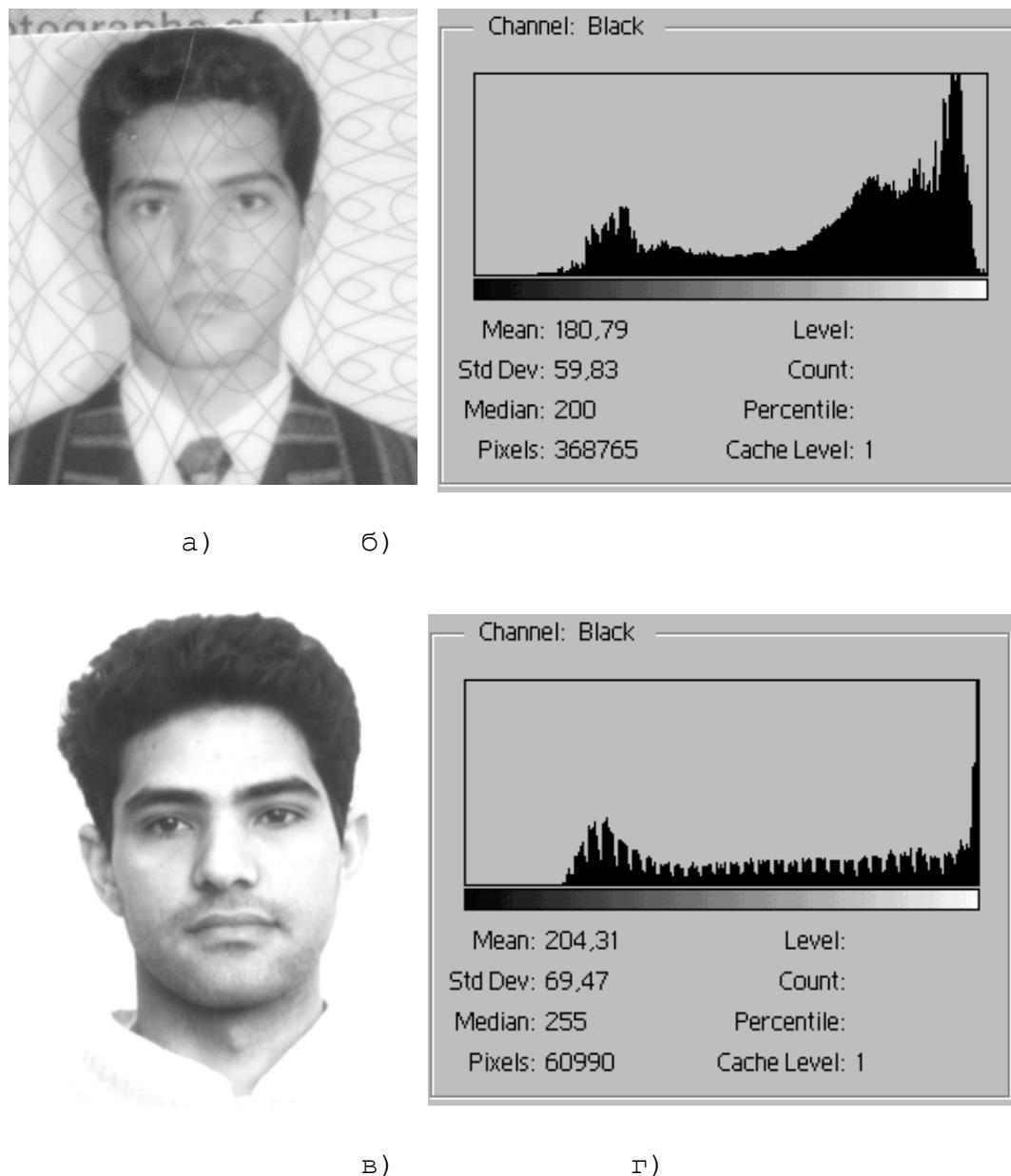


Рис. 1.1. Пример фотопортретов человека и их гистограмм, полученных: а) сканированием паспорта; в) с видеокамеры; б) и г) гистограммы изображений (а) и (в) соответственно

нескольких изображений, принадлежащих одному человеку (с вариациями в ракурсе, мимике, освещённости), вплоть до сравнения видеопоследовательностей некоторых определённых движений головы и мимических мышц лица. Кроме того, при решении задачи такого рода не возникает проблема учета возрастных изменений (поиска и выбора инвариантных во времени признаков). Как правило, у разработчика системы имеется возможность предусмотреть обновление базы изображений в случае повышения количества ложных отказов в доступе. Основное ограничение на выбор алгоритма – это необходимость функционирования системы в режиме реального времени, а также тот факт, что допустимы ошибки только I-го рода – “ложно не идентифицирован”, а II-го рода – “ложно разрешён доступ” – должны быть сведены к нулю [6,42].

1.1.3 Поиск фотопортретов заданного человека в базе данных

Данную задачу можно сформулировать следующим образом: имеется база данных, содержащая N фотопортретов (от нескольких тысяч и больше), требуется найти в ней n изображений, наиболее похожих на искомое. Выбор окончательного решения предоставляется эксперту (оператору). Следует отметить, что по статистике в реальной базе данных, содержащей 100 000 портретов, 417 человек будут иметь двойников [68]. Так как базы могут содержать сотни тысяч, миллионы изображений, то существующие системы работать в режиме реального времени не в состоянии. Их цель – решение задачи за разумное время.

Существует разновидность данной задачи – поиск изображений в базе данных по словесному описанию, т.е. на вход системы подаётся не изображение, а некоторый набор ключевых слов, описывающих внешность. Например: «мужчина», «европеец», «25-35 лет», «крупный нос», «открытый лоб» и т.д. Данная разновидность задачи не рассматривается в настоящей работе, так как её решение сводится к построению классификатора на основе двоичного дерева и соответствующей индексации изображений, хранящихся в базе. Она может являться дополнительной опцией к системе автоматического поиска изображений заданного человека.

Портреты, хранящиеся в базе данных, могут быть получены как в одинаковых условиях (например, база данных ГАИ РБ), так и в отличающихся (если база формировалась из разных источников).

Спецификой решения данной задачи, на взгляд автора, является критичность выбора значимых признаков, с тем чтобы портреты похожих друг на друга людей были близки в признаковом пространстве. Трудность

заключается в том, что сам человек не может однозначно определить степень похожести (или непохожести) двух различных людей, а также сказать, какие именно черты лица играют решающую роль в идентификации одного человека другим человеком. Насколько известно автору, исследования по определению степени важности для идентификации одной черты лица по отношению к другой ни криминалистами, ни кем-то другими не проводились. Известно лишь предположение, что значимость черт для идентификации уменьшается с верха лица вниз, т.е. такие признаки, как волосы и глаза важнее для распознавания, чем подбородок или губы [140]. Данное предположение частично подтверждает и рис. 1.2, на котором изображена траектория движения зрачков человека при рассматривании портрета в процессе психофизических экспериментов. Видно, что основное внимание человека концентрируется в районе глаз представленного изображения лица.

1.2 Подходы к распознаванию по фотопортретам

Наибольшее количество публикаций, относящихся к теме анализа изображений лиц, посвящены нахождению оптимальных методов решения задачи второго типа. Существует около десятка принципиально разных методов применяющихся для решения рассматриваемой проблемы (рис.1.3). Основными критериями при оценке приведённых ниже методов являлись вычислительная стоимость алгоритмов, построенных на их основе, и достоверность распознавания.

1.2.1 Антропометрический метод распознавания

Люди существенно различаются своей внешностью, и в частности чертами лица (например, расположение глаз и их мельчайшие характеристики различаются даже у близнецов). Поэтому не удивительно, что исторически первый подход к решению проблемы автоматической идентификации человека по изображению его лица был основан на выделении и сравнении некоторых антропометрических характеристик лица. Этот подход давно используется в практической криминалистике, однако замеры и сравнение выполнялись и выполняются вручную.

Главная проблема, которая возникает перед разработчиками систем распознавания основанных на данном подходе, - выбор совокупности характерных точек, наилучшим образом описывающих человеческое лицо, и способов надёжного нахождения таких точек. При этом необходимо учитывать следующие требования: точки на лице или черты лица, на которых основывается идентификация, не должны закрываться причёской, бородой,

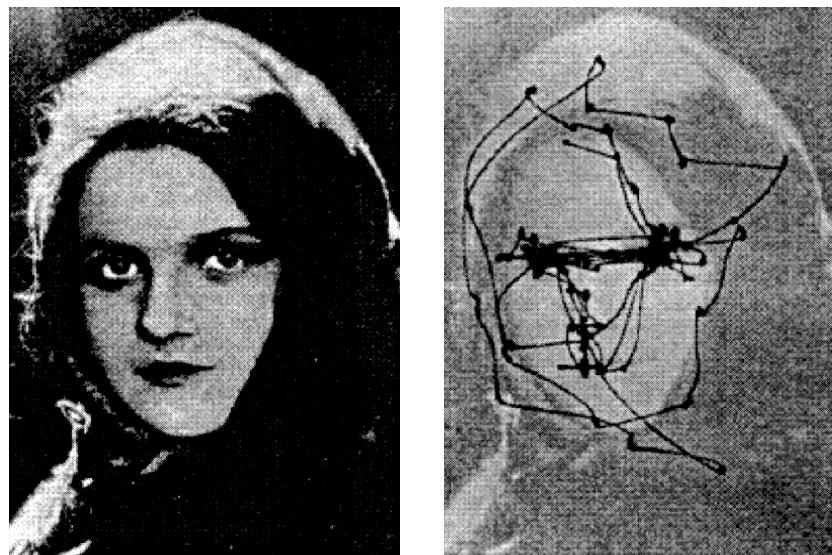


Рис.1.2. Траектория движения взгляда при распознавании человека по портрету



Рис.1.3. Основные подходы, применяемые при построении систем идентификации личности

маской и т.п. Для обеспечения независимости процесса распознавания от масштаба изображения целесообразно описывать систему идентификационных признаков в относительных единицах; выбранная система точек должна обеспечивать достаточную устойчивость процесса распознавания при незначительном изменении как ракурса съёмки (легкий поворот головы, наклон), так и мимики (улыбка, сосредоточенное выражение лица и т.д.); количество характерных точек системы, удовлетворяющей вышеизложенным требованиям, должно быть минимальным, так как вычислительная стоимость алгоритмов обычно пропорциональна количеству вычисляемых признаков.

К настоящему времени имеется много работ, посвященных распознаванию с помощью различных совокупностей характерных точек и анализу эффективности работы системы, построенных на их основе [35, 49, 57, 65, 94, 96, 103, 105, 107, 178].

Общие требования к набору выбираемых для описания лица точек сформулированы в разд. 2.3. Дополнительно следует отметить, что использование лишних параметров может не только затормозить работу алгоритма, но иногда и снизить точность распознавания. Так, в одной из первых работ по автоматизированной идентификации человека [96] при использовании набора из 16 параметров лица, среди которых были соотношения между расстояниями, площадями и даже определялись углы между выбранными комбинациями точек, коэффициент распознавания колебался от 45 до 75% в зависимости от набора параметров. Причем лучшие результаты были получены, когда некоторые признаки не использовались. При проверке схожести между изображениями из тестового набора и основной базы Канад [96] использовал фотографии 20 человек - по две на каждого. Снимки были сделаны с перерывом в месяц, но без изменений в условиях съёмки (ракурс, освещение и т.п.).

Брунелли и Поджио в работе [57] использовали 35 автоматически выделяемых геометрических характеристик, среди которых:

- толщина бровей над центрами зрачков;
- приблизительное описание арки брови над левым глазом;
- вертикальная координата кончика носа и его ширина;
- вертикальная координата местоположения рта, ширина и высота верхней и нижней губ;
- одиннадцать радиусов, описывающих форму подбородка;
- ширина лица на уровне кончика носа;
- ширина лица посередине между линией глаз и кончиком носа.

Применяя к описанным подобным образом лицам классификатор Байеса [9], авторы добились 90%-ного распознавания на базе данных 47 человек. Следует отметить, что метод сравнения с эталоном на той же самой базе изображений дал 100%-ую точность распознавания [57].

Кокс и другие в [65] предложили применять «смешанные расстояния» (различные арифметические комбинации расстояний), которые позволили получить точность распознавания 95% при поиске 95 изображений в базе, содержащей портреты 685 человек. Однако при экспериментах информативные точки для определения 30 «смешанных расстояний» выделялись на изображениях вручную. Таким образом, остаётся открытым вопрос об эффективности использования рассматриваемого в [65] набора признаков в автоматических системах.

В [35] описывается патент США №3805238 системы идентификации человека, основанной на анализе характерных черт лица. Данный патент описывает систему, построенную для решения задачи автоматического контроля доступа. Для настоящего исследования представляет интерес приведённая в описании система характерных точек и получаемых на их основе расстояний (рис.1.4, 1.5), а также её устойчивость к незначительным изменениям условий съёмки (поворот, мимика, освещение и т.д.).

На рис.1.4,а схематически показано фронтальное изображение головы человека. Ориентация фронтальной проекции определяется линией, перпендикулярной воображаемой оси, проходящей через центр ушей человека. Из рис.1.4,а следует, что фронтальные проекции идентификационных точек обозначены следующим образом: центр зрачка (15а и 15b); кончик носа (20), который определяется как центральная точка между носовыми отверстиями, центр рта (25) - как точка пересечения линии, разделяющей верхнюю и нижнюю губы объекта, и перпендикуляра, опущенного из точки, определяющей кончик носа объекта; левый угол левого глаза, обозначенный точкой 30а, правый угол правого глаза - 30b; нижние точки окончания мочек ушей - 35а и 35b. При желании могут быть использованы и другие характерные точки лица, например крайняя точка правого угла левого глаза.

Расстояния между соответствующими идентификационными точками на лице авторы называют «лицевыми параметрами». Номера признаков на рис. 1.5 означают следующее:

- 1 - расстояние между центром сетчатки правого глаза и точкой кончика носа (REN);
- 2 - расстояние между центром сетчатки правого глаза и центром ротового отверстия (REM);
- 3 - расстояние между центрами сетчатки глаз (LER);

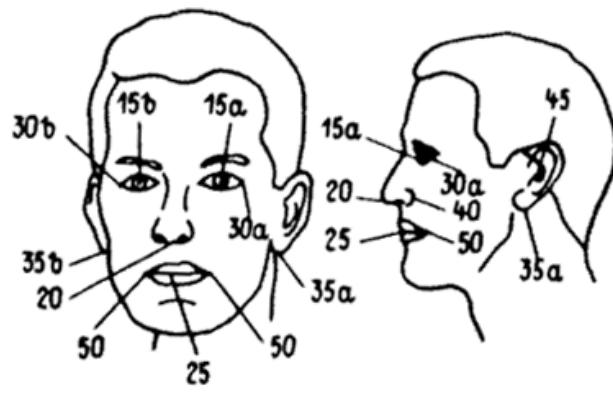
- 4 - расстояние между центром сетчатки левого глаза и центром кончика носа (LEN);
- 5 - расстояние между центром сетчатки левого глаза и центром ротового отверстия (LEM);
- 6 - расстояние между центром ротового отверстия и кончиком носа (DMN).

Все выбранные и отмеченные выше параметры приняты в качестве основных, поскольку на них не влияют такие факторы, как причёска, макияж, наличие ювелирных украшений и пр. Естественно, что количество параметров может быть увеличено. Однако следует отметить, что для приближённой идентификации с определённой степенью надёжности достаточно использовать всего несколько параметров.

Используя шесть отобранных параметров, можно составить до 30 линейных соотношений между ними, которые могут использоваться в качестве признаков. При этом 15 из них будут прямыми, а 15 - обратными соотношениями. Основной причиной использования отношений между параметрами является то, что они не зависят от расстояния, с которого делается фотография человека, а по фотографии невозможно определить действительные размеры головы и её участков. Таким образом, если используются отношения параметров, а не их реальные величины, то для системы совершенно не важно, на каком расстоянии находится измеряемый объект.

В соответствии с вышесказанным для установления идентичности человека чаще других используются следующие отношения параметров: LEN/LER; LEM/LER; REN/LER; REM/LER; DMN/LER, кроме того, в качестве дополнительных были выбраны усредненные отношения: $(LEN/LER+REN/LER)/2$ и $(LEM/LER+REM/LER)/2$.

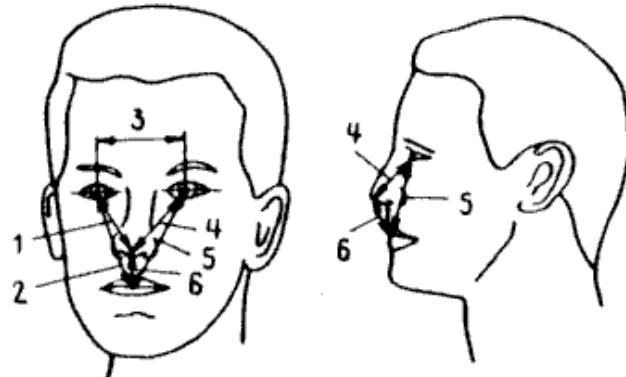
При проведении опытов с системой на девяти испытуемых было показано, что при самом простом методе получения изображения лица человека и простейшем оборудовании точность распознавания измерения достаточно хорошая. При этом не выдвигалось требование, чтобы лицо человека находилось при проверке в строго определённом положении. Испытуемые просто опускали свои пропуска в считывающее устройство, находясь в стандартном положении. Кроме того, семь приведённых отношений изменяются для девяти испытуемых не более чем на 21%. Различные серии проведённых на одних и тех же объектах измерений показали, что отклонение для всех семи показателей составляет не более 3%. Таким образом, используя шесть вышеуказанных параметров и вычисляя семь отмеченных выше соотношений (включая два усредненных) авторы получали коэффициент распознавания до 98% (на не очень большом тестовом множестве).



а)

б)

Рис. 1.4. Проекции лица человека: а) фронтальная и б) профильная [35]



а)

б)

Рис. 1.5. Идентификационные параметры: а) на фронтальной и б) на профильной проекции [35]

В [49] опубликованы результаты сравнения четырёх подходов. Один из них был геометрическим и основывался на расстояниях между точками лица. Эксперименты проводились на базе данных Исследовательской лаборатории фирмы Оливетти, состоящей из 400 изображений 40 человек [107, 141]. Для данного подхода был получен коэффициент распознавания 87,36%. Кроме этого, в этой работе представляется интересной экспериментальная проверка влияния количества выбранных характеристик на точность распознавания. Так, на 47 тестовых портретах было установлено, что при уменьшении размерности описывающего изображение вектора с 30 до 15 точность распознавания снизилась лишь на 2 %. О временных характеристиках подходов в работе не сообщалось.

Геометрический подход был использован Бартоном с коллегами в [59], но координаты антропометрических точек лица, на основании которых вычислялись признаки, определялись вручную. С помощью дискриминантного анализа они отбирали 12 из 18 исходных расстояний. Целью экспериментов являлось решение более простой, чем идентификация, задачи – отличить мужчину от женщины. Правильный результат работы системы был достигнут в 85,5% случаев на базе данных, состоящей из 91 мужского фотопортрета и 82 женских.

1.2.2 Метод главных компонент

Одним из наиболее проработанных и популярных является метод главных компонент (иногда его называют методом «собственных векторов» либо «собственных лиц»), основанный на преобразовании Карунена-Лоэва [15, 41, 80, 136]. Он является примером того, как статистические методы, успешно применяющиеся в других областях, оказались эффективно адаптированными к идентификации человека по портрету.

Простейший вариант представления любого цифрового изображения в виде вектора в пространстве признаков предусматривает учёт значения каждого пикселя как координаты отдельной оси. Так, если изображение описывается матрицей пикселей $m \times n$, то размерность такого пространства будет равна произведению m на n . Однако в связи с тем, что все человеческие лица схожи между собой (овальная форма с носом, ртом, глазами и т.д.), 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.

Таким образом, возникает задача построения пространства меньшей размерности, в котором описываются только изображения человеческих

лиц. Одним из вариантов является пространство, базисными векторами которого служат главные компоненты всех содержащихся в нем изображений лиц. Размерность такого оптимального пространства заранее определить невозможно, но она намного меньше размерности исходного пространства всех изображений. Из вышесказанного вытекает, что главной целью метода главных компонент является значительное уменьшение размерности пространства признаков таким образом, чтобы оно как можно лучше описывало «типичные» образы, принадлежащие множеству портретов. В случае применения данного метода для идентификации лиц образами будут служить фотопортреты из обучающей выборки.

Иными словами, при помощи анализа главных компонент удается выявить различные изменчивости в обучающем наборе изображений лиц и описать эту изменчивость при помощи нескольких переменных. Эти переменные представляют собой $m \times n$ -размерные векторы, которые называются собственными. Если преобразовать подобные векторы в изображения, то получаемые картинки будут отражать главные компоненты представленного обучающего множества (также называемые «собственные лица»).

Таким образом, за счёт снижения размерности пространства базисных векторов, в котором находятся изображения, добиваются хороших показателей в скорости распознавания изображений, близких к обучающей выборке.

Полученный один раз на основе представительной обучающей выборки набор собственных векторов (рис.1.6) используется для кодирования всех остальных изображений, которые представляются взвешенной комбинацией этих собственных векторов (рис.1.7, 1.8). Иными словами, используя ограниченное количество собственных векторов, можно получить улучшенную аппроксимацию к входному изображению, которая затем хранится в базе данных в виде вектора коэффициентов, служащего одновременно ключом поиска.

Обоснование выбора одного из нескольких возможных наборов базисных векторов описано в работе [152]. На базе из 1316 изображений 504 человек авторы продемонстрировали преимущество их подхода к выбору множества базисных векторов и получили точность распознавания 95-99% (в зависимости от вида опытов).

Тюрк и Пентланд [160] провели комплексное исследование данного метода на базе изображений 16 человек, 2500 фотографий которых были получены при различных условиях освещения, при съёмке с различных расстояний, при разных поворотах головы. Однако полученные изображения были одинаковы по таким критериям, как выражение лица, элементы

Обучающий набор



Рис. 1.6. Формирование k векторов для последующего определения главных компонент обучающего набора

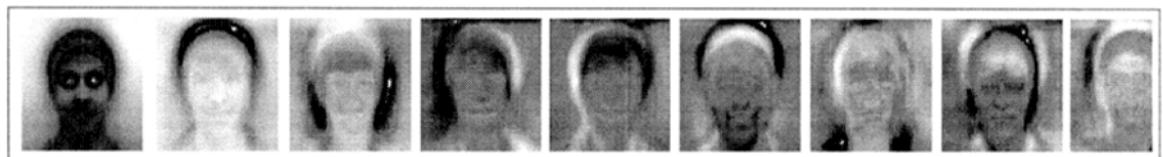


Рис. 1.7. Девять первых главных компонент лица человека [136]

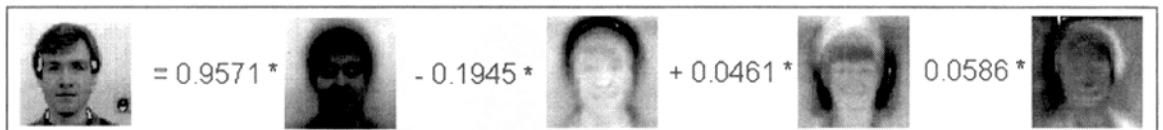


Рис. 1.8. Портрет, представленный в базисе главных компонент [136]

внешности (борода, очки и т.д.). При изменении освещения, ракурса съёмки и масштаба коэффициент распознавания составил 95, 85 и 64% соответственно. Для уменьшения влияния на результат идентификации изменений причёски и фона обрабатывалась только центральная часть лица. Скорость работы системы, реализованной авторами на рабочей станции SUN 3/160, приближалась к режиму реального времени.

Наиболее представительные по объёму эксперименты были описаны Пентландом с коллегами в [127]. Они исследовали эффективность применения метода при работе с большими базами данных, одна из которых состояла из 7562 изображений, принадлежащих почти 3000 человек. На данный момент это самая большая база изображений лиц, о работе с которой было объявлено. В ходе экспериментов были получены 20 собственных векторов из 128 случайно выбранных обучающих изображений. В дополнение к векторам в базе хранилась информация о поле, расе, приблизительном возрасте человека, выражении его лица. В отличие от полицейских баз данных, где используются лишь два изображения одного объекта - в фас и профиль, в экспериментальной базе было много (более десятка) изображений, принадлежащих одному человеку, но отличающихся разными выражениями лица, причёской и т.д.

Однако эксперименты, проведённые с базой, можно назвать ограниченными: проверялась лишь возможность интерактивного поиска по всей базе данных. Эксперт задавал системе словесное описание лиц определённого типа (например, «негритянка 50 лет»). Изображения, которые удовлетворяли данным запросам, выдавались на экран группами по 21. Пользователь выбирал среди них одно и система выдавала лица из базы, которые были наиболее похожи на требуемое, в порядке убывания сходства. Во время экспериментов с 200 выбранными изображениями коэффициент распознавания составил 95%, т.е. для 180 из них наиболее похожей была выбрана фотография того же самого человека. Учитывая тот факт, что выбранное изображение сначала раскладывалось на собственные векторы, а затем по ним производился поиск, ошибка распознавания прямо зависит от погрешности разложения как выбранного, так и хранящихся в базе изображений. Для оценки точности распознавания как функции от расы были проведены опыты с изображениями бело- и чёрнокожих, а также азиатских лиц мужского пола. Для европейской и африканской рас точность распознавания была 90 и 95% соответственно, а для азиатской – лишь 80%.

В работе [115] Могхаддам и Пентланд объявили о получении хороших результатов при экспериментах с базой данных FERET- только одно неправильное определение на 150 фронтальных изображений. Система

проводила расширенную предобработку по положению головы, определению черт лица, а также нормализацию по геометрии лица, освещённости, контрастности, масштабу и повороту.

На основе вышесказанного можно отметить, что хотя алгоритм распознавания изображений, представленных комбинацией собственных векторов, и является относительно быстрым, простым и практичным, при его работе с большими базами данных могут появиться проблемы с точностью. Один из способов преодоления последних - повышение степени соответствия между интенсивностями пикселей изображений из тренировочного набора и из базы. Подобные шаги, в свою очередь, требуют больших усилий по предобработке и нормализации.

1.2.3 Статистическое распознавание изображений

Подобно предыдущему методу в вероятностных моделях также используется обучающий набор. При этом формируются два класса из всех вариантов представления объектов: внутриобъектной и внешней изменчивости, т.е. отбираются признаки, по которым все портреты делятся на два класса: 1) портрет данного человека, 2) все другие портреты. Функции плотности вероятности для каждого класса оцениваются при помощи упомянутого выше обучающего множества и впоследствии используются для вычисления меры схожести, которая таким образом основывается на полученных опытным путём вероятностях. Кроме того, для получения более точных результатов иногда используется вероятностная модель некоторого физического процесса, при помощи которой и формируется окончательная мера схожести двух изображений.

Одними из первых исследователями данного подхода являются Б. Могхаддам, С. Настар и А. Пентланд. При идентификации по фотопортрету [116, 119] авторы определяют два класса возможных изменений изображений лиц: внутриобъектный Ω_I (различные выражения лица одного человека) и внеобъектный Ω_E (разница во внешности двух различных индивидуумов). Тогда мера схожести в терминах теории вероятности может быть выражена следующим образом:

$$S(I_1, I_2) = P(d(I_1, I_2) \in \Omega_1) = P(\Omega_1 | d(I_1, I_2)) , \quad (1.1)$$

где $P(\Omega_1 | d(I_1, I_2))$ является вероятностью, полученной по правилу Байеса [10] на основе опытов с применением оценки подобия $P(d(I_1, I_2) | \Omega_1)$ и $P(d(I_1, I_2) | \Omega_E)$, вычисленных из обучающих данных с использованием эффективного метода оценки плотности многомерных данных. Кроме того, авто-

ры используют новое представление для различий в сравниваемых изображениях $d(I_1, I_2)$, которое сочетается как с пространственными (X, Y), так и с яркостными (I) компонентами в унифицированной XYI структуре (непохожей на предыдущие подходы, которые по существу обрабатывали форму и структуру независимо). В частности, I_1 моделируется трёхмерной физической поверхностью в XYI -пространстве, которое деформируется в соответствии с привлекаемыми «физическими силами», заданными с помощью множества I_2 . Динамика подобных систем эффективно вычисляется при использовании «аналитических методов вибрации», при которых получают трёхмерное поле соответствия для искривления I_1 в I_2 . Дополнительно авторы используют параметрическое представление $d(I_1, I_2) = \bar{U}$, где \bar{U} – это спектр модальных амплитуд окончательной деформации. Это множество методов подгонки для соответствия изображений может рассматриваться в более общей формулировке, которая, в отличие от оптического потока, не требует постоянной оценки яркости.

Эксперименты по проверке данной системы проводились на изображениях базы данных FERET. Обучающий набор состоял из 38 пар изображений (по две фотографии на одного человека), тестовый набор – из 76 пар. Авторы добились хороших показателей в скорости (сравнимых с «собственными лицами») и точности распознавания – 92% (70 из 76).

К статистическому подходу можно также отнести опыт применения скрытых марковских моделей для идентификации человека. В работе [74] говорится о достижении 100%-ной точности распознавания с помощью системы, построенной на их основе.

1.2.4 Сопоставление изображения с эталоном

В данном подходе процесс идентификации разбивается на этапы, соответствующие анализу отдельных черт лица [138]. Каждая фотография, поступающая на вход распознающей системы, должна представлять собой фронтальное изображение лица человека с определённым для конкретной базы данных количеством масок, представляющих основные для идентификации регионы лица (например, глаза, нос, рот и нижняя часть лица). Кроме того, расположения данных масок должны быть одинаково нормализованы (например, относительно положения глаз) для всех изображений в базе данных. Во время процесса распознавания, когда части входного изображения по очереди сравниваются с частями изображения, хранящегося в базе, вычисляется оценка, отражающая результат сравнения в баллах (один балл за каждую совпавшую черту лица), на основе нормализованной взаимной корреляции (впрочем, методы сравнения могут быть разными).

После этого входное изображение классифицируется в соответствии с максимально набранными баллами. Имеются также некоторые разновидности данного подхода, например, с изменяющимися в процессе сравнения эталонами [57].

Наиболее известной работой, объявляющей об успехах в применении сопоставления с эталоном, является [57]. На базе данных из 47 человек авторы добились 100%-ной точности распознавания. В своих экспериментах они использовали 4 области лица для сравнения: глаз, носа, рта и нижней части лица (рис. 1.9). Однако, несмотря на успехи в точности системы, авторы признают высокую вычислительную стоимость работы алгоритма. Для сравнения двух изображений с расстояниями между зрачками в 27 пикселей и использованием эталонов лишь глаз, носа и рта требуется 25 мс времени работы SPARCStation IPX. Кроме того, к недостаткам этого подхода относится необходимость обеспечения одинаковых условий получения изображений (освещения, ракурса, масштаба).

Другой вариант этого подхода описан в [118]. Авторы используют изолинии уровней яркости, которые не принадлежат к трёхмерным структурам, но передают изображение рельефа лица. Сначала, применив к изображению лица на однородном (чёрном) фоне оператор Собеля [98] и другие операторы предобработки, они получают границы области лица, а затем при помощи гистограммы яркостей разделяют контуры головы на изолинии. После этого используется процедура сопоставления с эталоном. На базе данных, состоящей из двадцати изображений, были получены коэффициенты распознавания 80-90%.

1.2.5 Распознавание с помощью нейронных сетей

Существует около десятка основных разновидностей нейронных сетей. Принципы функционирования систем, построенных на автоассоциативной памяти (АП), заключаются в том, чтобы в ответ на некоторую входящую совокупность данных, называемую «ключом», выдать на выход хранящуюся в сети и наиболее близкую к входной по значениям совокупность такой же размерности [80, 155]. В случае идентификации по фотопортрету ключом служит изображение лица человека [77, 80, 101, 104, 162, 169].

Одним из самых широко используемых вариантов является сеть, построенная на многослойном перцептроне, которая позволяет классифицировать поданное на вход изображение/сигнал в соответствии с предварительной настройкой/тренировкой сети.

Порядок работы нейронных сетей на основе автоассоциативной памяти следующий: первый шаг - изображение оцифровывается и кодируется в



Рис. 1.9. Эталоны некоторых областей лица, использовавшиеся в [57]

виде вектора; второй шаг - каждая координата вектора располагается в отдельной ячейке, связанной со всеми остальными ячейками (обучение или настройка системы происходит путём изменения весов связей между ячейками); третий шаг - векторы обрабатываются нейросетью, при этом на её выходе формируется ближайшее к входному хранящемуся в АП изображение.

Линейная автоассоциативная память представляет собой один слой нейронной сети. Каждый нейрон этого слоя ассоциируется с одним компонентом, получившимся из разложения изображения лица вектора (аналогично методу «собственных лиц»). Таким образом, при размере изображения $m \times n$ пикселей каждый слой данной сети будет содержать $m \times n$ нейронов. Кроме того, каждый нейрон связан со всеми остальными и линейная АП строится при вычислении $m \times n$ весов связей этой нейронной сети. Данные веса определяются на этапе обучения, при котором несколько обучающих изображений, представленных линейно АП, запоминаются в ней [104].

Несмотря на то, что точно оценить вычислительную сложность алгоритмов, основанных на нейронных сетях, затруднительно, сообщается о достаточно эффективном использовании нейронных сетей в области анализа изображений лиц по трем направлениям: классификация людей по полу, непосредственно идентификация и определение эмоциональных выражений лиц [57, 59, 69, 82, 86, 99, 104, 107, 163].

Об одном из первых применений нейронных сетей для извлечения из базы изображений, близких к заданным, стало известно из работы [99]. При проверке данного подхода на малых наборах данных были получены хорошие показатели точности распознавания даже тогда, когда входное изображение было очень сильно зашумлено и когда части изображения были вырезаны. В [69] авторы выделяли 50 первых принципиальных компонент изображений и уменьшали их размерность до 5, применяя автоассоциативную нейронную сеть. Результирующую классификацию получали при помощи стандартного многослойного перцептрана. Были получены хорошие результаты, но база данных была достаточно проста: изображения нормализовались вручную и среди них не было различий по свету, повороту или наклону. Иерархическая нейронная сеть, которая растёт автоматически, также применялась для распознавания лиц в [167]. Авторы добились хороших результатов различия десяти человек.

Используя в своей системе алгоритмы, основанные на самоорганизующихся картах (SOM) и свёрточных нейронных сетях (Convolutional Networks), Лоренс и другие в [107] провели её комплексное исследование по следующим пунктам: изменение числа классов, по которым проводилось

распознавание; изменение размерности самоорганизующихся карт; изменение уровня квантизации SOM; замена SOM на разложения Карунен-Лоэва; замена указанных нейросетей на многослойный перцептрон; проверка зависимости между коэффициентом распознавания и порогом отвержения; сравнение с другими системами на одних базах изображений; изменение числа изображений одного человека, применяемых для идентификации. Эксперименты проводились на изображениях из ORL-базы с использованием пяти изображений одного человека для обучающего набора и столько же для тестового - всего по 200 изображений в обоих множествах. Наилучшая достоверность распознавания составила 94,25%. Кроме того, авторы утверждают, что по производительности их система почти в 500 раз превосходит вторую после нее систему 2D-HMM (разработанную Самария [141]), однако никаких временных параметров не приводится.

В статье [82] описывается система идентификации личностей по их изображениям, хранящимся в базах данных, построенная по так называемой гибридной архитектуре классификации. Данная архитектура включает в себя как нейронные сети, так и деревья выбора. Подобная организация позволяет добиться 96-97% точности распознавания в зависимости от типов поиска, осуществляемых системой. Эксперименты проводились на базе данных FERET, состоящей из 1084 изображений 350 человек.

Гибридные нейронные сети, отличающиеся, однако, от описанных в [82], положительно характеризуются и в работе [86].

Более подробную информацию о некоторых направлениях развития нейронных сетей и моделях лиц, использующихся в построенных на указанных выше принципах системах, можно найти в обзоре [163].

Брунелли и Поджио в [56] экспериментировали с нейросетью, основанной на «гипербазисной» функции. В качестве признаков, описывающих исходные фотопортреты, использовались такие характеристики лица как: расстояния от зрачков до носа, ширина носа, радиус подбородка, ширина бровей и т.п. Общее количество признаков составило 18. Все они вычислялись автоматически. База данных состояла из 168 изображений 21 мужчины и 21 женщины. Система обучалась на всех изображениях базы за исключением одного для каждого класса, на которых она впоследствии тестировалась. Средний коэффициент распознавания составил 87,5 %.

1.2.6 Анализ оптических потоков изображений

Применение этого метода в целях идентификации человека признано довольно эффективным, но дорогим с вычислительной точки зрения и в практических приложениях почти не используется. Суть метода в

следующем: сравниваемые фотопортреты A и B превращаются в многослойные пирамиды - A^i и B^i . В основе пирамид лежит исходное изображение A или B , значения четырёх соседних пикселей которого формируют значение одного пикселя изображения следующего уровня A^1 или B^1 . После применения рекурсии i раз будут получены A^i и B^i .

После завершения данного процесса на соответствующих слоях двух разных пирамид производят поиск подходящих между собой наилучшим образом групп пикселей. Для каждого блока изображения A^i определяется вектор смещения. Этот вектор уточняет смещение между центрами блока из A^i и наиболее близким к нему блоком из B^i . Аналогично строятся векторы и для изображения B^i . Анализируя получившиеся системы векторов, можно сделать вывод о степени схожести сравниваемых изображений.

Несмотря на относительную простоту метода, исследований его применения для распознавания лиц мало. В [102] опубликованы результаты работы системы идентификации, основанной на данном подходе. Точность распознавания данной системы составила 92,7%. Эксперименты проводились на базе данных из 78 изображений 38 персон. Учитывая тот факт, что на сравнение одной пары уходило около 1 мин машинного времени на ЭВМ с параллельной структурой СМ-5, можно предположить, что для проверки присутствия одного изображения в такой малой, по практическим меркам, базе требовалось около 1 ч или больше. Таким образом, хотя поведение системы полностью предсказуемо при увеличении базы данных её производительность неудовлетворительна с практической точки зрения, так как при 6000 изображений время поиска составит порядка 25 дней.

1.2.7 Распознавание путём эластичного сопоставления графов

В ряде работ [58, 62, 72, 94, 100, 104, 171] описаны исследования применения метода сравнения фотопортретов путём эластичного сопоставления описывающих их графов. Сравниваемые объекты – изображения лиц – описываются в виде графов с взвешенными вершинами и рёбрами. При последующем сопоставлении один из графов – эталон – остаётся неизменным, в то время как другой деформируется с целью наилучшей подгонки к первому. Эталонный график может представлять собой как прямоугольную решётку [72], так и структуру, образованную характерными точками лица [169, 170]. Веса вершин соответствуют значениям признаков, характеризующих близлежащие пиксели изображения, в то время как веса рёбер – расстояниям между смежными вершинами. Различие между двумя графиками $d(Q,R)$ вычисляется при помощи некоторой функции, учитывающей как

значения признаков – веса вершин, так и степень деформации рёбер графа.

Деформация графа происходит путём смещения каждой из его вершин на некоторое расстояние в определённых направлениях относительно её исходного местоположения и выбора такой её позиции, при которой разница в весе вершины Q_{vi} деформируемого графа и соответствующей ей вершине R_{vi} эталона будет минимальной. Данная операция выполняется поочерёдно для всех вершин графа до тех пор, пока не будет достигнуто наименьшее (для данной пары графов) значение $d(Q, R)$.

Дюк и др. в работе [72] предполагали, что веса вершин и веса рёбер независимы и расстояние между графиками вычисляется по формуле:

$$d(Q, R) = \sum_{i=1}^{N_v} d_v(Q_{vi}, R_{vi}) + \lambda \sum_{j=1}^{N_e} d_e(Q_{ej}, R_{ej}), \quad (1.2)$$

где Q_{vi} – значение i -й вершины графа Q ; R_{ej} – значение длины j -го рёбра графа R ; N_v , N_e – количество вершин и рёбер графа соответственно и λ – коэффициент, характеризующий жёсткость графа. Функции $d_v(Q_{vi}, R_{vi})$ и $d_e(Q_{ej}, R_{ej})$ отражают разницу весов i -х вершин графов Q , R и j -х рёбер соответственно. Самый пластичный график будет иметь место при $\lambda=0$, в то время как жёсткие графы могут быть получены при больших значениях λ .

Чаще всего [58, 101, 104, 170, 171] в качестве весов вершин графа используются комплексные значения фильтров Габора или их упорядоченных наборов – Габоровских вейвлет [94, 108], которые вычисляются локально путём свёртки значений яркости в разных окрестностях вершин графа (см. п. 1.2.8.). В работе [100] предлагается формировать признаки, описывающие вершины, с помощью операций математической морфологии – расширения и эрозии.

Следует заметить, что в “чистом” виде подход эластичного сопоставления графов практически не используется. Так, системы, представленные в [58] и [104], основываются на архитектуре динамических связей (Dynamic Link Architecture - DLA), реализованной в виде нейросети. Для представления изображений используется график в виде прямоугольной решётки. Распознавание нового изображения (эталона) происходит при его преобразовании в график и сравнении с графиками, хранящимися в сети. Качество сравнения оценивается при помощи ценовой функции деформации. Размер решётки деформируемого графа может изменяться, но её центр зафиксирован, если общая цена сопоставления графа с эталоном снижается, то принимается новый размер. Этот процесс повторяется до тех пор, пока не будет найдена наименьшая цена подгона хранящегося графа к эталону.

Распознавание производится после определения полной цены для большинства из представленных изображений каждого класса. Таким образом, система идентифицирует человеческое лицо путём сопоставления выделенного из него графа с набором хранящихся графов [62].

Однако вычислительная стоимость подобной системы очень высока. Так, Лэйдс и другие при экспериментах с базой данных, состоящей из 87 человек, добились 95%-го коэффициента распознавания при том, что изображения лица имели различные выражения и повороты до 15° . Но на сравнение одного изображения с 87 тратилось приблизительно 25 с при работе системы на параллельной ЭВМ с 23 транспьютерами.

Вискотт с коллегами использовал усовершенствованную версию подхода [171] и сравнивал 300 изображений лиц с другими 300 изображениями тех же людей, полученными из базы данных FERET. Авторы получили точность распознавания 97,3%, однако временные характеристики системы не указали.

Котрополос и др. в [100] говорят о существенном улучшении быстродействия системы, использующей модифицированную, по сравнению с DLA, архитектуру морфологических динамических связей. На базе данных, содержащей по 4 фотопортрета 37 человек, авторы получили коэффициент распознавания 96,3%. Время работы системы также не приводится.

В своих экспериментах Дюк и др. [72] использовали многомасштабное представление исходных изображений. Для каждого из фотопортретов строилась так называемая пирамида Гаусса, где самым нижним уровнем являлось оригинальное изображение, а верхние уровни получались путём низкочастотной фильтрации и уменьшением разрешения изображения в $(1/2)^n$ раз, где n – номер уровня пирамиды. Используя решётку-граф с 8×8 вершинами, в каждой из которых вычислялись 18 комплексных значений фильтров Гabora, авторы получали вектор, состоящий из $N_G = 1152$ признаков. Для минимизации признакового пространства применялись несколько подходов. На базе изображений в количестве 551 шт. (в среднем по 15 фронтальных изображений одного человека) коэффициент распознавания составил от 74,8 до 95,9% в зависимости от используемых в итоге признаков. Наилучший результат был получен при уменьшении разрядности признакового пространства с помощью метода линейного дискриминантного анализа Фишера [9], наихудший – метода главных компонент. О временных характеристиках работы алгоритма в работе не сообщается, но упоминается о высокой вычислительной стоимости подхода.

1.2.8 Применение фильтров Габора для распознавания

В 1946 г. Д. Габор [79] предложил подход, описывающий некоторую временную функцию, с одновременным учётом частоты сигнала, который впоследствии стал носить его имя. На основе этого подхода Дагманн [67] в 1988 году сформулировал двумерное преобразование Габора [131], которое применил для идентификации человека по изображению радужной оболочки глаза.

За последние 5-10 лет фильтры Габора – компактная форма упомянутого преобразования – стали широко используемым инструментом разработчиков систем обработки изображений. Так, например, они применяются для оценки симметричности текстур и их классификации [55, 73, 153], обнаружения движения на видеопоследовательностях [90, 142]. Упорядоченная группа таких фильтров, используемых с разными параметрами, часто называется Габоровскими *вейвлетами* [108].

Окрестность, окружающая пиксель $a(i,j)$ может быть описана значениями фильтров Габора, которые в совокупности формируют вектор-признак, характеризующий эту окрестность.

Общий вид фильтра может быть представлен как:

$$h_{\theta_0}(x, y) = g_{\theta_0}(x, y) \exp\{i(x \cos \theta_0 + y \sin \theta_0)\}, \quad (1.3)$$

где $g_{\theta_0}(x, y) = \frac{1}{\sqrt{\pi \sigma_x \sigma_y}} \exp\left\{-\frac{(x \cos \theta_0 + y \sin \theta_0)^2}{2\sigma_x^2} - \frac{(-x \cos \theta_0 + y \sin \theta_0)^2}{2\sigma_y^2}\right\}, \quad (1.4)$

$g_{\theta_0}(x, y)$ является двумерным гауссианом со стандартными отклонениями $\sigma_x \sigma_y$, повернутым на угол θ_0 .

Из анализа выражений (1.3) и (1.4) следует, что фильтр характеризуется следующими параметрами – радиусами эллипса σ_x и σ_y , углом ориентации θ . В зависимости от периода \cos и \sin у фильтров может быть различное количество пиков (локальных максимумов) t (рис.1.10). Результатом свёртки фильтра с изображением в некоторой точке (x, y) является комплексное число.

В настоящее время, судя по количеству появляющихся публикаций, фильтры Габора являются одним из наиболее исследуемых и обещающих подходов в области распознавания образов и обработки изображений. В ряде случаев они применялись и для идентификации человека по фотографии [101, 104, 170, 171]. Однако все возможности применения данного инструмента пока не раскрыты из-за большого количества возможных вариантов фильтров, а также их упорядоченных и произвольных наборов.

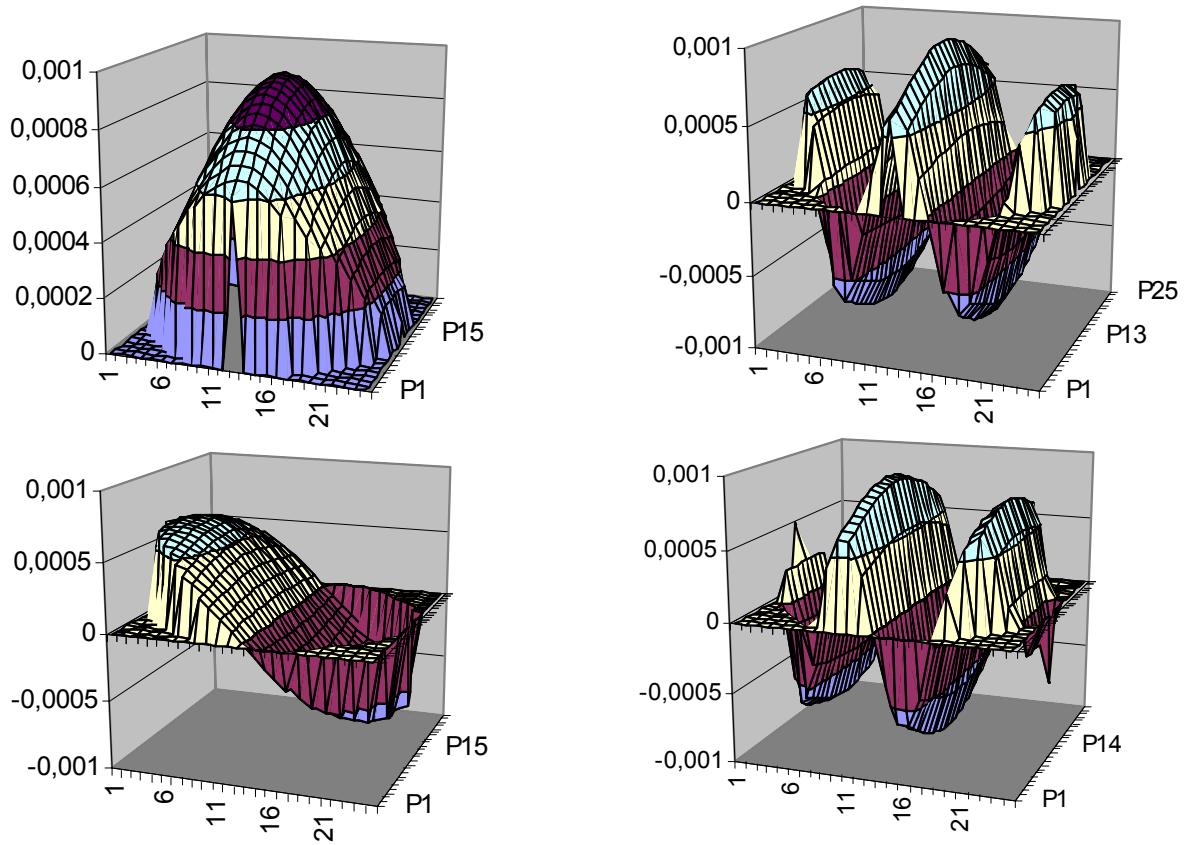


Рис.1.10. Два варианта фильтра Габора: функции, соответствующие действительной части фильтра (верхний ряд) и мнимой (нижний ряд). Приведённые фильтры отличаются лишь количеством пиков - один пик (слева) и три (справа). Остальные параметры: $\sigma_x=\sigma_y=13$, угол ориентации $\theta=0^\circ$, нач. угол сектора $\beta=0^\circ$, угол сектора $\alpha=360^\circ$, радиус отверстия кольца $r=0$

В большинстве работ [101, 104, 171] выбор параметров фильтров оставался на усмотрении авторов, и полновесного исследования о целесообразности использования того или иного фильтра для конкретной задачи не производилось. Кроме того, в области идентификации человека по фотопортрету фильтры Габора применялись в основном в комбинации с другими подходами, например: методом сопоставления графов и нейронными сетями.

Так, Вискотт [169] проводил эксперименты по определению трёх признаков тестируемого фотопортрета: пола (мужской/женский), наличия очков и бороды на лице. Система основывалась на сопоставлении графов (графы – прямоугольные сетки 7×4 узлов – позиционировались на портрет вручную), веса вершин которых вычислялись с помощью фильтров Габора. На базе изображений, состоящей из 111 фотопортретов (63% мужского пола, 30% из них имели бороду, а 44% – носили очки) коэффициент распознавания по упомянутым трём признакам составил 92, 96 и 94% соответственно.

Недостатком системы являлась её скорость – одно сопоставление графов, включая преобразование изображений и вычисление значений фильтров Габора, требовало около 1 мин машинного времени на SUN SPARCStation 10-512 с 50 MHz процессором (1996-1997 гг.).

1.3 Аппарат, применяемый при нормализации фотопортретов

На вход систем идентификации человека могут поступать фотопортреты, различающиеся по таким параметрам, как размер изображения, масштаб лица, наклон головы. Для того чтобы иметь возможность нормализовать фотопортрет, прежде всего требуется найти на изображении само лицо, а также центры зрачков – основные точки привязки, расстояние между которыми нормируется.

1.3.1 Методы поиска области лица на фотопортрете

В последнее время публикуется большое количество работ, посвященных поиску областей лиц на фотографиях [15, 57, 61, 63, 71, 76, 77, 78, 91, 93, 97, 109, 113, 123, 132, 133, 138, 151, 155, 156, 165, 166, 168, 174, 179]. Однако большинство из них рассматривает вариант одновременного поиска более чем одного лица на групповых фотографиях либо в видеопотоке.

Чаще всего для решения данной задачи применяются вариации метода сопоставления с эталоном [9]. Данный подход может быть достаточно эффективен в определённых условиях, но время работы велико из-за больших объёмов перерабатываемых данных. При использовании указанного метода для поиска лиц требуется оптимизировать количество анализируемых

частей изображения и установить непосредственные границы обнаруженного лица [64].

Существует множество альтернативных подходов. Так, например, в статьях [77, 78] описана система, построенная на нейронных сетях. Точность её распознавания составляет 84 - 86,5% в зависимости от качества исходных данных. Время обработки одной групповой фотографии с произвольным фоном от 1,3 до 2,9 с на DEC Alpha-333 MHz. В [132] авторы обрабатывают изображение с помощью методов математической морфологии, а затем анализируют совокупности локальных экстремумов, применяя метод главных компонент.

Поиск лиц путём анализа контуров изображения описывается в [112, 165]. Точность поиска таким методом колеблется от 80 до 87,5%.

В работе [82] характеристики выполнения этапа выделения лица не приводится, но итоговый коэффициент распознавания системы составила 95-97%, что говорит о хорошей предварительной обработке изображений. Кроме того, имеется также и несколько других подходов. Например подход, основанный на анализе по принципу «сверху-вниз», который использует несколько отдельных моделей лица, выстроенных по порядку от самой грубой до самой точной. Сначала изображение проверяется на предмет соответствия какой-либо его части самой грубой модели. Затем, если такой фрагмент найден, то он сравнивается с более точной моделью. Так продолжается до сравнения с самой точной моделью или же пока алгоритм не забракует фрагмент после проверки на соответствие какой-либо модели. Более сложные (объёмные) модели лица применялись в [133]. На базе изображений, состоящей из 88 фотографий с пёстрым фоном, была получена 91%-ная точность обнаружения искомого объекта.

Наиболее основательной, на взгляд автора, работой в данной области исследований можно назвать [166]. В двух экспериментах Вебер и др. испытывали свой подход на базе данных состоящей из более чем 5,5 тыс. и 32 тыс. изображений и достигли 96 – 99%-ной точности обнаружения. Суть их подхода состоит в сопоставлении частей пирамиды, слоями которой является исходное изображение, взятое с различным разрешением, с некоторым эталоном лица. Оценка близости и выбор областей-кандидатов производился с применением методов дискриминантного анализа. Время обработки одного изображения без построения пирамиды составляло около 0,3 с на ЭВМ Pentium II-400.

Подходы, использующие принцип «снизу-вверх», сначала проверяют изображение на наличие в нем совокупности признаков человеческого лица, затем группируют их в области-кандидаты для более тщательной

проверки [50, 176, 179]. Так, в работе [92] достигается 86%-ная достоверность обнаружения лиц при общем времени обработки одного изображения менее 5 секунд на ЭВМ “SUN Sparc 20”.

К данной области исследований можно отнести и работу Ву, Чена и Яшиды [174], результатом которой явилось построение системы проверки лиц, т.е. системы, которая устанавливает, действительно ли фрагмент изображения является человеческим лицом или же представлены просто похожие на него объекты. Когда данная система получает предполагаемое изображение лица, то из изображения сначала выделяются края. Затем система устанавливает приблизительные регионы поиска черт лица путём выявления зон, где средняя плотность горизонтальных краёв высока. При помощи метода интегральных проекций уточняются координаты каждой зоны. Полученные данные сравниваются с геометрической моделью лица и в итоге определяется является ли данная часть изображения действительно лицом или нет.

Основное новшество работы [174] состоит в том, что исходные изображения (в количестве 50 шт.) были цветными. В итоге 90% фрагментов изображений, на которых действительно были представлены человеческие лица, были распознаны корректно. Поиск области лица с использованием цветовых характеристик исследовался также и в [38, 109].

Несмотря на очевидный прогресс, вычислительная стоимость и соответственно время работы предлагаемых алгоритмов не позволяет их использовать в системах реального времени на этапе предварительной обработки. Кроме того, в большинстве случаев авторы решают задачу обнаружения лиц на групповых снимках либо на фотопортретах с однотонным фоном.

1.3.2 Методы поиска координат центров зрачков на фотопортрете

Большое количество исследований посвящено поиску различных черт лица на изображении [61, 71, 76, 97, 138, 155], и в первую очередь – центров зрачков [57, 75, 91, 92, 106, 117, 142, 180]. Одной из основополагающих работ по данной теме была статья Брунелли и Поджио [57]. Области изображений глаз они находят при помощи метода сопоставления с эталоном, предварительно обнаружив координаты окон, содержащих нос и рот, и таким образом сузив область поиска.

В статье [180] авторы предлагают геометрический подход к точному выделению зрачка и контура глаза. Процесс разбивается на три этапа. Первый - это выделение контура лица. Он основан на модели активных контуров, т.е. на постепенном сжатии контура области, заведомо

содержащей изображение лица, к контуру собственно лица. По получении координат контура лица из него на основе антропологической статистики извлекают два окна расположения глаз. На втором этапе, опираясь на априорные знания о том, что зрачок круглый и чёрный, строят гистограммы яркостей для этих окон, выбирают 25% пикселей с наименьшей яркостью и, используя методы кластерного анализа, определяют количество областей, к которым эти пиксели принадлежат, а также вычисляют центры этих областей. На последнем этапе для каждой области проводят k прямых линий из центра до границы, вычисляют для каждой линии градиент интенсивности и определяют несколько точек, находящихся на границах резкого перепада яркости для каждой области. После чего координаты полученных таким образом контурных точек подставляют в уравнение окружности, проверяя их принадлежность одной окружности. Из этих областей выбирается та, контурные точки которой менее других отклоняются от окружности. Она и считается зрачком.

Предлагаемый в работах [91, 92] подход определяет черты лица, используя геометрическую модель, которая конструируется согласно геометрическим соотношениям между частями лица. Благодаря этой модели трудности идентификации, вызванные эмоциональным выражением и ориентацией лица, могут быть легко преодолимы. На первой стадии работы алгоритма используется высокочастотная фильтрация, которая подчёркивает (усиливает) изображения. После чего производится бинаризация изображения, получившегося в результате фильтрации. Далее пиксели полученного на первом этапе бинарного изображения группируются и опознаются. Во-первых, ищутся глаза, которые должны находиться на одной линии, близкой к горизонтальной. Во-вторых, расстояние между глазами приблизительно равно двум длинам глаза. Учитывая эти, а также и другие признаки, авторы проверяют все пары регионов и выбирают одну пару, соответствующую упомянутым условиям. После чего, основываясь на антропометрической статистике, находят относительное местоположение рта и других частей лица. Скорость работы данного алгоритма на рабочей станции SUN-20 составила около 5 с на один портрет, причем более 70% времени тратилось на предварительную обработку.

В [75] предложен подход, основанный на комбинации изменяющихся эталонов с генетическими алгоритмами, для определения контуров глаз и губ. Эшме, Санкур и Анаrim сначала находят ось симметрии лица, затем производят предобработку (с использованием выравнивания гистограммы и модифицированной бинаризации), после чего вычисляют центры масс темных областей. Исходя из априорных знаний, производят отбор

получившихся регионов, вычисляют энергетические поля оставшихся областей, значения которых потом используются при проверке так называемых «хромосом» на близость с соответствующей областью. Хромосома представляет собой код нескольких величин, описывающих форму глаза или рта. Авторы определяют также правило спаривания хромосом, после применения которого исходные величины, представляемые каждой хромосомой, либо увеличиваются, либо уменьшаются, либо остаются неизменными. Первоначально генерируются 100 случайных хромосом, являющихся собой по сути 100 разных по форме и размеру видов эталона глаза или рта. После каждой операции спаривания части хромосомы - гены - подставляются в специальные функции подгонки, которые некоторым образом изменяют сравниваемый эталон предварительно полученной области глаза или рта. После каждой итерации выбираются четыре наиболее подходящие к оригиналу хромосомы и все остальные делятся на четыре группы, каждая из которых затем спаривается с одной из четырёх главных хромосом. После каждой итерации 10% слабых генов убирается и хромосомы пополняются случайными новыми. Таким образом, авторы добиваются постепенного и точного сжатия эталона вокруг области глаза или рта.

Самым быстрым, надёжным и оригинальным можно назвать алгоритм выделения зрачков на изображении, получаемом с видеокамеры, описанный в [117]. Он основан на известном эффекте “красных глаз”, возникающем при фотосъёмке человека со вспышкой. Разработанная Моримото и др. система подсвечивает лицо инфракрасным излучением в момент съёмки и выделяет зрачки человека по порогу после цветового разделения снимка. К сожалению, данный способ не может применяться для анализа паспортных фотопортретов, так как физические свойства объекта копии не присущи.

1.3.3 Определение условий получения изображения - ракурса съёмки

Определение ракурса съёмки – это мало исследованная, однако часто возникающая задача при идентификации человека по фотопортрету [53, 95, 105, 158, 159, 162, 177]. Для проведения более надёжного распознавания желательно, знать с одного ли ракурса были сфотографированы люди, и если нет – учитывать разницу для коррекции признаков.

В работе [53] авторы предлагают методику определения положения 3D объекта в пространстве, основываясь на его 2D проекции – изображении. Она основана на предположении, что если 3D объект представить в виде пространственного преобразования Фурье, то срез этого преобразования будет эквивалентен двумерному преобразованию Фурье проекции данного

объекта на плоскость среза. Для экспериментальной проверки использовалась построенная при помощи трёхмерного сканера “Cyberware range” модель лица. При определении положения головы человека по фотопортрету погрешности составили 4.05° , 5.63° и 2.68° соответственно при повороте относительно вертикали, наклоне и повороте головы относительно оси, проходящей через нос и затылок.

Оценивался также масштаб изображения лица – стандартное отклонение отношения вычисляемого по методике масштаба к реальному, составляло 0.0856. Данных о временных характеристиках алгоритма, а также о возможности применения данной методики/модели для изображений людей разных возрастов, рас и полов (опыт производился на изображениях одного человека) в статье не приводится.

1.4 Анализ систем идентификации человека по фотопортрету

Основными недостатками большинства систем, как экспериментальных, так и коммерческих, являются их неустойчивость к вариациям условий съёмки и неизбежному изменению внешности человека с течением времени. Так, по данным, опубликованным в отчёте по испытанию известных систем на базе изображений FERET [128], утверждается, что коэффициент распознавания тестировавшихся систем резко падал – до 60%, если на вход подавались фотопортреты (база данных из снимков 200 человек), сделанные с временной разницей в полгода.

Временные характеристики вышеописанных систем приведены для ПЭВМ класса Pentium III, 500 MHz.

Наиболее широко рекламированным за рубежом «решением проблемы распознавания лиц» является *FaceIt* – программное обеспечение фирмы *Visionics* [134]. Оно представляет собой набор алгоритмов, интегрируемых в продукты третьих компаний и предназначенных для решения задач ограничения доступа и поиска в базе данных. В зависимости от типа задачи, как варианты предобработки изображений, так и сами алгоритмы идентификации могут меняться, а поэтому конкретные характеристики системы неизвестны. В основе распознавания лежит так называемый анализ локальных признаков лица (Local Feature Analysis). О его сути не сообщается. Известно лишь о предварительной настройке алгоритмов с помощью нейронных сетей для определения значимости каждого локального признака. Из распространяемых демоверсий программ видно, что глаза на изображениях выделяются весьма приблизительно, а также что подход, лежащий в основе, сильно уменьшает разрешение исходных изображений и не определяет точные координаты каких-либо черт или точек лица.

Утверждается, что изображение лица может кодироваться вектором длиной от 100 до 300 байт. Однако в предлагаемых примерах уже нормализованные и использовавшиеся для идентификации изображения занимали от 3100 до 3600 байт. Время полной обработки лица неизвестно, однако сообщается, что поиск области лица на изображении 400×300 занимает от 0.3 до 1 с. Время формирования вектора признаков – 1с. В информации разработчика говорится о менее чем 1% вероятности общей ошибки, т.е. как первого, так и второго рода. Однако по вышеупомянутым результатам теста FERET данная информация не подтверждается.

Достоинства системы: оценка качества исходного изображения и вероятности неправильного результата, обнаружение произвольного количества лиц на фотографии.

Недостатки: требуется представление нескольких изображений тестируемого лица для надёжной верификации; алгоритмы чувствительны как к примитивным изменениям внешности (причёска), так и к возрастным. Изменение фона и условий освещения отрицательно влияет на коэффициент распознавания.

Фирма eTrue [135] предлагает как набор алгоритмов, посвященных обработке и идентификации фотопортретов, в виде некоторого ядра для интегрирования в программные продукты покупателя TrueFaceEngine™, так и готовое решение для идентификации человека по базе данных на основе представленной фотографии либо кадра, полученного с видеопотока – TrueFaceID™. Судя по информации, предоставляемой фирмой, в основе указанных разработок лежит комбинация метода «собственных лиц» и анализа локальных признаков лица, а настройка системы производилась с использованием нейросетевых технологий. Размер нормализованных портретов, по которым производится сравнение, варьируется от 500 до 1500 байт. Однако для большей надёжности рекомендуется использовать последний вариант. Утверждается, что верификация, т.е. сравнение двух нормализованных изображений занимает менее 1 с, а поиск в базе данных намного быстрее – 500 сравнений в секунду. Утверждается, что TrueFaceID™ была победителем теста подобных систем по версии ICSA (Международной ассоциации по компьютерной безопасности) и «в зависимости от вида приложения точность распознавания может быть выше, чем 99,95%» [135].

Характерным примером состояния дел в коммерческом использовании систем идентификации человека по фотопортрету можно назвать FaceCam™TA – продукт компании Vision Sphere Technologies. Он представляет собой пропускное устройство, снабжённое камерой, микрофоном, цифровой клавиатурой и считывателем информации с магнитных карт.

Верификация пользователя производится сразу по нескольким критериям: информация с карты сравнивается с цифровым паролем, изображением с видеокамеры и характеристиками речи человека. Устройство работает только в тандеме с ПЭВМ класса Pentium II и выше. Гарантируется нулевая вероятность ошибки первого или второго рода. Пропускная способность – от 8 до 10 человек в минуту. К достоинствам предлагаемого фирмой решения относится встраиваемая защита от представления фотографии вместо живого человека, т.е. либо используется стереоизображение, либо анализ нескольких кадров из видеопотока. Фирмой предлагаются также иные варианты системы распознавания человека по фотопортрету для ограничения доступа к компьютеру It's Me™ и более сложный вариант - UnMask™ . Информация о принципах работы и характеристиках систем не предоставляется. Говорится лишь о 1% вероятности ошибки любого рода.

Из отечественных экспериментальных систем идентификации человека известна разработка сотрудников и студентов МГТУ им. Баумана под названием "ВидеОко" [17]. Система предназначена для решения задачи ограничения доступа. При настройке системы необходимо создать список людей, имеющих право доступа, и ввести в систему изображения каждого из них. Максимальное количество классов (людей) и минимальное количество изображений класса, необходимых для надёжной работы системы, не известно. Программа функционирует в режиме реального времени и автоматически реагирует на изменение (движение) во входном потоке видеоданных. В среднем на обработку изображения 256×256 и распознавание уходит от 4 до 5 с (в случае трёх классов). Если человек не относится к одному из классов, то он считается неопознанным и ему отказывается в доступе. Предположительно используется метод «собственных лиц». Количество случаев правильной идентификации прямо пропорционально количеству изображений для каждого из зарегистрированных людей: при трёх портретах – 30-35%, при 20 – 80-95%. Однако повышается время распознавания (до 8 с).

Достоинства системы: в экспериментах с тремя людьми (правда, с сильно различающейся внешностью) ни разу не была допущена ошибка второго рода, т.е. доступ незарегистрированному человеку не был разрешён; быстрое нахождение центров зрачков – за 0.5 – 1с;

Недостатки: в 25% случаев система не находит зрачки; сильно зависит от условий освещённости и изменения внешности человека, в частности изменение причёски уменьшало количество случаев правильного распознавания до 15-20% от общего числа попыток; чувствительна к мимике – с улыбкой не распознала ни разу. Система требует доработки, так как не

отличает статическое изображение от живого человека и может быть обманута путём поднесения к камере фотографии зарегистрированного пользователя.

Следует отметить, что существуют ещё 3-4 фирмы, которые заявили о своих разработках в указанной области. Однако информация об их разработках в специализированной прессе не публиковалась. В Интернете также не доступна. Проанализировав доступную по теме информацию, можно заключить, что в настоящий момент существует стабильный интерес на рынке к автоматическим системам идентификации и верификации на основе фотопортретов. Предлагаемые же решения рекомендуется использовать как вспомогательное (дополнительное) средство защиты либо вкупе с обязательным контролем со стороны человека.

1.4.1 Достоинства и недостатки перечисленных методов

Основными недостатками практически всех упомянутых методов является чувствительность к вариациям условий съёмки распознаваемых/сравниваемых людей, возрастным изменениям, а также высокая вычислительная стоимость алгоритмов. Однако, если последний недостаток является всё более слабеющим на фоне постоянно повышающейся мощности ЭВМ фактором, то первые два сильно сдерживают широкое применение уже имеющихся систем и наработок. Так, Вискотт со своими коллегами [171] заявляет: «Эластичное сопоставление графов хорошо работает, если изображение распознаваемого лица похоже по параметрам (ракурс, освещение) на какие-то из имеющихся в базе и уже описанных в виде графа, но метод сбоят, если лица выглядят достаточно разными».

Метод «собственных лиц» и анализ локальных признаков хотя и являются на сегодняшний день наиболее исследованными и применяемыми практически, но они не позволяют получать компактных векторов признаков, по которым можно производить поиск, либо требуют постоянной дополнительной настройки и тренировки, что в ряде случаев не представляется возможным.

По сравнению со всеми остальными методами нейросети являются наиболее приближённой к человеческому зрению моделью восприятия и обработки информации. Однако так как обучение человеческого мозга происходит на протяжении всей жизни и, тем не менее, при распознавании люди иногда ошибаются, данную модель нельзя назвать совершенной. То же самое можно сказать и о нейросетях - для обучения/настройки систем на их основе требуется представительная выборка распознаваемых объектов, а

также необходимы постоянное уточнение и дообучение в процессе функционирования системы.

1.4.2 Выбор подхода для решения задач контроля удостоверений личности и поиска фотопортретов заданного человека в базе данных

Названные задачи представляются более трудными, чем контроль доступа. Однако с коммерческой точки зрения удовлетворительным на данный момент являлось бы и частичное их решение, т.е. вариант, при котором система лишь выдаёт своё предварительное с оговоренной точностью заключение, – является ли предъявителем документа тем человеком, на чьё имя выдан документ или нет, а в случае с поиском в базе данных – выборку наиболее похожих на заданного индивида. Окончательное решение всё равно будет приниматься человеком – экспертом.

Наиболее проверенным за длительное время практики подходом является геометрический - на основе антропометрических характеристик лица. Он использовался многие десятилетия в криминалистике, однако выделение точек и все вычисления производились вручную.

С точки зрения автора самым главным достоинством геометрического подхода является принципиальная возможность распознавания с помощью него человека по фотопортретам, сделанным с большой разницей во времени.

В то же время основными трудностями по его применению являются автоматическая локализация антропометрических точек лица с требуемой точностью, оптимизация набора точек и вычисленных на его основе признакам.

Возможными путями преодоления данных трудностей могут быть использование фильтров Габора как для локализации определённых точек, так и для вычисления признаков, а также автоматическая оптимизация наборов полученных признаков.

1.5 Выводы

В первой главе выполнен анализ проблемы «распознавания лиц», основных подходов и методов, применяемых для её решения, алгоритмов поиска лица и центров зрачков на фотопортретах. Проведён анализ и сравнение характеристик известных на данный момент коммерческих и экспериментальных систем идентификации человека по фотопортрету. Основные выводы могут быть сформулированы следующим образом:

1. Проблема идентификации по фотопортретам может быть рассмотрена с позиции трёх ключевых задач - в зависимости от области применения систем и возможности использования различных методов для их реализации: контроль удостоверений личности, контроль доступа, поиск фотопортретов заданного человека в базе данных. Задачи автоматического контроля удостоверений личности и поиска фотопортретов в базе данных являются наименее исследованными (п.1.1).

2. Установлено, что для решения задачи контроля удостоверений личности либо сравнения фотопортретов без использования априорной информации наиболее устойчивым к возрастным изменениям, освещению, ракурсу съемки, причёске, макияжу и т.п. является геометрический подход. Основными препятствиями для его машинной реализации являются проблема автоматической локализации антропометрических точек лица и выбор признаков на их основе. Для поиска указанных точек, а также для определения признаков в их окрестностях целесообразно использовать фильтры Габора с различными комбинациями параметров (п.1.2).

3. На момент анализа, информации об алгоритмах, позволяющих обрабатывать слабоконтрастные портреты с произвольным фоном в режиме реального времени, в научной и широкой печати найдено не было (п.1.3).

4. Выявлено отсутствие систем, решающих задачу контроля удостоверений личности. Задача поиска фотопортретов в базе данных исследовалась для малых по объёму баз, – содержащих не более 10^4 записей. Подавляющее большинство систем неустойчивы к возрастным изменениям внешности распознаваемого человека – коэффициент распознавания падает до 60%, если исходные фотопортреты имеют разницу во времени более полугода (п.1.4).

Результаты проведённых исследований опубликованы в работе [30].

ГЛАВА 2

ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ИСХОДНЫХ ИЗОБРАЖЕНИЙ И ВЫДЕЛЕНИЕ ПРИЗНАКОВ

В данной главе рассматриваются преобразования исходных изображений для последующего использования в системах автоматического контроля документов и поиска фотопортретов в базах данных.

2.1 Требования к исходным изображениям

Цифровые фотопортреты одного и того же человека могут отличаться по таким критериям, как условия съёмки (расстояние до камеры, ракурс, освещённость), эмоциональное состояние (выражение лица – нейтральное, напряжённое, весёлое, грустное и т.д.), параметры оцифровки (разрешение, количество уровней яркости, контраст, насыщенность (интенсивность)), шум (вносимый как во время фотопроцесса, так и на стадии оцифровки изображения). Благодаря огромному количеству возможных вариаций исходных представлений портрета одного и того же человека автоматическое распознавание может быть крайне затруднено, если не сформулировать ограничения, налагаемые на них.

Для алгоритмов и подходов, разработанных в рамках данной работы, такие требования приведены ниже. В случае если на вход алгоритма поступают изображения, не удовлетворяющие сформулированным ниже ограничениям, результат работы алгоритмов может быть некорректен.

Для задачи автоматической верификации на исходном фотопортрете не должно быть изображено более одного человека. Центральная часть лица (брови - нос) не должна закрываться волосами, очками или иными предметами.

Левой и правой частями лица на фотопортрете будут считаться области изображения, соответственно находящиеся слева и справа относительно центральной оси симметрии лица с точки зрения человека, смотрящего на фотопортрет.

Ограничения приведены по группам в порядке убывания значимости фактора для автоматического распознавания:

2.1.1 Размер лица на изображении

Лицо на фотопортрете не должно занимать меньше 20% и больше 60% от площади всего снимка. Иными словами, высота головы должна быть не менее 25% и не более 125% высоты исходного изображения. Глаза не должны располагаться ниже горизонтальной линии, делящей снимок

пополам. Область всего лица (ширина и высота) должна быть не менее чем 60×80 и не более чем 210×280 пикселей.

2.1.2 Ракурс

Нормальным положением головы, в рамках данной работы, будет считаться такая её ориентация на фотопортрете, при которой выполняются следующие условия:

ось центральной симметрии лица совпадает с вертикальной осью ОУ и проходит через точку ЕМ - середину отрезка (ЕL,ER), соединяющего центры левого и правого зрачков (рис. 2.1).

соотношение отрезка (ЕМ,NN) – расстояния от линии уровня глаз до кончика носа – к отрезку (ЕМ,LN) – расстоянию от линии уровня глаз до линии смыкания губ – при нейтральном выражении лица будет 0,66.

Нулевым положением лица будет называться слегка наклоненная, относительно нормального положения, позиция, при которой $(EM,NN)/(EM,LN) = 0,69$. При этом надкозелковые точки (правая O_1 и левая O – см. рис.2.20, стр. 77) будут находиться на уровне глаз, т.е. – на прямой (ЕL,ER). Данное положение будет являться крайним при наклоне головы, что соответствует углу приблизительно в 15° .

Подъём головы вверх, относительно нормального положения головы, допускается на угол до 15° ($(EM,NN)/(EM,LN) = 0,55$).

Отклонение головы при повороте вправо/влево относительно нормального положения допускается также в пределах $\pm 15^\circ$. Другими словами, если у человека оба уха не скрыты причёской и видны полностью или частично, то изображение удовлетворяет данному условию, в противном случае – нет.

2.1.3 Освещённость

Лицо человека должно быть освещено достаточно равномерно, так как резкие тени, затрудняющие визуальную локализацию уголков глаз, крыльев носа, контуров губ и т.п., недопустимы.

2.1.4 Яркостные характеристики цифрового изображения

Фотопортреты, поступающие на вход предлагаемых алгоритмов, должны быть полутооновыми с уровнями яркости от 64 до 256. Яркостные характеристики оцениваются с помощью следующих величин [2, 21]:

- средняя яркость M ;
- среднее квадратичное отклонение σ ;
- медиана Md .

Чтобы исключить нежелательное воздействие фона при анализе, указанные выше величины высчитываются по центральной части лица. При этом удовлетворяющими требованиям изображениями будут считаться те, у которых:

- 1) $0,45 \geq \sigma/M \geq 0,2;$
 - 2) $80 \geq M \geq 130;$
 - 3) $M/Md \leq 0,95.$
- (2.1)

2.1.5 Эмоциональное выражение лица

На большинстве фотопортретов для документов, удостоверяющих личность, выражение лица нейтральное (либо сосредоточенно-напряжённое) и доброжелательно-весёлое (с небольшой улыбкой). Система для автоматического контроля документов должна уметь работать с такими изображениями.

2.2 Предварительная обработка исходных изображений

Целью предварительной обработки является приведение исходных изображений к единому масштабу, яркостным характеристикам и, желательно, единому ракурсу (одноименные черты лица должны располагаться в областях портрета с одинаковыми координатами).

Последовательность шагов такой обработки может быть следующей:

- поиск области лица на изображении;
- обнаружение центров зрачков на портрете;
- поворот изображения (если требуется) – центры зрачков должны находиться на горизонтальной прямой;
- масштабирование (нормализация всех портретов по определённому расстоянию между зрачками);
- кадрирование (вырезание прямоугольной области с заданными размерами из полученного в результате предыдущих шагов изображения);
- выравнивание яркостных характеристик исходного изображения, т.е. применение различных фильтров, изменяющих контраст, интенсивность и т.д., в зависимости от начальных параметров (2.1);
- обработка фотопортрета в соответствии с требованиями дальнейших этапов идентификации, например выделение перепадов яркости (с помощью оператора Собеля, Дериша [98, 126]) либо подчеркивание некоторых черт лица (например, используя оператор “High Boosting” [92]).

Пример изображений до нормализации и после показан на рис. 2.2.

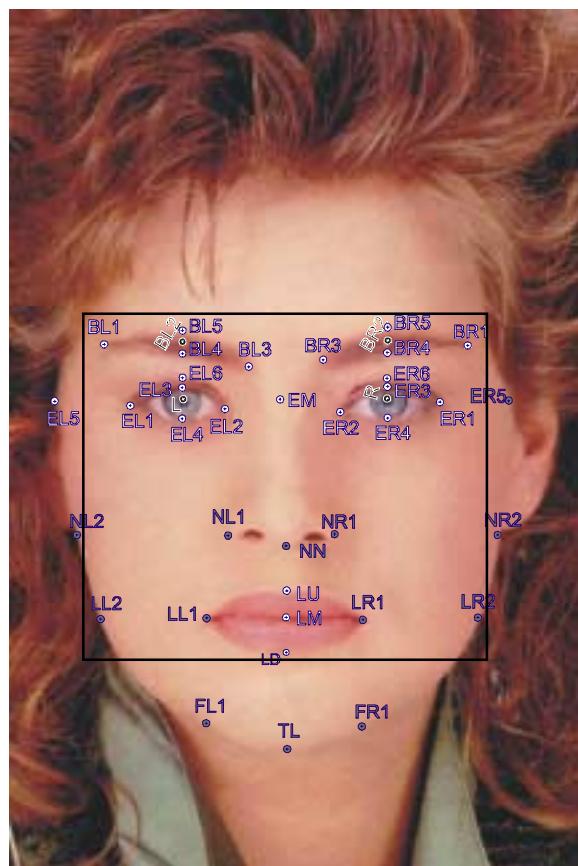


Рис. 2.1. Точки и область лица, используемые при идентификации человека по фотопортрету



Рис. 2.2. Исходные фотопортреты: а) предъявителя паспорта; б) фотография на документе. Результаты предобработки: в) нормализованное изображение (а); г) соответствует нормализованному (б).

2.2.1 Алгоритм поиска области лица на фотопортрете

Несмотря на очевидный прогресс, вычислительная стоимость и соответственно время работы рассмотренных в разд. 1.3.1 алгоритмов не позволяет их использовать в системах реального времени на этапе предварительной обработки. Кроме того, в большинстве случаев авторы решают задачу обнаружения лиц на групповых снимках либо на фотопортретах с однотонным фоном.

В рамках настоящей работы предлагается оригинальный алгоритм поиска области лица, построенный на основе метода сопоставления с эталоном. Алгоритм не имеет указанного выше недостатка - большой вычислительной стоимости и позволяет работать всей системе в режиме реального времени. Рис. 2.3 иллюстрирует шаги алгоритма.

Голова человека представляет собой объект с достаточно устойчивыми внешними признаками (в основном это брови, глаза и нос). Верхняя и нижняя части головы могут существенно отличаться у людей разного возраста, пола, расы в основном за счёт наличия/отсутствия волосяного покрова и, в меньшей степени, из-за особенностей анатомического строения. Следует отметить также, что цвет/оттенок кожи и вариации в освещении могут в той или иной мере «маскировать» лицо, особенно на пестром фоне. Дополнительные затруднения вызывает отсутствие априорной информации, например о размерах головы на фотопортрете.

Для преодоления указанных сложностей на первом шаге выполняется операция выделения краёв на изображении [60, 137]. Пусть $A(x,y)$ – изображение размером $m \times n$ пикселей. Под *свёрткой изображения* $A(x,y)$ с некоторой *маской* $H(x,y)$ размером $k \times l$ будем подразумевать преобразование $A(x,y) \rightarrow A'(x,y)$, при котором каждый элемент изображения $A'(x,y)$ будет получен следующим образом:

$$a'_{ij} = \sum_{i'=1}^n \sum_{j'=1}^m a_{i-i'j-j'} \times h_{i'j'} \quad (2.2)$$

Хорн в своей книге [46] даёт следующее определение: «Маской называется набор используемых для свёртки весовых коэффициентов, расположенных таким образом, чтобы отражать пространственные отношения между элементами, к которым они применяются».

За последние 30 лет разработано много операторов выделения краёв на изображении [51, 89, 98, 126, 129, 130, 161]. Специфика данной задачи требует применения быстрого алгоритма, дающего надёжные результаты, в

том числе и на слабоконтрастных изображениях. После сравнения ряда алгоритмов был выбран оператор Собеля.

Оператор Собеля [98] представляет собой свёртку исходного изображения с двумя масками S_h и S_v размером 3×3 по отдельности и суммирование результатов (2.3):

$$S_h = \begin{array}{|c|c|c|} \hline -1 & -2 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 2 & 1 \\ \hline \end{array} \quad S_v = \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array}$$

$$A' = \sqrt{(S_h \otimes A)^2 + (S_v \otimes A)^2}. \quad (2.3)$$

1-й шаг: В результате применения оператора Собеля получается изображение A' с выделенными краями (рис. 2.3,в). Применив инверсию к A' , получим изображение

$$\bar{A}' = G - A', \quad (2.4)$$

где G – константа, равная количеству уровней яркости изображения, в данном случае $G = 255$.

На **2-м шаге** \bar{A}' бинаризируется в соответствии с методом Отсу [124] $A'' = Ot(\bar{A}')$:

$$A'' = Ot(\bar{A}') = \begin{cases} 0, & \bar{a}'_{ij} \leq T^{opt}; \\ 1, & \bar{a}'_{ij} > T^{opt}; \end{cases} \quad (2.5)$$

где T^{opt} – оптимальный для \bar{A}' порог бинаризации, вычисленный в соответствии с дискриминантным критерием максимальной межклассовой (объект/фон) дисперсии уровней яркости [36,98].

3-й шаг. Для A'' строится пирамида, состоящая из множества изображений $Mz = \{A''_1, A''_2, \dots, A''_t\}$, где A''_t – изображение с размерами $(m/c) \times (n/c)$, $c = 2t+p$ (в экспериментах $p=14$), каждый пиксель a''_t вычисляется в соответствии с выражением

$$a''_t(i, j) = \frac{1}{c^2} \sum_{x=i-c/2}^{i+c/2} \sum_{y=j-c/2}^{j+c/2} a''(x, y). \quad (2.6)$$

Пример такого множества для $t=5$ приведён на рис. 2.3,д-к.

4-й шаг: Размеры изображений A''_t отличаются, соответственно и размеры лица (в пикселях) будут варьироваться. При сопоставлении эталона

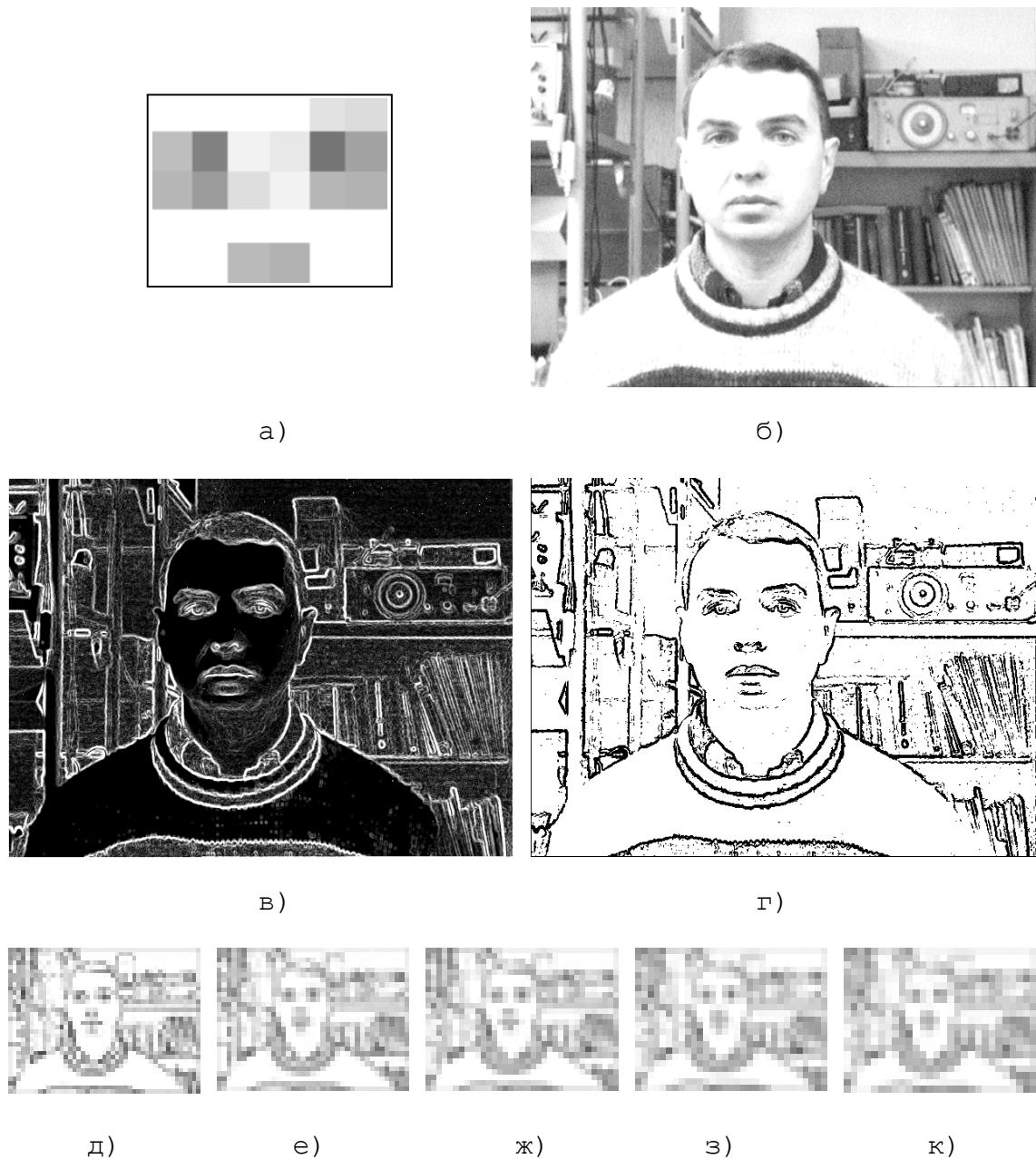


Рис. 2.3. Поиск области лица на фотопортрете: а) эталон 5×6 пикселей; б) исходный фотопортрет \mathbf{A} ; в) результат применения оператора Собеля \mathbf{A}' ; г) инверсия и бинаризация по методу Отсу \mathbf{A}'' ; д)-к) мозаичное представление \mathbf{A}'' , $n = 14, 16, 18, 20, 22$, соответственно

фиксированного размера (в результате экспериментов выбран эталон С (6×5) (рис. 2.3, а)) с каждым из A''_t разница в масштабе не играет существенной роли: там, где мозаичное изображение области лица будет совпадать по размеру с эталоном, функция сравнения $F(A''_t, C)$ будет иметь минимальное значение. В качестве $F(A''_t, C)$ была выбрана городская метрика [36, 89, 98, 126, 161]:

$$F(A''_t, C) = d(A''_t, C) = \sum_{\substack{x= \\ i-Tm/2}}^{i+Tm/2} \sum_{\substack{y= \\ j-Tn/2}}^{j+Tn/2} |a''_t(x, y) - C(x, y)|. \quad (2.7)$$

В результате, оценив каждую из областей изображения, где может находиться лицо, выбирается та, при сравнении с которой $F(A''_t, C)$ даёт минимальный отклик. Если разница в значениях оценки F для нескольких областей минимальна или отсутствует, то проверяются соответствующие им области на смежных мозаиках A''_{t-1} , A''_{t+1} и вычисляется комплексная оценка. Примеры результатов работы алгоритма приведены на рис. 2.4.

Временные характеристики и результаты сравнения предлагаемого алгоритма с иными приведены в главе 4.

2.2.2 Алгоритм поиска координат центров зрачков

В разделе 1.3.1 описан ряд подходов, использовавшихся исследователями для решения названной задачи. Однако их существенными недостатками являются в ряде случаев высокая вычислительная стоимость и, соответственно, большое время работы либо неудовлетворительное качество работы на слабоконтрастных изображениях и цифровых изображениях с низким разрешением. В процессе исследований был разработан оригинальный алгоритм, избавленный от указанных выше изъянов.

Имея координаты центральной области лица A_F , район поиска центров зрачков значительно сужается.

1-й шаг: Производится обработка области лица с помощью тех же инструментов, что и в алгоритме 2.2.1: оператора Собеля, инверсии и метода Отсу. Итоги бинаризации на данном шаге, как правило, отличаются от результата применения того же оператора на 3-м шаге алгоритма 2.2.1, так как метод Отсу учитывает гистограммы, а обрабатываемые изображения разные по размеру. В результате преобразования $A_F(x, y) \rightarrow A'_F(x, y) \rightarrow \bar{A}'_F(x, y) \rightarrow A''_F(x, y)$ получается бинаризованное изображение с подчеркнутыми чертами лица и осветленными областями волосяного покрова (рис. 2.5).



Рис. 2.4. Пример результатов работы алгоритма поиска области лица на фотопортретах: на паспортных и на поступающих с видеокамеры изображениях

В работе Брунелли и Поджио [57] для нахождения местоположения носа и рта были использованы интегральные проекции - горизонтальная и вертикальная. Вычислялись они следующим образом: если $I(x, y)$ - исходное изображение, то вертикальная проекция изображения $I(x, y)$ в квадрате с угловыми координатами $[x_1, y_1]$ и $[x_2, y_2]$ может быть определена как

$$V(x) = \sum_{y=y_1}^{y_2} I(x, y). \quad (2.8)$$

Аналогично вычисляется горизонтальная проекция:

$$H(y) = \sum_{x=x_1}^{x_2} I(x, y). \quad (2.9)$$

2-й шаг: С помощью интегральных проекций центральной части лица выделяются несколько полос A^B_i шириной приблизительно 4-5 диаметров зрачка, которым соответствуют значительные впадины на графике $f = V(x)$. В таких полосах находятся глаза, а иногда брови и нос (рис. 2.5). Глубина впадин на графике $f = V(x)$, достаточная для принятия решения об анализе соответствующей ей области, устанавливается эмпирически. При соблюдении ограничений раздела 2.1 таких полос выделяется не более трёх, в 70% случаев – одна, в 25% - две $A''_F \rightarrow U\{A^B_1, A^B_2, \dots A^B_i\}$.

3-й шаг: Каждая из выделенных полос $U\{A^B_1, A^B_2, \dots A^B_i\}$ подвергается анализу с помощью преобразования Хоха [23, 52, 87]:

$$H_{iR}^B = A_i^B \otimes O_R,$$

где

$$O_R(m, n) = \begin{cases} 1, & (R-0,5)^2 < m^2 + n^2 < (R+0,5)^2; \\ 0. & \end{cases} \quad (2.10)$$

Результатом данного преобразования является отображение $A^B_q \rightarrow H^B_{qR}$, при котором каждому чёрному пикслю $a^B_q(i, j)$ ставится в соответствие окружность/полуокружность радиуса R пикселей с центром в $h^B_{qR}(i, j)$. Для каждой из областей проводится несколько подобных преобразований с различными $R = \{9, 12, 15, 17\}$. Примеры результатов подобных преобразований для $R=9$ и $R=12$ приведены на рис. 2.6.

Геометрический смысл данного преобразования состоит в том, что при наличии в области A^B_q элементов в виде полуокружности или окружности радиуса R пиксель, соответствующий её центру, будет иметь наибольшее значение на множестве H^B_{qR} (рис. 2.7).

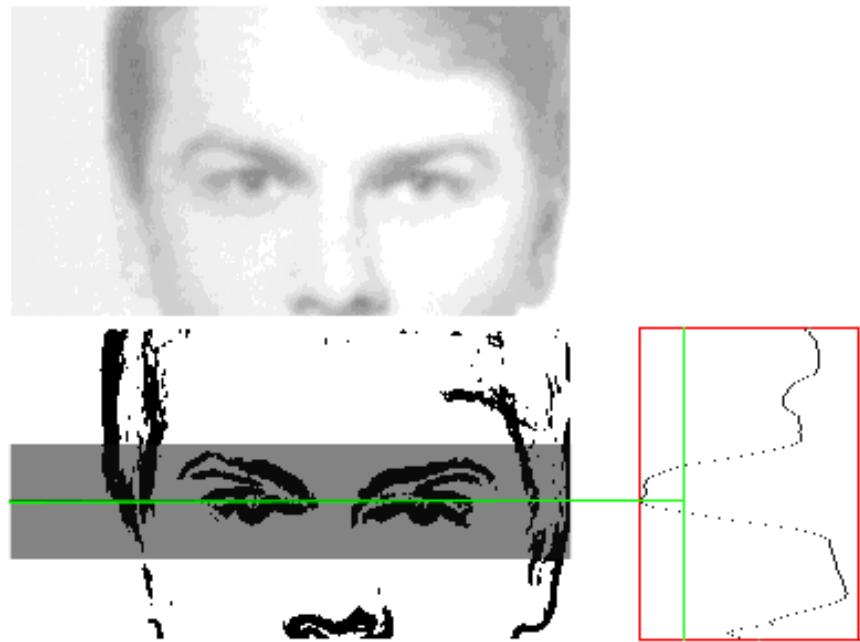


Рис. 2.5. Область лица на исходном изображении (вверху) и после выделения краёв и бинаризации (внизу). Справа приведён график вертикальной интегральной проекции, серым выделена найденная по ней область глаз

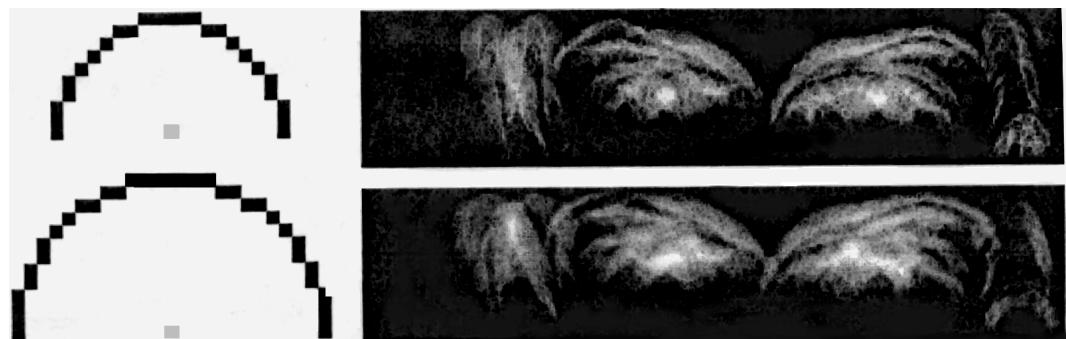


Рис. 2.6. Примеры масок с радиусами $R=9$ и $R=12$ и представлений области поиска зрачков в виде поверхностей H^B_{qR} при использовании указанных масок

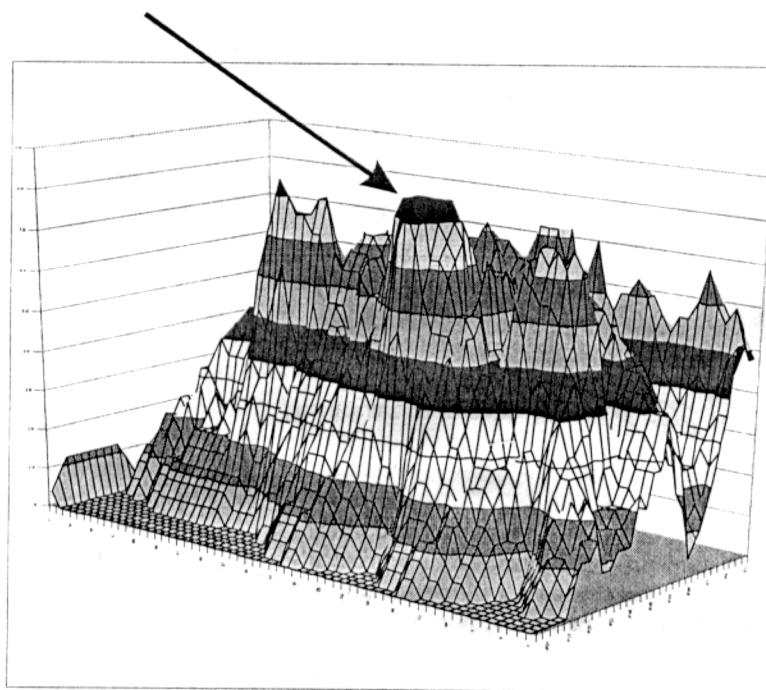


Рис. 2.7. Трёхмерное представление области левого глаза (рис.2.5) в виде поверхности H_{qR}^B при $R=9$. Стрелка указывает местоположение зрачка, которому соответствует глобальный экстремум



Рис. 2.8. Примеры вариантов изображений глаз на исходных фотографиях

Учитывая информацию о том, что у мужчин расстояние между зрачками составляет 63 ± 7 мм, а у женщин – 61 ± 6 мм, при диаметре радужки $d = 12$ мм [180], на изображении H^B_{qR} ищутся пики (локальные максимумы), удалённые друг от друга по горизонтали на расстояние $4,5 - 6d$. Из полученных пар-претендентов выбирается одна с наибольшим суммарным значением на различных множествах H^B_{qR} .

Благодаря независимости от возможных вариантов изображений глаз (рис. 2.8) и постепенному сужению области поиска зрачков данный алгоритм позволяет надёжно находить центры зрачков на слабоконтрастных изображениях (рис. 2.9).

2.2.3 Выбор алгоритма поворота цифровых изображений

Для нормализации исходных фотопортретов требуется алгоритм поворота изображения на произвольный угол α . Для выбора алгоритма, вносящего в исходное цифровое изображение минимальное количество искажений, был выполнен сравнительный анализ разных алгоритмов поворота.

Поворот фигуры в евклидовой геометрии является изометрическим преобразованием, т.е. сохраняет расстояния и углы между точками. Если произвольную фигуру повернуть на угол $+\alpha$, а затем на $-\alpha$, то результат будет идентичен исходной фигуре. Однако не все свойства евклидовой геометрии имеют место в дискретном пространстве. Если цифровое изображение рассматривать как дискретную поверхность, то её поворот на произвольный угол не является инвариантным из-за необходимости аппроксимации этой поверхности при вычислении её нового дискретного образа. Из-за дискретности пространства не сохраняются углы и расстояния между пикселями и, как следствие, несколько искажаются фигуры. Так, например, в случае последовательного поворота на угол $+\alpha$, а затем на $-\alpha$ идентичность между исходной и результирующей фигурами наблюдаться не будет. Исключением являются углы кратные 90° .

При выполнении дискретного поворота возникают две основные проблемы. При вычислении новых координат отображения пикселей в конечной фигуре будут наблюдаться пробелы (рис. 2.10), т.е. появляются точки, требующие аппроксимации поверхности. С другой стороны, при повороте на 45° расстояния, вычисленные в пикселях дискретной плоскости, уменьшаются в $\sqrt{2}$ раз, т.е. при построении новой дискретной плоскости некоторые пиксели приходится отбрасывать.

Для решения описанных выше проблем предложены различные алгоритмы поворота цифровых изображений. Условно их можно разделить на три класса в зависимости от сложности аппроксимации. Самым простым

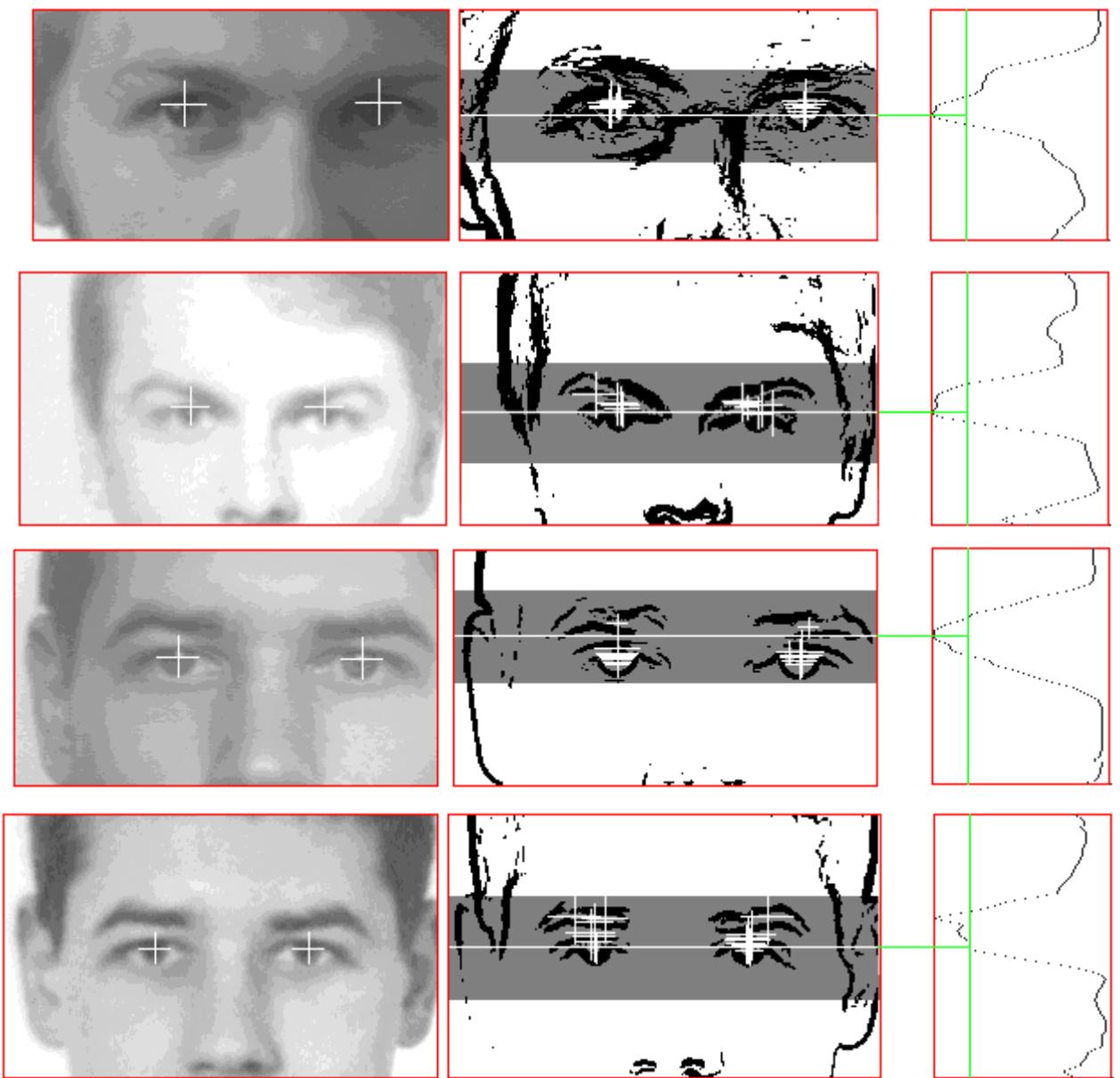


Рис. 2.9. Примеры результатов работы алгоритма поиска центров зрачков. Справа показаны интегральные проекции областей лица, по центру – обрабатываемые области (крестиками отмечены потенциальные центры зрачков), слева – результат поиска

является метод «ближайшего соседа», вторым по сложности вариантом – линейная/билинейная интерполяция. Более точного вычисления значений пикселей преобразованного изображения добиваются с помощью полиномов различной степени, а также сплайн-функций.

Для заполнения пробелов в повёрнутом изображении Паркер [126] предлагает использовать билинейную интерполяцию.

Абламейко и Лагуновский в [1] предлагают сначала преобразовывать пространство исходного представления изображения в пространство с изменённым шагом дискретизации. В своём алгоритме они уменьшают указанный шаг в $\sqrt{2}$ раз и присваивают каждому новому пикселью значение ближайшего к нему исходного пикселя. Затем изображение поворачивается в пространстве с уменьшенным шагом и выполняется обратное преобразование в пространство с исходным шагом дискретизации. Фактически появляющиеся пробелы заполняются значениями ближайших пикселей исходного изображения.

Оуэн и Македон в работе [125] обосновывают метод поворота цифровых изображений путём трёх последовательных сдвигов строк, столбцов и затем снова строк изображения. Декомпозицию данного преобразования $R(\alpha)$ можно выразить в матричной форме следующим образом:

$$R(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} = \begin{bmatrix} 1 & -\operatorname{tg}\alpha/2 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ \sin \alpha & 1 \end{bmatrix} \times \begin{bmatrix} 1 & -\operatorname{tg}\alpha/2 \\ 0 & 1 \end{bmatrix}. \quad (2.11)$$

В рамках настоящей работы с точки зрения минимизации вносимых искажений в исходное изображение и возможности применения на стадии предварительной обработки были исследованы пять алгоритмов поворота цифровых изображений. Три из них: алгоритм поворота сдвигом по Оуэну и Македону, алгоритм Лагуновского и алгоритм, использующий билинейную интерполяцию [126], – были запрограммированы автором на основе информации из указанных источников. Кроме того, был принят во внимание тот факт, что при обработке изображений в интерактивном режиме чаще всего используются такие коммерческие пакеты, как COREL Photo-Paint® и ADOBE Photoshop®. Несмотря на отсутствие информации об алгоритмах, используемых в названных программах, результаты поворота тестовых изображений при помощи данных пакетов также приводятся ниже.

Для объективной оценки вносимых алгоритмами искажений использовались четыре меры сравнения: два варианта коэффициента корреляции (2.12) и (2.13), среднеквадратичная ошибка (MSE) (2.14) и мера D, подробно рассмотренная в разделе 3.2. Все меры были нормализованы

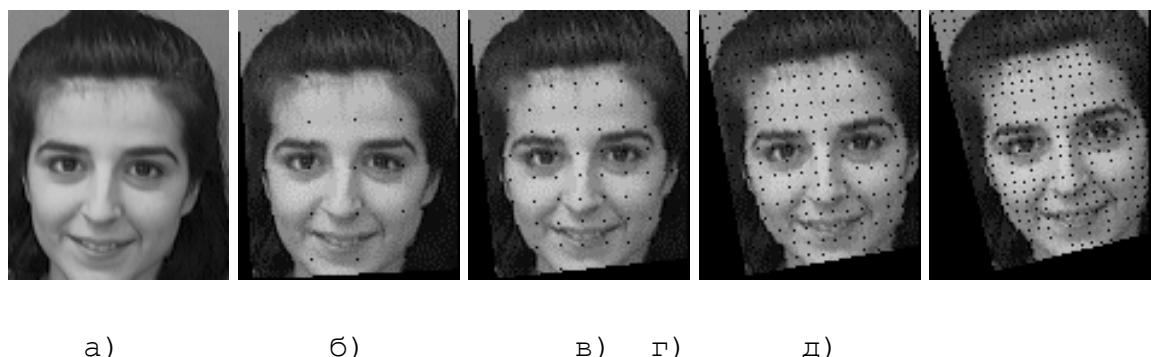


Рис. 2.10. Пример поворота изображения без коррекции на произвольный угол: а) исходное изображение; б) повёрнутое на 3° ; в) – на 6° ; г) – на 10° ; д) – на 14°

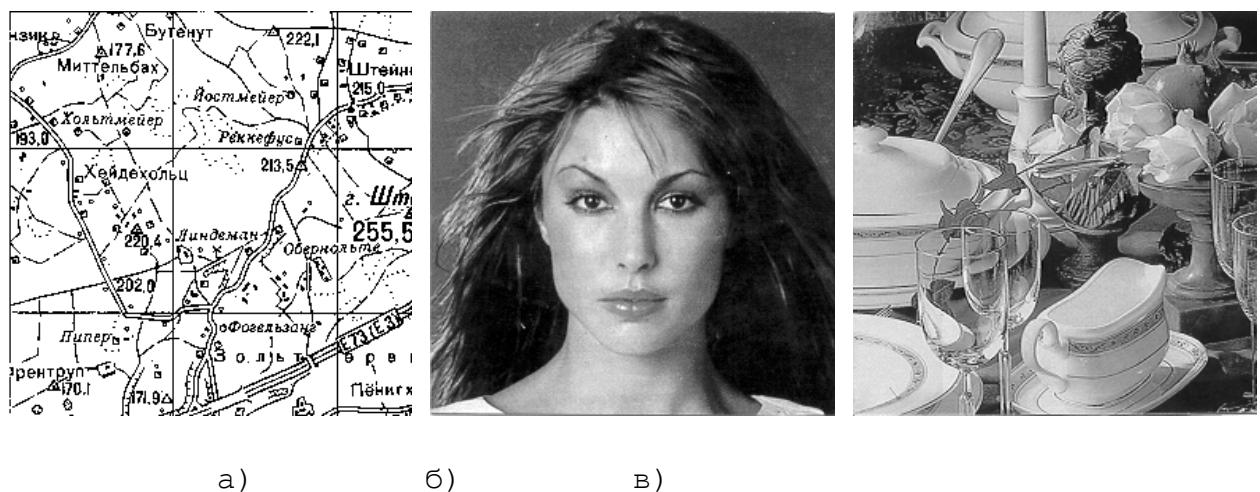


Рис. 2.11. Пример изображений а) \mathbf{M}^0 ; б) \mathbf{G}^0 и в) – \mathbf{N}^0 , на которых производилась проверка исследуемых алгоритмов

таким образом, чтобы значение 0.0 соответствовало случаю идентичных изображений A и B , а 1.0 - максимально различных, т.е. полностью белого и полностью чёрного изображений.

$$Cor1(A, B) = 1 - \frac{\sum_{i,j=1}^N a_{ij} b_{ij}}{\sqrt{\sum_{i,j=1}^N a_{ij}^2 \sum_{i,j=1}^N b_{ij}^2}}, \quad (2.12)$$

$$Cor2(A, B) = 1 - \frac{\sum_{i,j=1}^N (a_{ij} - a)(b_{ij} - b)}{\sqrt{\sum_{i,j=1}^N (a_{ij} - a)^2 \sum_{i,j=1}^N (b_{ij} - b)^2}}, \quad (2.13)$$

здесь a и b – средние значения изображений А и В соответственно, $N \times N$ – размер изображений.

$$MSE(A, B) = \frac{1}{NG} \sqrt{\sum_{i,j=1}^N (a_{ij} - b_{ij})^2}. \quad (2.14)$$

В качестве тестовых были выбраны три существенно различающиеся между собой изображения: участок карты (**M**), портрет девушки (**G**) и натюрморт (**N**). Для получения их цифровых представлений использовался сканер “HP ScanJet 5P” и программное обеспечение, поставляемое с ним в комплекте.

Фотографии размещались на сканере с поворотом против часовой стрелки на угол от 0° (исходное положение) до 45° , с шагом в 5° , и затем сканировались. Поворот осуществлялся относительно левого верхнего угла фотографии. Полученные таким образом изображения обозначены как **M**^ρ, **G**^ρ и **N**^ρ, где ρ - угол поворота в градусах. Необходимо отметить, что оригиналы **G** и **N** являлись цветными, а **M** – бинарной фотографией. В результате сканирования изображения **M**^ρ оставались бинарными, а **G** и **N** преобразовывались в полутонаовые **G**^ρ и **N**^ρ с 256 уровнями яркости. Таким образом было получено 30 эталонных изображений (размером 255×255 пикселей), повёрнутых вручную на различные углы ρ . Изображения **M**⁰, **G**⁰ и **N**⁰ показаны на рис. 2.11.

Исследуемые алгоритмы поворота были обозначены следующим образом: Al_1 - алгоритм с использованием билинейной интерполяции, Al_2 - алгоритм из пакета ADOBE Photoshop®, Al_3 - алгоритм из пакета COREL Photo-Paint®, Al_4 - алгоритм Лагуновского, Al_5 - алгоритм поворота сдвигом по Оуэну и Македону. Исходные изображения **M**⁰, **G**⁰ и **N**⁰ после поворота при

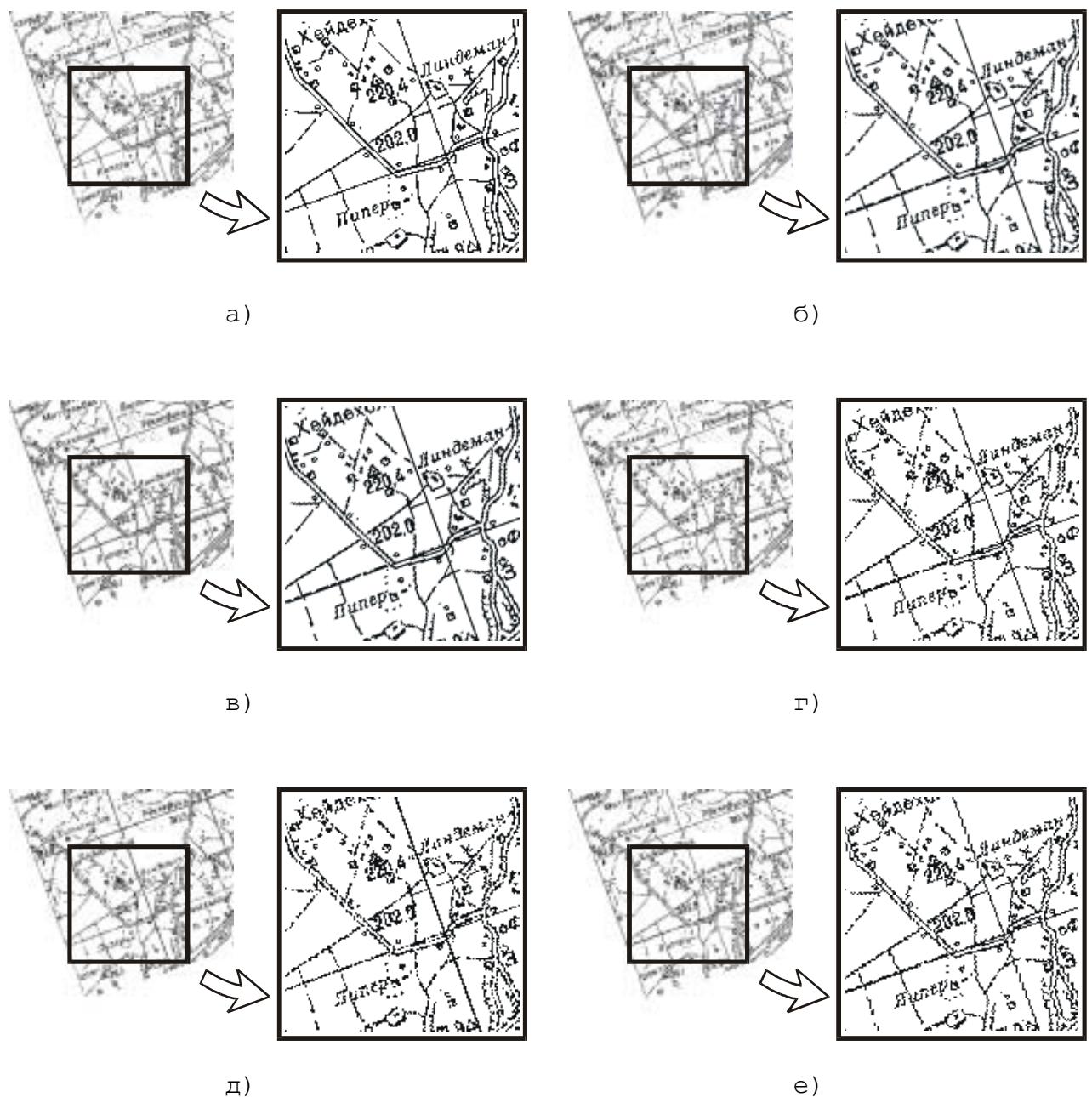


Рис. 2.12. Примеры изображений и их увеличенных фрагментов, полученных в первом опыте ($\rho=20^\circ$): а) повёрнутого «вручную» – M^0 ; б) с использованием билинейной интерполяции – M_1^0 ; в) с помощью Adobe Photoshop® – M_2^0 ; г) – Corel Photopaint® – M_3^0 ; д) алгоритмом Лагуновского M_4^0 ; е) сдвигом M_5^0

помощи k -го алгоритма на угол ρ обозначались как \mathbf{M}^ρ_k , \mathbf{G}^ρ_k и \mathbf{N}^ρ_k . Изображения \mathbf{M}^ρ_k , \mathbf{G}^ρ_k и \mathbf{N}^ρ_k , дополнительно повёрнутые на угол $-\rho$ тем же самым алгоритмом k , обозначались $\mathbf{M}^{(\rho)}_k$, $\mathbf{G}^{(\rho)}_k$ и $\mathbf{N}^{(\rho)}_k$ соответственно.

Было выполнено два эксперимента. В первом сравнивалось сходство изображений \mathbf{M}^ρ_k , \mathbf{G}^ρ_k и \mathbf{N}^ρ_k с \mathbf{M}^ρ , \mathbf{G}^ρ и \mathbf{N}^ρ соответственно. Целью второго была оценка искажений, вносимых в изображения в результате многократного поворота тем или иным алгоритмом: $\mathbf{M}^{(\rho)}_k$, $\mathbf{G}^{(\rho)}_k$ и $\mathbf{N}^{(\rho)}_k$ сравнивались с \mathbf{M}^0 , \mathbf{G}^0 и \mathbf{N}^0 . В обоих случаях для оценки сходства изображений использовались указанные выше меры.

Вносимые искажения наиболее заметны в случае поворота бинарного изображения. На рис. 2.12 представлены фрагменты изображений карты, повёрнутых на угол $\rho = 20^\circ$ «вручную» (\mathbf{M}^{20}) и при помощи исследуемых алгоритмов (\mathbf{M}^{20}_k). Анализируя полученные изображения, можно заметить, что билинейная интерполяция и алгоритм, используемый в ADOBE Photoshop®, изменяют большинство пикселей на краях объектов поворачиваемых изображений, а не только те, которые соответствуют возникающим пробелам. Иными словами, указанные алгоритмы осуществляют избыточную интерполяцию. Остальные алгоритмы Al_3 - Al_5 не вносят в изображение пиксели с новыми уровнями яркости. Данные выводы подтверждает и анализ результатов сравнения изображений \mathbf{M}^ρ с \mathbf{M}^ρ_k , представленных на рис. 2.13. Так мера D сразу выделяет алгоритмы Al_1 , Al_2 в отдельную группу с большим отклонением от эталонных изображений \mathbf{M}^ρ .

Коэффициенты корреляции $Cor1$, $Cor2$ и среднеквадратичная ошибка MSE показывают несколько критических углов, при повороте на которые в изображения вносится наибольшее количество искажений: для изображений \mathbf{M} и \mathbf{N} – это $\rho = 10^\circ$ и $\rho = 20^\circ\text{-}25^\circ$, для \mathbf{G} эти углы слегка смещены – $\rho = 15^\circ$ и $\rho = 30^\circ\text{-}35^\circ$ (см. прил. 1 – рис. П.1.1, П.1.2).

Фрагменты изображений и их гистограммы, полученные в ходе второго опыта – путём двойного поворота на ρ и $-\rho$ градусов, представлены на рис. 2.14. Можно заметить, что билинейная интерполяция привносит больше искажений в изображение, чем алгоритм ADOBE Photoshop®, а изображения $\mathbf{M}^{(15)}_3$ и $\mathbf{M}^{(15)}_5$, повёрнутые при помощи COREL Photo-Paint® и сдвига соответственно, визуально отличаются меньше от исходного, чем $\mathbf{M}^{(15)}_4$ – результат работы алгоритма Лагуновского. Такие же выводы можно сделать при анализе графиков сравнения мерой D изображений \mathbf{M}^ρ с $\mathbf{M}^{(\rho)}_k$ (рис. 2.15), несмотря на то, что коэффициенты корреляции $Cor1$, $Cor2$ и среднеквадратичная ошибка MSE ставят алгоритм Лагуновского на более высокое место, чем Al_3 , Al_5 . Однако имеется одно исключение: все меры показывают

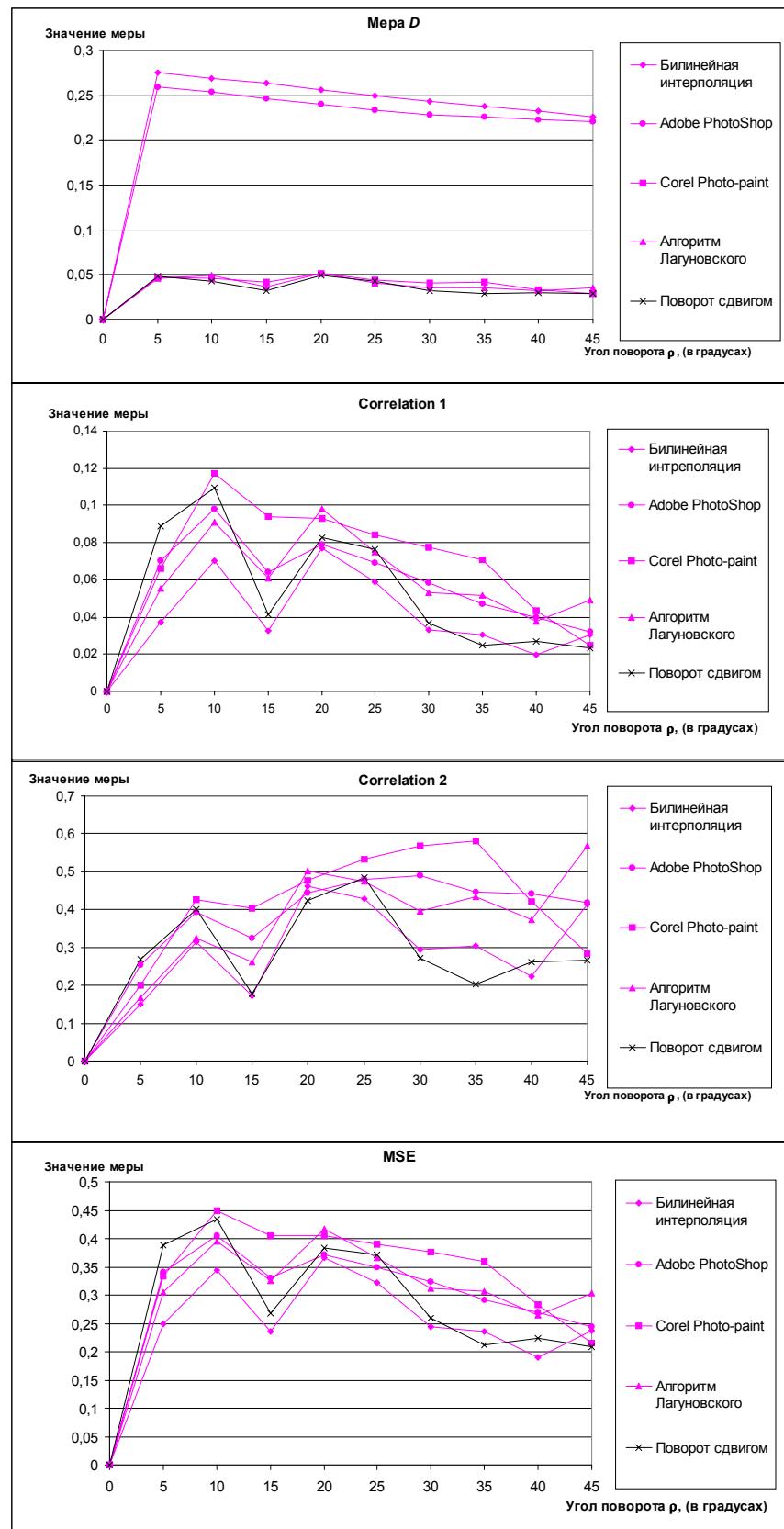


Рис. 2.13. Результаты сравнения изображений \mathbf{M}^0 и \mathbf{M}_k^0 с помощью меры D, коэффициентов корреляции Cor1 и Cor2, оценки среднеквадратичной ошибки MSE

лучшие оценки для поворота сдвигом и COREL Photo-Paint®, если $\rho = 45^\circ$. Алгоритм Лагуновского показывает худшие результаты для бинарных изображений, но лучше обрабатывает полутонаовые **G** и **N**, на которых много однородных областей.

Из экспериментов следует, что алгоритм пакета COREL Photo-Paint® лучше других обрабатывает изображения различных типов. Несмотря на то, что он не ориентирован, как алгоритм ADOBE Photoshop®, на минимизацию среднеквадратичной ошибки, он практически не изменяет гистограмму исходного изображения и даёт в результате лучшее для последующего автоматического анализа изображение. На основе полученных изображений можно сделать вывод о том, что алгоритм ADOBE Photoshop® учитывает психофизические особенности человеческой зрительной системы, из-за чего повёрнутые с его помощью изображения визуально кажутся более похожими на эталоны, чем другие.

На основании проведённого исследования можно сделать следующие выводы: на этапе предварительной обработки изображений с точки зрения минимизации вносимых в них искажений лучше использовать алгоритм поворота сдвигом по Оуэну и Македону. Он быстр, прост и даёт хорошие результаты. Обрабатывать изображения в интерактивном режиме для последующего автоматического анализа лучше с помощью пакета COREL Photo-Paint®.

2.2.4 Масштабирование и кадрирование исходных изображений

Более 70 лет в качестве нормы для приведения фотопортретов к единому масштабу в криминалистической экспертизе [11] применяется расстояние между зрачками человека. Это объясняется его неизменностью в течение жизни человека, а также надёжностью выделения на фотопортретах с различной чёткостью и яркостью.

В рамках настоящей работы не проводились исследования искажений, вносимых в изображение при использовании различных алгоритмов масштабирования. Это объясняется тем, что в любом случае искажения будут присутствовать, так как выполняется аппроксимация уровней яркости изображения. Кроме того, на данном этапе обработки разница между алгоритмами масштабирования не играет существенной роли, так как коэффициент масштабирования, как правило, мал и цифровое изображение лица представляет собой более-менее гладкую поверхность.

В результате экспериментальной проверки было выявлено, что лучшие для работы алгоритмов следующего этапа распознавания результаты даёт расстояние между зрачками $ED = 40$ пикселей. Поскольку в 98% случаев

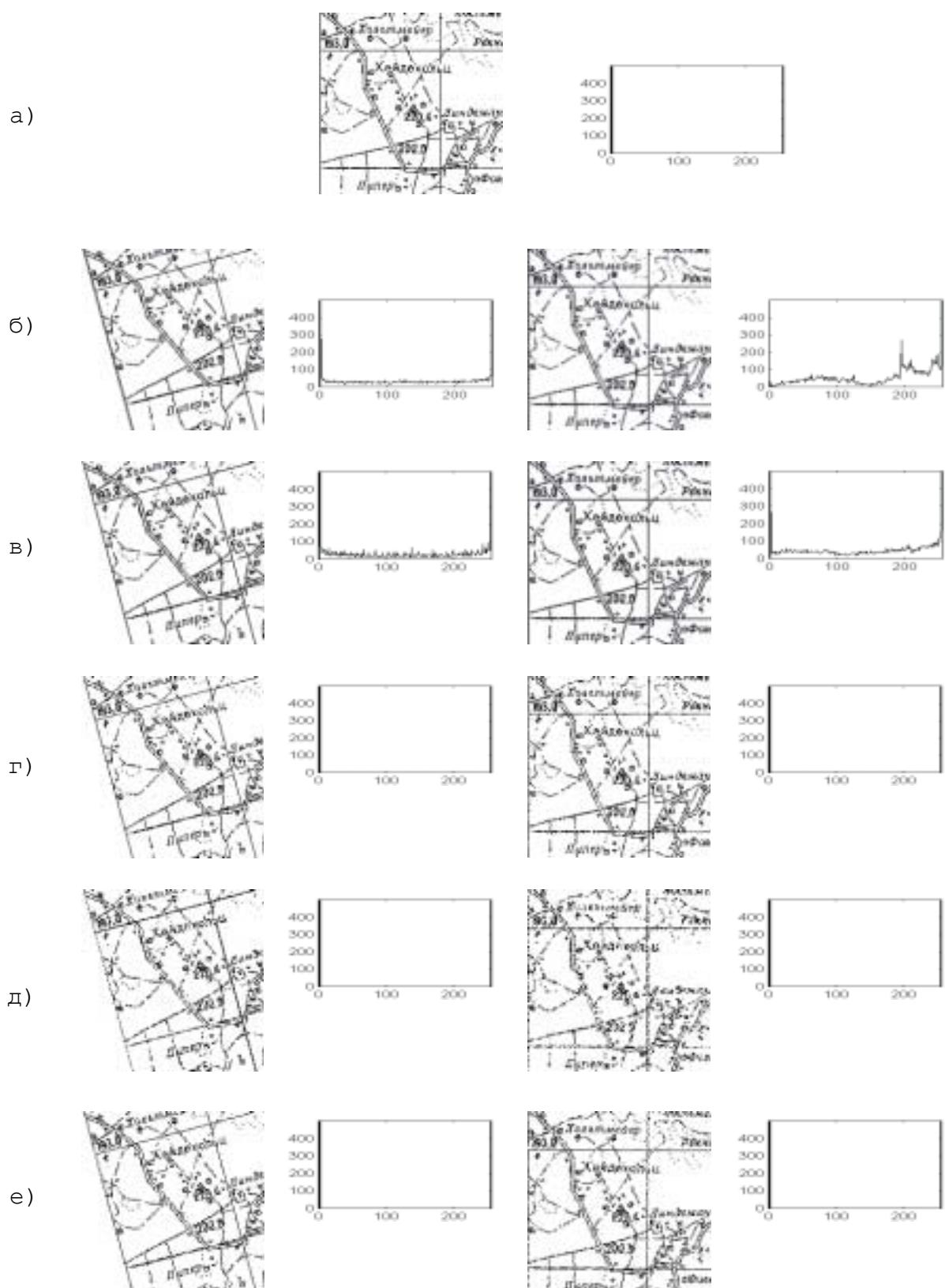


Рис. 2.14. Примеры фрагментов изображений M^0 , $M^{ρ_k}$, $M^{(ρ)_k}$ и их гистограмм, использовавшихся во втором опыте: а) исходного, M^0 . Рисунки б)-е): в левом столбце фрагменты $M^{ρ_k}$, $ρ=15^\circ$; в правом - $M^{(ρ)_k}$, значение k возрастает от 1 (б) до 5 (е) соответственно

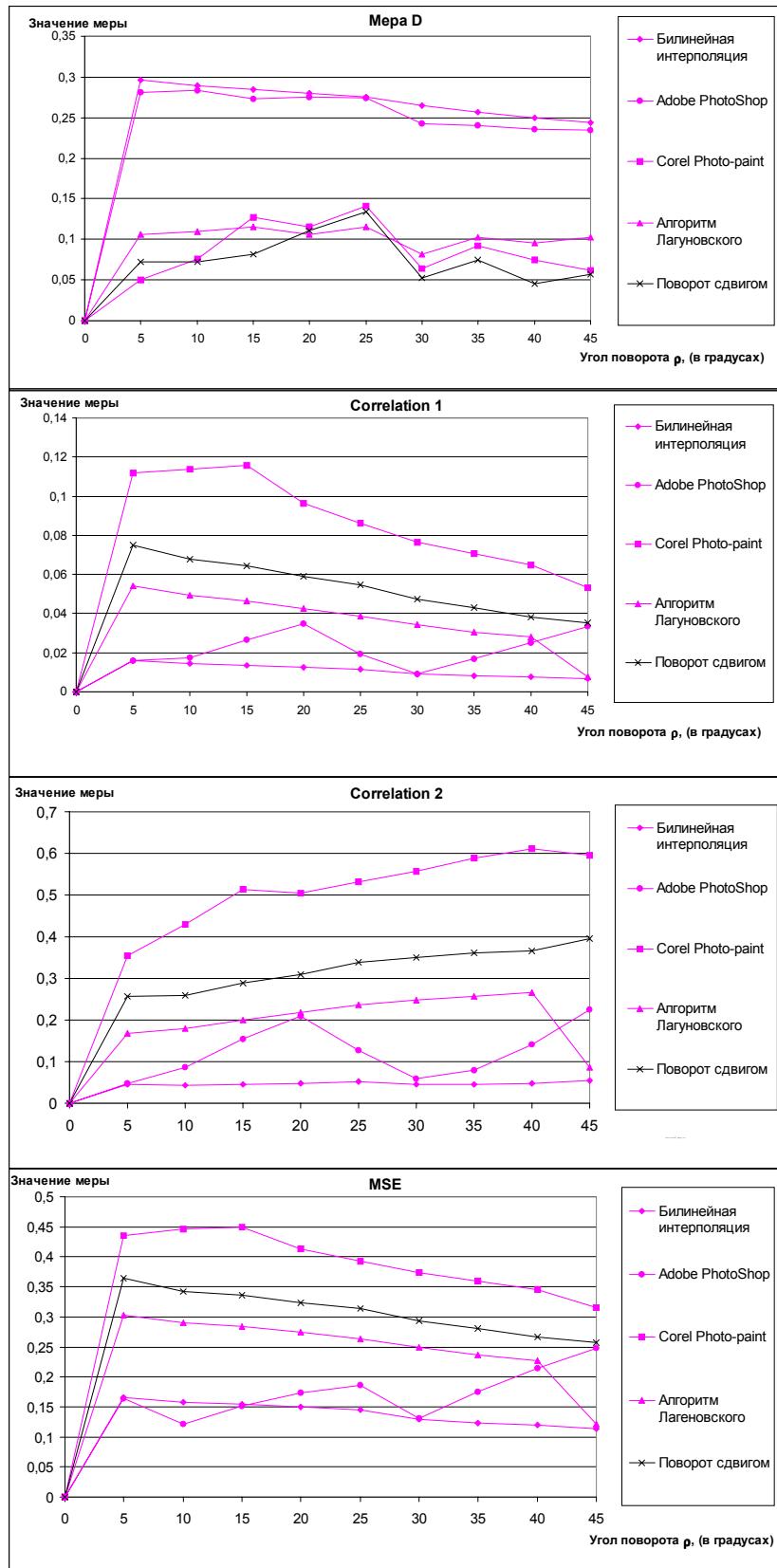


Рис. 2.15. Результаты сравнения изображений \mathbf{M}^0 и $\mathbf{M}^{(\rho)}$ с помощью меры D, коэффициентов корреляции Cor1 и Cor2, оценки среднеквадратичной ошибки MSE

исходные изображения требуется уменьшать, чтобы привести указанное расстояние к норме, был выбран алгоритм масштабирования с использованием билинейной интерполяции [126].

В зависимости от выбора алгоритма сравнения размер нормализованных изображений мог меняться. Основной целью при этом было получение изображения с минимальным размером, но сохраняющее максимальное количество информации для идентификации.

В связи с этим были исследованы различные варианты кадрирования изображений (рис.2.16). В дополнение к ним были проверены и иные решения, содержащие лишь полоску области глаз, половину лица (с учётом симметрии) и др.

2.2.5 Фильтрация и улучшение яркостных характеристик исходных изображений

В силу различия технических характеристик устройств ввода и оцифровки информации, разницы в условиях съёмки и типах представляемых документов часто требуется повысить качество исходных изображений. Пример фотопортрета, полученного с помощью специального считывателя паспортов (Мод. 7001, НПП «Regula»), приведён на рис. 2.17,*a*. Центральная часть лица, на основе которой вычисляются основные признаки для идентификации, показана на рис.2.17,*b*.

В настоящей работе было исследовано несколько десятков различных комбинаций фильтров и подходов с целью улучшения, поступающих на вход системы цифровых изображений. Ниже описаны результаты исследований оригинальных подходов и тех, которые дали хорошие результаты и были впоследствии использованы при построении экспериментальной системы идентификации.

Для принятия решения, следует ли улучшать поступившее изображение либо нет, анализировалась гистограмма центральной части лица рис.2.17,*b*, 2.18,*a,д*.

Установлено, что цвет зрачков во время фотографирования человека, даже при ярком свете, очень близок к чёрному цвету. Используя этот факт, а также определив координаты центров зрачков (EL, ER) на предыдущих этапах, проверялись значения яркости соответствующих пикселей. Если значения указанных пикселей были отличны от 0, т.е. не соответствовали чёрному цвету, то производились следующие действия.

Выбиралось меньшее значение яркости G_m из $\{G_{ER}, G_{EL}\}$. Находилось максимальное значение яркости G_M из всех пикселей центральной части

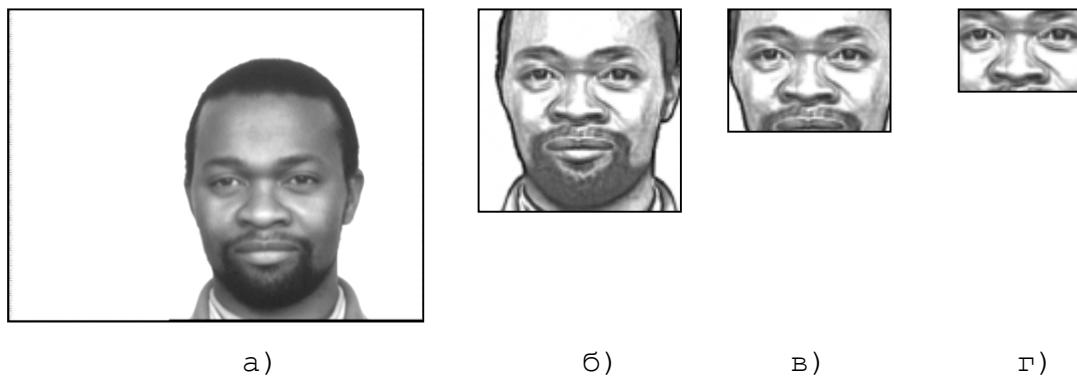


Рис. 2.16. Примеры вариантов кадрирования:
а) исходное изображение 384*288; б) нормализованное 100*100; в) нормализованное 80*60; г) нормализованное 60*40. Расстояние между зрачками $ED = 40$ пикселей

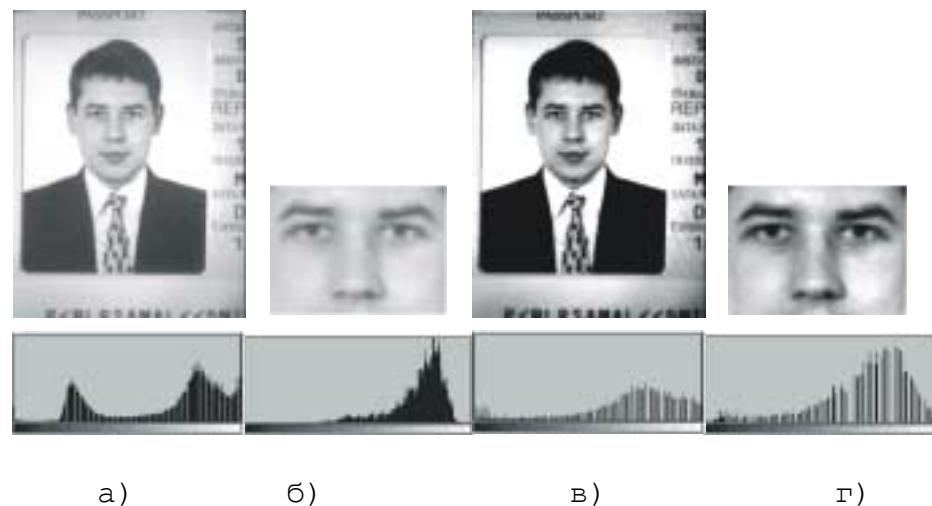


Рис. 2.17. Примеры улучшения контраста исходных изображений:
а) исходное изображение; б) центральная часть лица исходного фотопортрета; в) - г) изображения (а) и (б) после улучшения контраста. Снизу приведены их гистограммы

лица ($m \times n$) $G_M = \max\{G_i\}$, $1 \leq i \leq m \times n$. Новые значения пикселей G'_i вычислялись в соответствии с формулой (2.15).

$$S_i = (G_i - G_m) \frac{255}{(G_M - G_m)},$$

$$G'_i = \begin{cases} 0, & S_i \leq 0; \\ S_i, & 0 < S_i < 255; \\ 255, & S_i \geq 255. \end{cases} \quad (2.15)$$

Эксперименты показали, что, несмотря на отдельные положительные результаты, такие как, например, рис. 2.18,*b,e*, недостатками данного подхода являются чувствительность к шуму и относительная неустойчивость результатов из-за разницы оттенков и цвета кожи у разных людей. Метод не позволяет прогнозировать изменение среднего значения яркости M и стандартного отклонения σ после обработки портрета.

Похожий метод, основанный на том же принципе «растяжения гистограммы», не дал устойчивых положительных результатов. Мы исходили из того, что гистограммы исходных изображений имеют очень узкую форму, и разброс значений пикселей практически укладывается в интервал $[M-\sigma, M+\sigma]$. В результате $G_m = M-\sigma$, $G_M = M+\sigma$, а растяжение производится в соответствии с формулой 2.15. Некоторые из обработанных таким образом изображений улучшались, но основными результатами были портреты, похожие по своим характеристикам на рис. 2.18,*e, ж*.

Как один из вариантов предобработки изображений был протестирован подход “High Boosting” [91, 92]. Его отличительной чертой является совмещение информации, извлекаемой как из карты краёв, так и из обычного представления изображения. Для вычисления значения яркости обработанного пикселя с координатами (x,y) используется свёртка с маской, центрированной в данной позиции и имеющей переменный коэффициент w , вычисляемый в соответствии с (2.16):

-1	-1	-1
-1	w	-1
-1	-1	-1

$$w = [z(1-s) + (1-z)(1-g)] * (f - 8) + 8,$$

$$\text{где } s = \text{Grad}(x,y)/255, \quad g = G(x,y)/255, \quad z = \begin{cases} 0.4, & s > 0.5; \\ 0.8, & s \leq 0.5. \end{cases} \quad (2.16)$$

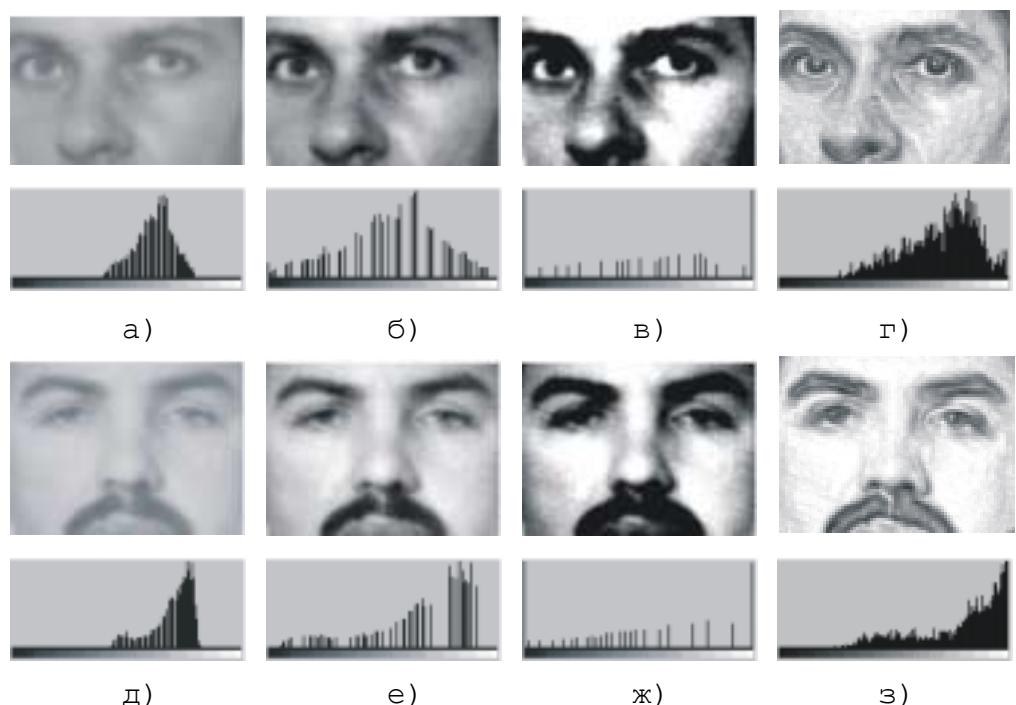


Рис. 2.18. Изменение яркостных характеристик исходных изображений а), д) с помощью «растяжения гистограммы» (2.10): б), е) $G_m = \min\{G_{ER}, G_{EL}\}$, $G_M = \max\{G_i\}$; в), ж) $G_m = M - \sigma$, $G_M = M + \sigma$. Изображения г), з) – результат применения подхода “High Boosting” (2.11) при $f=9,6$

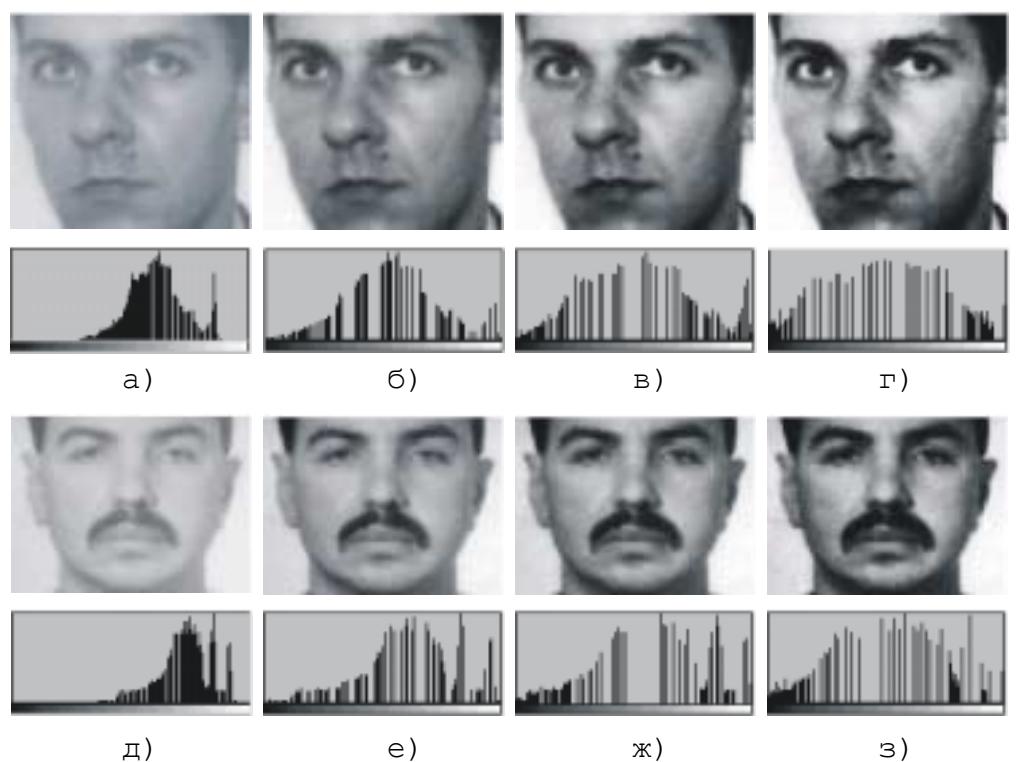


Рис. 2.19. Изменение яркостных характеристик исходных изображений а), д) с помощью метода «улучшения гистограммы» (2.12). Для изображений б), е) параметр $r = 0,1$; в), г) $r = 0,5$; г), з) $r = 1,0$

Здесь $Grad(x,y)$ – значение пикселя с координатами (x,y) карты краёв, полученной с помощью оператора Собеля; $G(x,y)$ – значение пикселя с координатами (x,y) исходного изображения; f – выбираемый параметр.

Результаты экспериментальных исследований показали, что данный подход целесообразно использовать в комбинации с фильтрацией и иными методами улучшения изображений на этапах последующего анализа изображений. Одиночное же его применение не даёт ожидаемых положительных результатов (рис. 2.18, g,z , $f = 9.6$).

Стабильное улучшение контраста и яркостных характеристик изображений было получено с помощью алгоритма нелинейного растяжения гистограммы, его формальная запись имеет следующей вид:

Если H_i – количество пикселей изображения со значением яркости i ($0 \leq i \leq 255$) и $G(x,y)$ – значение яркости пикселя (x,y) , то

$$Q = \sum_{i=0}^{255} (H_i)^r, \quad S_{G(x,y)} = \frac{255}{Q} \sum_{j=0}^{G(x,y)} (H_j)^r,$$

$$G'(x,y) = \begin{cases} 0, & S_{G(x,y)} \leq 0; \\ S_{G(x,y)}, & S_{G(x,y)}; \\ 255, & S_{G(x,y)} \geq 255. \end{cases} \quad (2.17)$$

Данный метод хорошо зарекомендовал себя на большом количестве протестированных изображений (более 300). Его плюсом является прямая зависимость стандартного отклонения значений пикселей итогового изображения от параметра r . На основании указанных положительных моментов он был выбран для использования на этапе предварительной обработки в экспериментальной системе идентификации человека по фотопортрету.

Примеры изображений приведены на рис. 2.19. Кроме того, рис. 2.17 v,g демонстрирует результат применения данного преобразования к рис. 2.17 a,b соответственно.

2.3 Анализ антропометрических точек лица

2.3.1 Антропометрические признаки в криминалистике

На основе 70-летнего опыта криминалистами была разработана процедура проведения фотопортретной экспертизы [11, 19, 37], а также методики криминалистического описания внешности человека [7, 32]. В данных работах приводится описание антропометрических точек, которые используются в процессе идентификации человека по фотопортрету (рис. 2.20, а). Однако ошибочно было бы считать, что криминалистическая фотопортретная экспертиза проводится только по некоторым

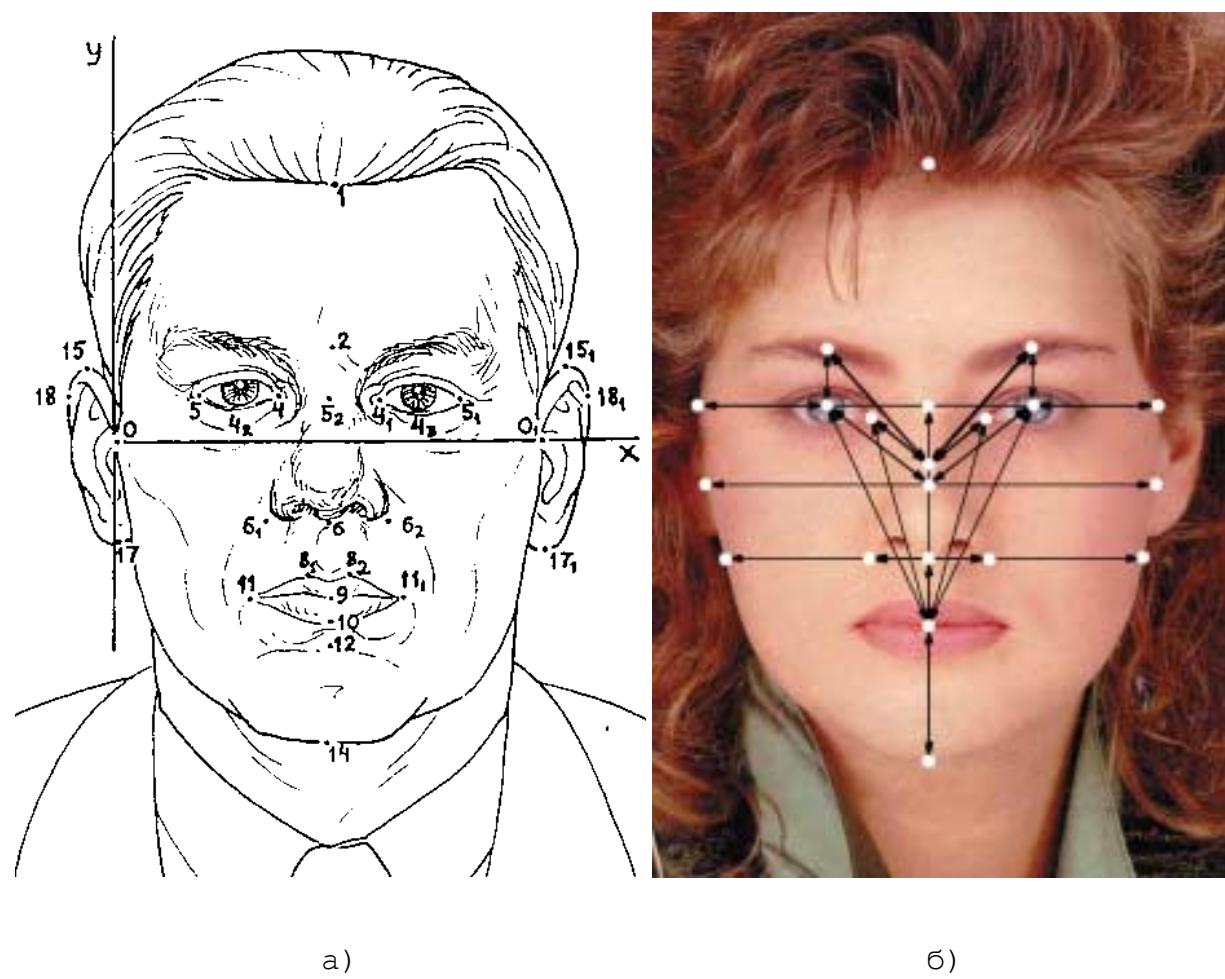


Рис. 2.20. Идентификационные точки: а) точки, используемые при криминалистической фотопортретной экспертизе; б) точки и расстояния между ними, наиболее часто применявшиеся при построении автоматизированных систем идентификации

количественным характеристикам, описывающим расстояния между данными точками. В процессе криминалистической идентификации по портрету изучаются такие характеристики элементов внешности, как форма, величина (относительная), положение, цвет (если сравниваемые снимки цветные), наличие (отсутствие) индивидуальных особенностей лица (морщины, родинки, шрамы и др.) и их количество, степень симметрии парных элементов, степень выраженности тех либо иных черт лица. Следует отметить, что измерение расстояний между антропометрическими точками проводится экспертами только по горизонтали или по вертикали. В методике [11] упоминаются лишь 13 основных расстояний лица, изображенного в фас, но при необходимости количество анализируемых расстояний может быть увеличено экспертом. Несмотря на то, что в некоторые системы автоматической идентификации базируются на аналогичных принципах, не все точки, выделяемые криминалистами, могут в них использоваться. При интерактивном режиме работы эксперт может выделять идентификационные точки, даже если они не видны на снимке (закрыты, например, волосами) или же видны плохо (верхнелобовая, надкозелковые, носовые - 1, 0, 0₁ и 6_{1,6₂} на рис. 2.20,а). В то же время автоматическое определение координат подобных точек крайне затруднено, а порой и вовсе невозможно.

Хотя указанный подход имеет значительные преимущества перед остальными, для его реализации требуется точное автоматическое выделение точек лица на цифровом изображении. Первые работы по распознаванию человека с помощью ЭВМ также базировались на данном методе. Однако полностью его формализовать не удалось из-за сложности автоматического определения местоположения большинства антропометрических точек с требуемой точностью.

На основе вышеизложенного имеет смысл произвести анализ антропометрических точек лица с точки зрения их автоматического выделения, используя криминалистическую методику идентификации как исходную.

2.3.2 Отбор антропометрических точек лица для автоматического распознавания

На основе анализа работ по построению систем идентификации человека по фотопортретам [35, 57, 59, 49, 65, 164] с использованием антропометрических признаков были выявлены наиболее часто использовавшиеся точки. Ниже приведены признаки - расстояния между характерными точками лица, которые чаще всего использовались при построении систем идентификации личности по портрету (рис. 2.20,б):

- между подбородком и линией глаз;
- между серединой линии смыкания губ и подбородком;
- между центром сетчатки глаза и линией волосяного покрова (в разных местах);
- между центром сетчатки глаза и центром носа;
- между кончиком подбородка и серединой лица;
- между внутренним уголком глаза и центром лица;
- между центром сетчатки глаза и центром брови;
- от центра брови до центра лица;
- между центром сетчатки глаза и серединой линии смыкания губ;
- между внутренним уголком глаза и уголком рта;
- между серединой линии смыкания губ и нижней точкой носа.
- ширина лица на уровне линии глаз.

Кроме того, часто при идентификации учитывались следующие области лица, вернее, их площади:

- треугольник с вершинами во внутренних уголках глаз и середине линии смыкания губ;
- треугольник с вершинами в центрах бровей и середине линии смыкания губ;
- пятиугольник с вершинами в следующих точках: складки верхних век, находящихся вертикально над центрами зрачков, центры зрачков и центр лица.

Рассмотрим упомянутые точки подробнее с точки зрения их инвариантности при изменении условий съёмки. Верхняя часть лица более статична и меньше подвергается изменениям (возрастным, косметическим и пр.), чем нижняя.

Группа 1. Точки, описывающие глаза:

- центры зрачков;
- внутренние уголки глаз;
- точка складки верхнего века, расположенная вертикально над центрами зрачков;
- точка края верхнего века, расположенная вертикально над центрами зрачков;
- наружные уголки глаз.

Расстояние между зрачками является самой важной измеряемой величиной: оно является базисом, по отношению к которому определяются все основные пропорции лица. Вторыми по степени инвариантности можно назвать внутренние уголки глаз. Их положение и видимость меньше меняются

со временем. Они меньше маскируются косметикой, чем наружные. Однако форма глаза (как и любая кривая) не может быть описана лишь двумя точками. Для полноты представления необходимо добавить не менее четырёх точек, что позволяет минимально описывать вертикальные размеры глаза. Пиксели, расположенные по вертикали над центрами зрачков и принадлежащие нижнему краю верхнего века и складке верхнего века, являются приемлемыми как по легкости выделения, так и по инвариантности. Веки могут быть опущены в момент съёмки, но изменить форму складки верхнего века без пластической операции практически невозможно. Изображение нижнего века менее инвариантно, так как его визуальная форма и местоположение изменяются относительно легко как умышленно («прищур», косметика), так и в силу независимых факторов (морщины, «мешки» под глазами). Визуальное изменение протяженности глаза по горизонтали является часто применяемым приемом в современном макияже. Поэтому целесообразно координатами наружных уголков глаз считать абсциссу и ординату точки глазного яблока, примыкающей к границе верхнего и нижнего век, максимально удалённой от вертикальной оси симметрии лица.

Группа 2. Точки, описывающие брови:

- крайние внутренние точки бровей;
- ординаты самой нижней и самой верхней точек бровей;
- верхняя, серединная и нижняя точки брови, располагающиеся на вертикали, проходящей через центр сетчатки глаза;
- крайние наружные точки бровей.

С точки зрения инвариантности длина бровей может быть отнесена на первое место лишь при распознавании мужских портретов, однако и женщины не так часто изменяют координаты внутренних крайних точек бровей, поэтому данные антропометрические точки упомянуты первыми как наиболее инвариантные. Полное описание формы бровей, а также последующее их сравнение хоть и является важным при идентификации, но дорого с точки зрения вычислительной стоимости алгоритма. Компромиссом может быть учёт лишь вертикальной протяженности брови, т.е. разности ординат крайней верхней и крайней нижней точек брови. Крайние наружные точки брови, ширина и середина брови в месте её пересечения с вертикальной осью, проходящей через центр сетчатки глаза, могут служить надёжными признаками при описании внешности мужчин и менее надёжными при описании женщин. Хотя женщины часто изменяют толщину и длину брови, полное видоизменение её формы, а также существенное изменение координат середины и крайних наружных точек брови происходит очень редко и встречается преимущественно лишь у девушек.

Группа 3. Точки, описывающие область носа:

- нижняя точка носа;
- носовые – точки пересечения вертикальных касательных к крыльям носа с линией, соединяющей нижние точки слияния крыльев носа с верхней губой;
- крайние точки лица, лежащие на горизонтали, проходящей через нижнюю точку носа.

Нижняя точка носа на большинстве фотографий может быть выделена достаточно точно. Существенно изменить её местоположение может только пластическая операция. Носовые или крайние боковые точки крыльев носа могут быть плохо заметны, особенно при бесстеневом освещении в момент съёмки. Однако ширина носа является важным признаком при идентификации человека человеком и игнорировать его нельзя. Ширина лица по горизонтали, проходящей через нижнюю точку носа, относительно надёжна при её выделении для мужчин, так как может иногда закрываться бородой или бакенбардами, и абсолютно инвариантна для женщин. Так называемая серединная точка носа, как, впрочем, и точка центра лица, используемая многими авторами в работах по автоматическому распознаванию, не приведены в данном списке по той простой причине, что методика определения их координат не является достаточно чёткой. Также следует отметить, что в большинстве случаев кончик носа не может быть достаточно точно автоматически выделен по фотографиям человека в фас (или же его выделение требует больших вычислительных затрат).

Группа 4. Точки, относящиеся к области рта:

- верхняя точка губ;
- центральная точка рта - точка пересечения серединной вертикали лица и линии смыкания верхней и нижней губ;
- крайние точки лица, лежащие на горизонтали, проходящей через центральную точку рта;
- нижняя точка лица;

Середина края верхней губы является самой инвариантной точкой рта. Расстояние между ней и нижней точкой носа является относительно устойчивым параметром, так как не изменяется при напряжении мимических мышц. Центральная точка рта тоже полезна при распознавании. Однако изменить её местоположение можно достаточно легко – плотно сжав губы, слегка приоткрыв рот или слегка улыбнувшись. Ширина лица, измеряемая по горизонтали, проходящей через центральную точку рта, является точно определяемой для большинства изображений женщин, но недостаточно

точно у мужчин в случае наличия бороды, усов. То же относится и к нижней точке лица.

Указанные выше точки расположены в порядке убывания инвариантности и надёжности точного их выделения. Группы точек имеют следующий приоритет: глаза, брови, нос, рот. На основе указанных точек можно выделить множество разных параметров для идентификации. Основные были уже названы, другие требуют дополнительной экспериментальной проверки эффективности их применения. Следует отметить, что расстояния, измеряемые вдоль горизонталей и вертикалей, требуют меньше вычислительных затрат на их получение и обработку, а точность измерения подобных параметров будет выше (криминалисты используют именно горизонтально-вертикальные измерения).

2.4 Применение фильтров Габора для обнаружения антропометрических точек лица

В рамках настоящей работы была исследована возможность автоматического поиска на изображении антропометрических точек с помощью фильтров Габора (см. раздел 1.2.8) - уголков глаз (внутренних - точки 2,3; наружных - точки 0,1), крыльев носа (точки 4,5) и уголков рта (точки 6,7). Эти точки показаны на рис.2.22.

Для поиска координат таких точек используют различные алгоритмы, чаще всего основанные на методе сопоставления изображения с эталоном [41, 138]. Основным недостатком таких алгоритмов является большое время работы при малой точности и высокая зависимость от условий получения фотопортрета. Вариант такого алгоритма, разработанный совместно с греческими коллегами, описан в работе [150].

В данном разделе описаны исследования возможности нахождения антропометрических точек лица с помощью фильтров Габора.

2.4.1 Исходная гипотеза

Предположим, что существуют фильтры, которые имеют экстремум в искомой точке. Требуется определить, имеются ли такие фильтры, и если да, то найти их параметры для каждой из антропометрических точек в зависимости от вида предобработки и нормализации исходного изображения.

Основанием применения преобразования Габора для распознавания объектов, текстур на изображениях, является тот факт, что оно очень похоже на психофизическую модель восприятия визуальной информации человеком. Преобразование Габора, используемое в виде фильтров, можно

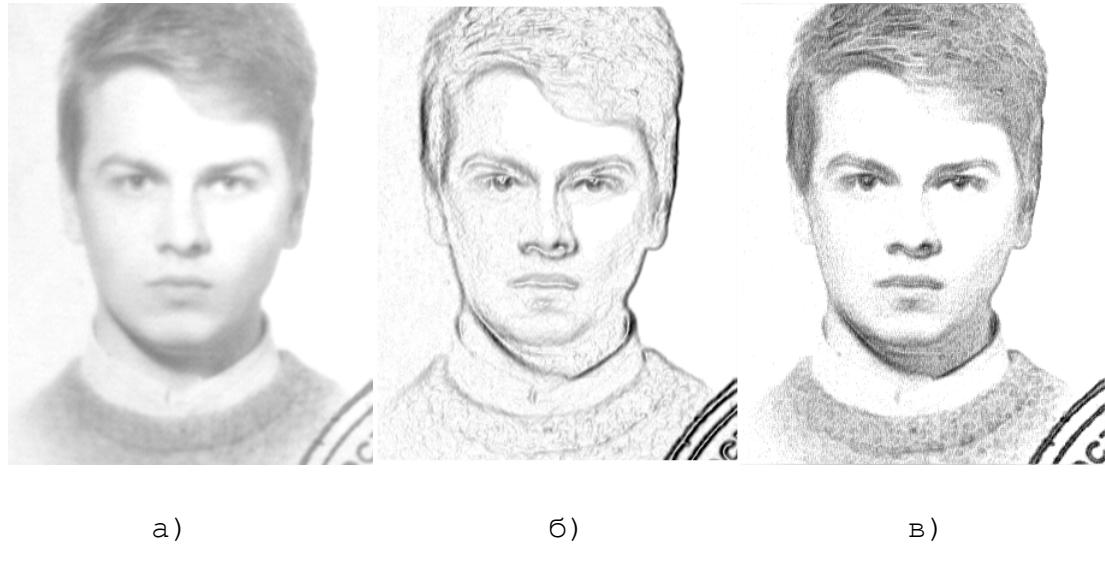


Рис.2.21. Исходное паспортное изображение (а), оно же после применения оператора Дериша (б) и после комбинированного подхода "High Boosting" (в)

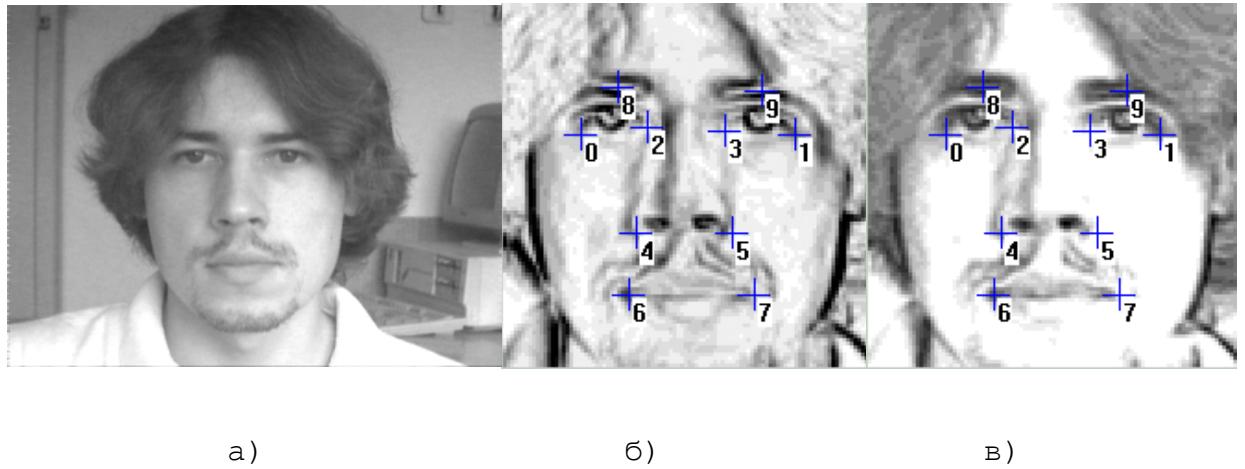


Рис.2.22. Исходное изображение, полученное с видеовхода (а), и нормализованные, с указанными антропометрическими точками, при помощи оператора Дериша (б) и оператора "High Boosting" (в)

рассматривать как «направленный микроскоп» с чувствительностью как к ориентации, так и к масштабу изображения [72].

Как уже упоминалось в разделе 1.2.8, фильтр Габора характеризуется следующими параметрами – радиусами эллипса σ_x и σ_y , углом ориентации θ . В зависимости от периода \cos и \sin в формуле Эйлера у фильтров может быть различное количество пиков (локальных максимумов) t (рис.1.10). Результатом свёртки фильтра с изображением в некоторой точке (x,y) является комплексное число.

Таблица 2.1

Параметры фильтров

Наименование параметров	Значения параметров
Радиус, σ	5, 9, 13, 17
Количество пиков, t	1, 3, 5
Угол ориентации фильтра, θ	0°, 30°, 45°, 60°, 90°
Начальный угол сектора, β	0°, 45°, 90°, 180°, 270°, 315°
Угол сектора, α	0°, 30°, 45°, 60°, 90°, 180°, 360°
Радиус отверстия кольца, r	0, $\sigma/2$

В настоящей работе исследовались различные варианты фильтров (рис. 1.10 и 2.23). Чтобы несколько уменьшить множество вариантов, вместо эллипса использовался круг, поэтому все фильтры имели $\sigma_x=\sigma_y$. Но при этом область определения фильтров могла иметь не только форму круга, но и сектора, кольца и их пересечений. Таким образом, в дополнение к (1.3) и (1.4) были введены новые параметры: угол начала сектора, угол сектора и радиус, определяющий ширину кольца. Каждый вариант фильтра (рис. 2.24) описывается шестью параметрами, указанными в табл.2.1.

2.4.2 Методика выполнения экспериментов и условные обозначения

Обозначим множество тестовых фотопортретов через $\mathbf{A}=\{A_k\}$, где $k \in [1, n]$. Эксперименты проводились на базе данных, содержащей 42 портрета ($n = 42$) двадцати человек. Все фотопортреты, использовавшиеся в ходе экспериментов, были предварительно нормализованы по масштабу и расположению зрачков, а центры зрачков располагались на горизонтали с расстоянием между ними 40 пикселей. Координаты зрачков определялись автоматически с помощью алгоритма, описанного в разделе 2.2.2. Исследовались два варианта предварительной обработки изображений: оператор

Дериша [98] и оператор “High boosting” [92]. Примеры исходного и обработанных портретов показаны на рис.2.21.

Через $A_k = \{a_{ij}^k\}$ обозначим отдельное изображение из множества \mathbf{A} , где a_{ij}^k – пиксель k -го изображения с координатами (i,j) , определёнными на плоскости изображения. После нормализации системы координат определена таким образом, что центр правого глаза у всех изображений A_k находится в фиксированной точке (i_0, j_0) .

Пусть $\mathbf{F} = \{F^p\}$, где $p \in [0, s-1]$ означает множество антропометрических точек лица человека. Были исследованы восемь точек (уголки глаз и рта, края носа), т.е. $s=8$. Через $F^p(A_k) = (i_j)^{p_k}$ обозначим координаты p -й антропометрической точки, найденной в плоскости k -го изображения. Следует отметить, что даже на фотопортретах одного человека координаты одной и той же антропометрической точки (например, уголка рта) могут отличаться при изменении выражения лица и положения головы относительно камеры. Для чистоты эксперимента координаты искомых точек для каждого портрета задавались вручную.

Пусть $\mathbf{G} = \{G^l\}$, где $l \in [1, m]$ – множество вариантов фильтра Габора. Проверка всех возможных значений параметров фильтров может быть значительно сокращена в силу плавного изменения функций Габора. Было выбрано несколько дискретных значений каждого из параметров (табл. 2.1). В ходе эксперимента поочерёдно генерировались все возможные комбинации указанных параметров. Общее их количество составляет $4 \times 3 \times 5 \times 6 \times 7 \times 2 = 5040$, однако было проверено $m=2160$ вариантов фильтров. Это объясняется тем, что при малых радиусах ($\sigma = 5, \sigma = 9$) фильтры с большим количеством пиков (т.е. малым периодом косинуса) не могут быть корректно реализованы из-за малого количества точек дискретизации.

Через $G^l(F^p(A_k))$ обозначим значение l -го варианта фильтра Габора, вычисленное в антропометрической точке F^p , имеющей координаты (i,j) на изображении A_k . Поскольку G^l – комплексное число, обозначим через Re^l и Im^l значения его действительной и мнимой частей, т.е. $G^l = Re^l + Im^l$ или

$$G^l(F^p(A_k)) = Re^l(F^p(A_k)) + Im^l(F^p(A_k)). \quad (2.18)$$

Окрестность точки $E = (i,j)$ радиуса h обозначим через $Q_h(E)$. В наших исследованиях использовалось квадратное множество точек размера $(2h+1) \times (2h+1)$ на плоскости изображения, удовлетворяющее условию

$$Q_h(E) = \{ E^* = (i^*, j^*) : (|i^* - i| \leq h, |j^* - j| \leq h) \}. \quad (2.19)$$

Такая окрестность соответствует кругу в шахматной метрике. Если используется другая метрика, круг будет иметь иную форму, что не принципиально для выполняемого исследования.

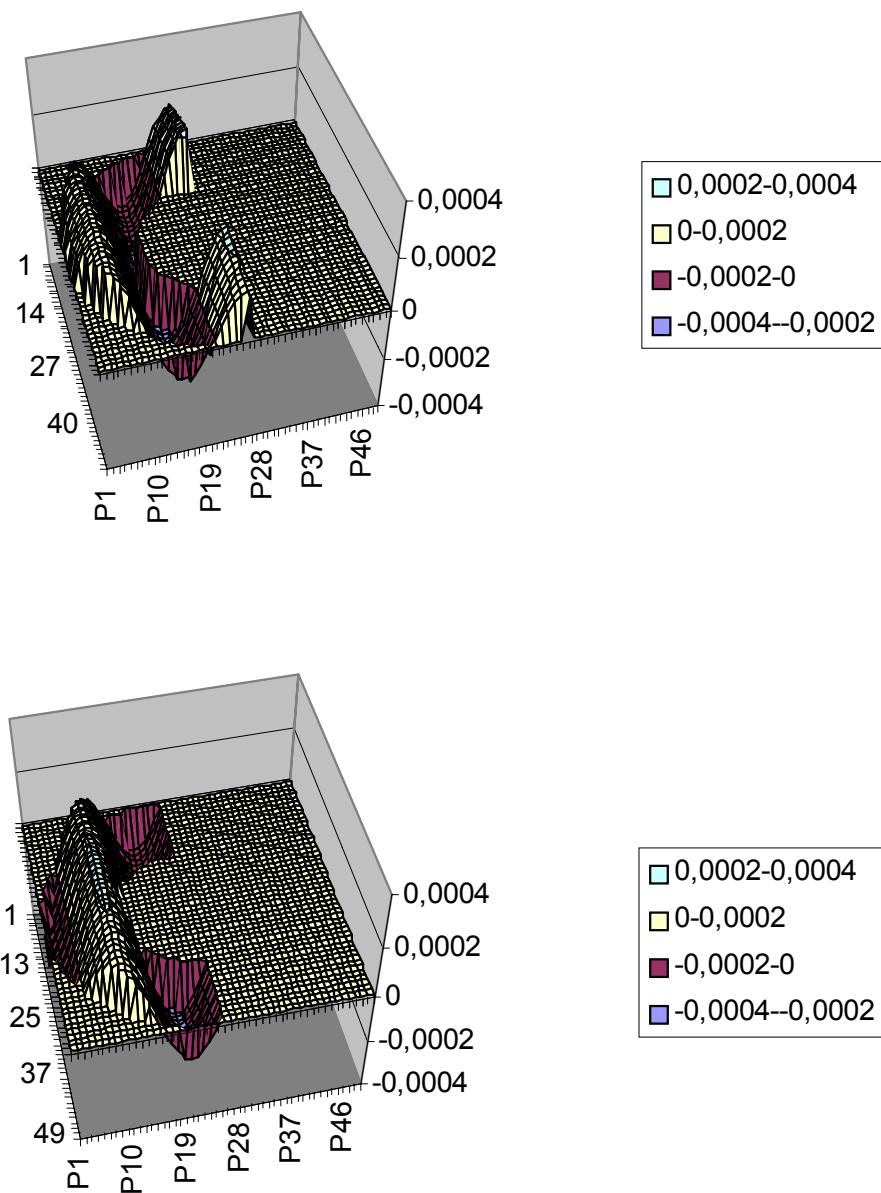


Рис. 2.23. Пример одного из вариантов фильтра Габора: функции, соответствующие действительной части фильтра (вверху) и мнимой (внизу). Параметры фильтра: $\sigma_x=\sigma_y=25$, $t=3$, угол ориентации $\theta = 180^\circ$, нач. угол сектора $\beta = 0^\circ$, угол сектора $\alpha = 180^\circ$, внутренний радиус кольца $r = 13$

Множество комплексных значений l -го варианта фильтра Габора, вычисленных в окрестности антропометрической точки F^p на изображении A_k , обозначим через $\mathbf{M}^l(Q_h(F^p(A_k))) = \{G^l(Q_h(F^p(A_k)))\}$. Это множество имеет максимальные и минимальные значения для действительной и мнимой частей фильтра l -го типа

$$\begin{cases} Re_{max}_{lhpk} = \max \{ Re^l(Q_h(F^p(A_k))) \}, \\ Re_{min}_{lhpk} = \min \{ Re^l(Q_h(F^p(A_k))) \}, \\ Im_{max}_{lhpk} = \max \{ Im^l(Q_h(F^p(A_k))) \}, \\ Im_{min}_{lhpk} = \min \{ Im^l(Q_h(F^p(A_k))) \}. \end{cases} \quad (2.20)$$

Точку E , в которой фильтр l -го типа даёт экстремальное значение, будем обозначать с учётом левой части уравнений (2.20), например $E^{Re_{max}lhpk}$. Уравнения (2.20) определяют четыре экстремальных точки, обозначим их в соответствии с порядком уравнений (2.20) E^1, E^2, E^3, E^4 . Таким образом, $E^1 = E^{Re_{max}lhpk} = (i_1, j_1)$, где (i_1, j_1) – координаты точки E^1 в плоскости изображения A_k , в которой фильтр l даёт максимальный отклик в окрестности Q_h антропометрической точки F^p .

Пусть $d(E, E^*)$ – расстояние между двумя точками $E = (i, j)$ и $E^* = (i^*, j^*)$, вычисленное в городской метрике. Тогда расстояние между антропометрической точкой F^p и точками E^1, E^2, E^3, E^4 из её окрестности на изображении A_k обозначим через d_1, d_2, d_3, d_4 . Такие расстояния были найдены для всех вариантов фильтра, для всех антропометрических точек на всех тестовых изображениях. С учетом наших обозначений получено четыре множества расстояний $\{d_1^{lpk}\}, \{d_2^{lpk}\}, \{d_3^{lpk}\}, \{d_4^{lpk}\}$, где верхние индексы означают вариант фильтра (l), номер антропометрической точки (p) и номер изображения (k). Для каждого из этих множеств вычислялось среднее отклонение точки, дающей экстремальное значение фильтра Габора по всем изображениям при фиксированных параметрах l и p :

$$d_z^{lp} = \frac{l}{n} \sum_{k=1}^n d_z^{lpk}, \quad (2.21)$$

где z меняется от 1 до 4.

Иными словами, в ходе экспериментов каждый из 2160 вариантов фильтра, сгенерированных в соответствии с возможными значениями параметров, вычислялся для всех пикселей окрестности каждой из восьми исходных антропометрических точек. Местоположение каждого из экстремумов $Re_{max}, Re_{min}, Im_{max}, Im_{min}$ отдельного фильтра относительно исходной точки оценивалось по городской метрике в пикселях и усреднялось по всем обработанным изображениям.

Было проведено три различных эксперимента. В каждом из них были различными либо варианты предобработки исходных изображений либо размеры анализируемых окрестностей искомых точек. В первом случае был использован оператор “High boosting” и окрестность 7×7 ($h=3$), во втором – оператор Дериша при той же окрестности. При проведении третьего эксперимента область поиска экстремума для каждой из точек была расширена до 15×15 ($h=7$) пикселей, а в качестве варианта предобработки был выбран оператор Дериша.

2.4.3 Результаты исследований

Результаты экспериментов анализировались отдельно по вышеуказанным антропометрическим точкам лица. Ввиду того, что они симметричны относительно вертикальной оси лица, то данные приведены ниже лишь для половины из указанных выше точек – внешнего уголка глаза (точка 0), внутреннего уголка глаза (точка 2), точки на крыле носа (точка 5), угла рта (точка 6). Для симметричных им точек результаты практически идентичные с учётом зеркального отображения фильтров.

Внешний уголок глаза (точка 0). Для данной точки вычислялись значения функций $d_1^{10} - d_4^{10}$ для $l \in [1,2160]$. На рис. 2.25 приведён график значений функции d_1^{10} для $l \in [1441,2160]$, из которого видно, что некоторые фильтры по сравнению с остальными имеют ярко выраженные локальные минимумы. В некоторых случаях (это зависит от условий эксперимента и вида функции) наблюдается чёткая периодичность упомянутых экстремумов, которая говорит о том, что определённый параметр доминирует при формировании экстремальных значений фильтров в окрестности искомой точки. Номера фильтров с 25 лучшими значениями указанных функций на протяжении трёх опытов приведены в табл. 2.2. Параметры для некоторых из них указаны в табл. 2.3. Для анализа результатов такие таблицы, а также графики по каждой из функций строились для всех антропометрических точек. За недостатком места приводятся лишь окончательные результаты и выводы в сокращённом виде.

Из анализа табл. 2.2 следует, что несмотря на хорошие результаты в первом и во втором опытах (значения функций находятся в пределах 2.1–2.5, что означает возможность локализации искомой точки с точностью до 2 пикселей), при увеличении окрестности, по которой производился поиск, до 15×15 ($h=7$ в третьем эксперименте) у большинства из приведённых в таблице фильтров расположение экстремумов стало реже (например, в 4 раза при $l = 2099$). Следует также отметить, что хотя в с точки зрения упрощения программной реализации алгоритма предпочтительно отбирать

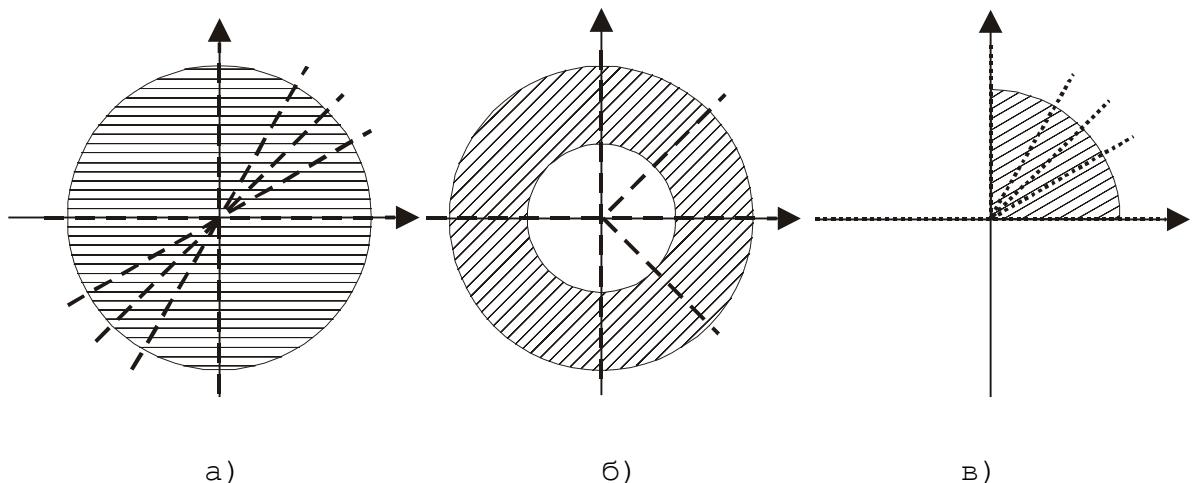


Рис.2.24. Примеры возможных параметров и форм фильтров:
 а) круг, б) кольцо, в) сектор. На (а) пунктирной линией показаны исследовавшиеся углы ориентации фильтра θ , на (б) - начальные углы секторов β , на (в) - возможные углы секторов α

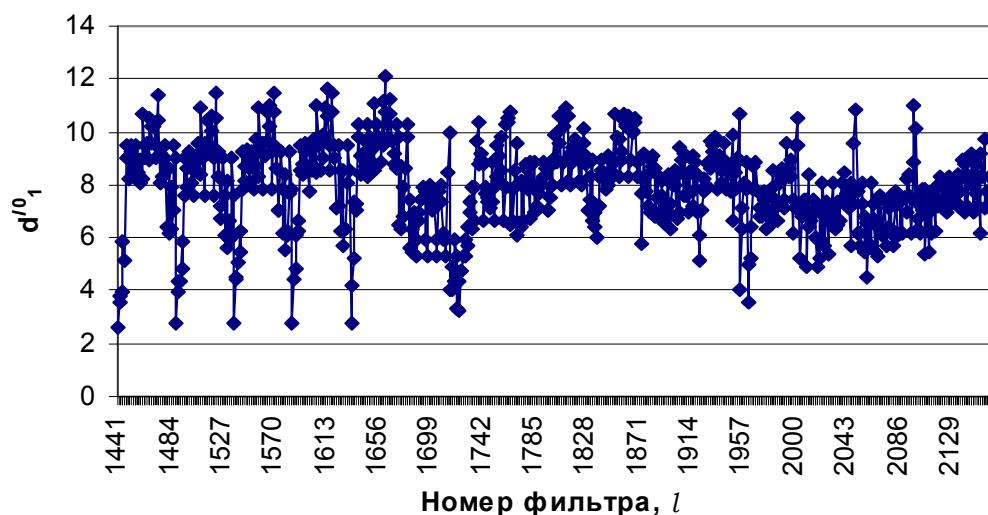


Рис.2.25. Пример значений функции d_1^{10} для множества фильтров $l \in [1441, 2160]$ в эксперименте №3

Таблица 2.2

25 лучших значений функций d_z^{10} в трёх опытах

Номер фильтра, l	Эксперимент №1				Эксперимент №2				Эксперимент №3			
	d_1^{10}	d_2^{10}	d_3^{10}	d_4^{10}	d_1^{10}	d_2^{10}	d_3^{10}	d_4^{10}	d_1^{10}	d_2^{10}	d_3^{10}	d_4^{10}
242	2.57	4.5	2.64	4.69	2.57	4.79	2.69	4.88	3.78	10.5	4.85	10.2
338	2.48	4.93	4.17	2.62	2.48	5.07	4.6	2.69	4.17	10.4	10.4	3.8
386	2.52	4.98	4.55	2.55	2.33	5.05	4.81	2.57	4	10.5	10.4	3.88
434	2.57	4.86	4.79	2.45	2.36	5	4.9	2.52	4	10.5	10.5	4
437	3.36	4.9	4.79	2.43	3.33	4.83	5.02	2.29	7.85	9.32	11.7	2.98
713	2.36	4	4.79	2.19	2.86	3.9	4.6	2.26	6.56	7.54	6.9	6.61
722	2.02	5.1	3.02	4.45	2.33	5.26	3.39	5.17	3.07	11.0	5.1	10.4
770	2.21	5.14	4.69	2.9	2.48	5.33	4.52	3.4	3.01	11.4	9.7	5.02
817	2.1	5.29	4.83	2.26	2.52	5.38	5.02	2.62	3.61	10.8	10.9	2.88
818	2.38	5.21	4.69	2.24	2.52	5.38	5.05	2.74	3.66	11.4	11.1	3.2
865	2.07	5.24	5	2.17	2.64	5.36	5.17	2.55	3.8	10.9	11.3	3
866	2.43	5.19	4.88	2.19	2.55	5.38	5.21	2.43	3.73	11.3	11.2	3.07
913	2.21	5.14	5.24	2.12	2.64	5.29	5.45	2.43	4.1	10.8	11.3	2.95
914	2.4	5.05	5.1	2.29	2.5	5.36	5.33	2.52	3.59	11.0	11.5	3.15
1242	2.36	3.5	3.36	2.5	2.21	3.5	3.62	2.55	4.93	7.9	9.22	3.66
1396	2.38	3.52	3.81	3.02	2.29	3.4	3.36	2.69	7.27	7.05	7.3	8.2
1441	2.17	4.88	3.79	4.62	2.33	4.95	3.74	4.64	2.63	11.1	5.73	10.6
1489	2.31	4.86	4.07	3.57	2.4	5.12	4.33	3.24	2.78	11.0	8.98	4.83
1537	2.12	4.95	4.26	3.02	2.29	5.19	4.69	3	2.73	10.9	10.9	3.41
1585	2.1	5.12	4.45	2.93	2.36	5.14	4.79	2.79	2.78	10.8	11.2	3.27
1633	2.02	5.14	4.6	2.81	2.4	5.19	4.98	2.81	2.8	10.5	11.1	3.39
2004	4.07	2.55	2.19	4.48	3.83	2.67	2.71	4.29	5.24	7.02	4.1	10.1
2099	2.29	4.33	4.6	2.71	2.98	4.4	4.57	3.05	11	9.2	9.76	9.9
2105	2.4	3.83	4.64	2.17	3.1	3.79	4.43	2.69	7.8	5.78	7.93	5.78
2153	2.5	3.74	4.69	1.9	2.83	3.64	4.29	2.33	6.15	6.85	8.02	5.15

Таблица 2.3

Параметры восьми фильтров,
имевших лучшие значения функций d_1^{10} либо d_4^{10}

Номер фильтра, l	Параметры фильтров, $(\sigma, t, \theta, \beta, \alpha, \tau)$	Номер фильтра, l	Параметры фильтров, $(\sigma, t, \theta, \beta, \alpha, \tau)$
437	$(9, 1, 90^\circ, 0^\circ, 180^\circ, 0)$	1489	$(17, 1, 30^\circ, 0^\circ, 45^\circ, 0)$
817	$(13, 1, 45^\circ, 0^\circ, 45^\circ, 0)$	1537	$(17, 1, 45^\circ, 0^\circ, 45^\circ, 0)$
913	$(13, 1, 90^\circ, 0^\circ, 45^\circ, 0)$	1585	$(17, 1, 60^\circ, 0^\circ, 45^\circ, 0)$
1441	$(17, 1, 0^\circ, 0^\circ, 45^\circ, 0)$	1633	$(17, 1, 90^\circ, 0^\circ, 45^\circ, 0)$

Таблица 2.4

Рекомендуемые параметры фильтров и значения
соответствующих расстояний $d_1^{l^*} - d_4^{l^*}$ в третьем опыте

Номер точки	Номер фильтра, l	Параметры фильтров, $(\sigma, t, \theta, \beta, \alpha, \tau)$	Эксперимент №3 *			
			$d_1^{l^*}$	$d_2^{l^*}$	$d_3^{l^*}$	$d_4^{l^*}$
Точка 2	554	$(9, 2, 30^\circ, 180^\circ, 45^\circ, 4)$	9.49	2.46	3.34	8.54
	602	$(9, 2, 45^\circ, 180^\circ, 45^\circ, 4)$	9.83	2.66	7.56	3.1
	650	$(9, 2, 60^\circ, 180^\circ, 45^\circ, 4)$	4.59	6.34	9.51	2.07
	986	$(13, 2, 0^\circ, 180^\circ, 45^\circ, 6)$	2.63	9.63	8.07	6.1
	1226	$(13, 3, 0^\circ, 180^\circ, 45^\circ, 6)$	4.9	7.44	9	2.73
Точка 5	363	$(9, 1, 45^\circ, 180^\circ, 90^\circ, 0)$	5.41	10.4	3.98	9.51
	638	$(9, 2, 60^\circ, 45^\circ, 180^\circ, 4)$	9	3.78	7.15	6.1
	1076	$(13, 2, 45^\circ, 90^\circ, 90^\circ, 6)$	10.2	3.32	8.1	9.07
	1757	$(17, 2, 30^\circ, 180^\circ, 180^\circ, 0)$	3.9	10.6	6.56	7.02
Точка 6	366	$(9, 1, 45^\circ, 180^\circ, 180^\circ, 4)$	3.66	8.61	6.78	9.17
	515	$(9, 2, 0^\circ, 270^\circ, 90^\circ, 0)$	4.9	5.59	7.83	3.63
	523	$(9, 2, 0^\circ, 315^\circ, 90^\circ, 0)$	3.39	8.68	5.29	5.29
	525	$(9, 2, 0^\circ, 315^\circ, 180^\circ, 0)$	2.73	7.07	4	9.32
	563	$(9, 2, 30^\circ, 270^\circ, 90^\circ, 0)$	9.12	4.95	7.46	3.63

* Прим.: $d_1^{l^*}$ означает d_1^{l2} либо d_1^{l5} , либо d_1^{l6} в зависимости от точки, к которой относится соответствующая строка таблицы.

варианты фильтра, дающие значения близкие к экстремальным около искомой точки одновременно как для действительной, так и мнимой части фильтра, на практике можно использовать по отдельности действительную часть одного фильтра и мнимую другого.

Из табл. 2.2 и 2.3 видно, что наиболее стабильные результаты ($l= 1441\text{--}1633$) достигаются при $\sigma = 17$, $t = 1$, угол ориентации θ произвольный ($0\text{--}90^\circ$), начальный угол $\beta = 0^\circ$, угол сектора $\alpha = 45^\circ$ и $r = 0$, т.е. форма фильтра круг, а не кольцо. Отсутствие влияния параметра θ на значения фильтра объясняются симметричной формой фильтра относительно его центра (при $t = 1$), т.е. фильтр Габора вырождался в двумерный Гауссиан.

Остальные точки. К ним относятся: внутренний уголок глаза (точка 1), крылья носа (точка 5), уголок рта (точка 6). В табл. 2.4 приведены сводные данные по лучшим значениям расстояний d_1^{l*} - d_4^{l*} . Для всех указанных в табл. 2.4 фильтров соответствующие значения d_1^{l*} - d_4^{l*} в первом и втором опытах значительно лучше. Поскольку окрестность поиска точки в третьем эксперименте была увеличена более чем в четыре раза, то соответствующие значения лучше отражают возможность применения фильтров Габора для обнаружения антропометрических точек. Относительно плохие значения d_1^{l*} - d_4^{l*} для точек на крыльях носа можно объяснить их слабой определённостью на портретах и соответственно большей погрешностью их выделения вручную, которая и отражается на результатах опыта.

Для точек, соответствующих уголкам глаз, лучшими являются фильтры с большими радиусами, - для края носа и уголков губ предпочтительнее выглядят фильтры с $\sigma \in [5, 9]$. Сравнивая между собой значения фильтров в первом и втором опытах, можно заметить, что в некоторых случаях (особенно для точки 6) вариант нормализации “High Boosting” позволяет находить точки с большей точностью, чем оператор Дериша.

2.5 Выводы

Вторая глава посвящена вопросам нормализации, предварительной обработки исходных фотопортретов, а также выделению из них информации, необходимой для дальнейших этапов распознавания. Были определены требования к качеству и параметрам исходных фотопортретов, разработаны алгоритмы поиска области лица и центров зрачков на слабоконтрастных фотопортретах, проведены исследования по предварительной обработке и нормализации слабоконтрастных изображений для последующего автоматического распознавания, произведено исследование возможности нахождения антропометрических точек лица с помощью фильтров Габора.

Результаты выполненной работы можно сформулировать следующим образом:

1. Разработанные алгоритмы поиска области лица и центров зрачков глаз отличаются от существующих способностью находить лицо и центры зрачков на слабоконтрастных фотопортретах со сложным фоном в режиме реального времени без предварительного обучения и настройки (п.2.2.1, п.2.2.2).

2. При экспериментальном исследовании алгоритмов поворота цифровых изображений было определено, что лучшим с точки зрения минимизации вносимых искажений является алгоритм Оуэна и Македона. С точки зрения визуального восприятия человеком лучшим является алгоритм ADOBE Photoshop®. Обрабатывать изображения в интерактивном режиме для последующего автоматического анализа лучше с помощью пакета COREL Photo-Paint®, так как он меньше искажает исходное изображение, по сравнению с ADOBE Photoshop® (п.2.2.3).

3. Алгоритм нелинейного растяжения гистограммы продемонстрировал стабильное улучшение контраста для более 300 слабоконтрастных фотопортретов. Его положительной стороной является прямая зависимость стандартного отклонения значений пикселей итогового изображения от задаваемого параметра. Результаты экспериментальных исследований показали, что для последующего анализа изображений целесообразно использовать подход “High Boosting” в комбинации с фильтрацией и нелинейным растяжением гистограммы. Одиночное же его применение к слабоконтрастным изображениям не даёт положительных результатов (п.2.2.5).

4. На основе анализа 5040 возможных комбинаций параметров фильтра Габора установлено, что с помощью фильтра можно находить координаты антропометрических точек лица. Экспериментально отобранные лучшие варианты фильтра позволяют определять восемь основных антропометрических точек лица с точностью до 2 пикселей. Варианты предварительной обработки фотопортретов существенно влияют на точность поиска. Для нахождения большинства из указанных точек использовалась предварительная обработка изображения оператором Дериша, однако в некоторых случаях (особенно при поиске уголков рта) предпочтительнее применение подхода “High Boosting”. Такой вариант предобработки фотопортретов позволяет находить точки с большей точностью, чем после использования оператора Дериша (п.2.4).

Результаты исследований, описанных в данной главе, опубликованы в печатных работах [29, 139, 145, 149, 150] и отчётах о НИР [20, 25, 33].

ГЛАВА 3 АЛГОРИТМЫ ИДЕНТИФИКАЦИИ ЧЕЛОВЕКА И СРАВНЕНИЯ ФОТОПОРТРЕТОВ

В третьей главе рассматриваются алгоритмы решения задач поиска портрета в базе данных и паспортного контроля, разработанные на основе проведённых автором исследований.

3.1 Поиск фотопортретов в базе данных

Имеется база данных, содержащая N фотопортретов (от нескольких тысяч и больше), требуется найти в ней n изображений (от одного до десятка), наиболее похожих на заданное. Выбор окончательного решения предоставляется эксперту (оператору).

При решении данной задачи необходимо учитывать следующие факторы. Пусть имеется база данных N изображений (порядка $10^3 - 10^6$) и фотография некоего человека. Для большой базы данных задача идентификации человека по его фото не имеет однозначного решения, поскольку по статистике в реальной базе размером $N = 10^6$ может находиться до 417 пар двойников [68]. Поэтому окончательное решение принимает оператор, анализируя автоматически сформированное множество K . В процессе исследований было принято $K = (\ln N)^2$. Если использовать для поиска информацию о человеке (пол, раса, возраст, регион и т.п.), организованную в виде индексов в соответствии с теорией реляционных баз данных [39, 45], то можно значительно уменьшить число проверяемых портретов $M < N$. Для отбора ближайших K портретов в случае несортированной базы надо проверить все M записей. Если критерий отбора будет иным, например, «отобрать первые K портретов, у которых значение функции сходства с заданным фотопортретом выше некоторого порога R », то вероятность P – того, что хотя бы один портрет из K будет находиться в числе последних K проверенных записей, тоже достаточно велика. Так как необходимо отобрать K портретов, удовлетворяющих указанному выше условию, то вероятность нахождения портрета на конкретном (например, на последнем месте) будет равна $p_n = 1/N$; так как таких мест K и искомых портретов тоже K , то $P = K^2/N = (\ln N)^2/N$. Таким образом, при $N = 10^4$ вероятность того, что какой-либо из $(\ln(10^4))^2 = 84$ искомых портретов окажется на одном из последних мест, будет равна $P = 0,72$. Из-за разницы скорости роста функций $f(N) = N$ и $g(N) = (\ln N)^2$ при $N=10^5$, $K = 132$ вероятность такого события будет значительно ниже – $P = 0,172$, но всё равно достаточно велика – почти каждый пятый случай. Допустим, что, применяя методы решения задач верификации и контроля доступа, мы добились получения точной оценки сравнения

двух фотопортретов (не используя априорную информацию и не учитывая вычислительные затраты на предобработку изображений) за время $t = 1$ с. Следует отметить, что для задачи «паспортного контроля» это был бы идеальный вариант. Используя подобные методы при поиске в базе данных, время работы алгоритма в худшем случае составило бы $T = 2,(7)$ ч (при $N = 10^4$) и $T = 27,(7)$ ч (при $N = 10^5$), что неприемлемо.

В процессе анализа возможных вариантов решения данной задачи был выбран двухэтапный подход. Его суть заключается в постепенном сужении области поиска: на первом этапе отбираются портреты в количестве K_1 , наиболее близкие к искомому в признаковом пространстве. Использование признаков, описывающих и хранящихся вместе с изображениями, предполагает большую скорость отбора близких портретов. На втором этапе все K_1 отобранных портретов проходят более тщательное сравнение с изображением искомого человека.

Таким образом, поиск производится сперва по индексам некоторой группы людей, затем по геометрическим признакам лица и на финальном этапе путём детального сравнения нормализованных фотопортретов. Схема такой технологии поиска приведена на рис. 3.1. Опишем её подробнее.

3.1.1 Заполнение базы данных

На этапе формирования базы данных выполняется нормализация всех фотопортретов. Это позволит существенно сэкономить время при последующем поиске. Подробно этапы нормализации рассмотрены в главе 2.

Выше уже отмечалось, что основой геометрического подхода являются признаки, формируемые на антропометрических точках лица. В разделах 2.3 и 2.4 проведён анализ таких точек для автоматического распознавания. Экспериментальная проверка [26, 49, 65, 144] показала, что их совокупность наиболее устойчива по сравнению с другими системами признаков как к вариациям ракурса, так и к возрастным изменениям. Из-за возможной разницы в обработке изображений для локализации точек и сравнения на финальной стадии поиска автоматическое выделение антропометрических точек лица производится уже после обнаружения местоположения центров зрачков на исходном фотопортрете. Во избежание ошибок предусмотрена возможность ручной коррекции координат антропометрических точек лица. Вычисленные на их основе признаки представляют собой вектор, который хранится в базе.

Заполнение информации (записи) об объекте выполняется оператором. Определённые признаки человека, как-то пол, раса, возраст, особые приметы и т.п., формируют индексы поиска.

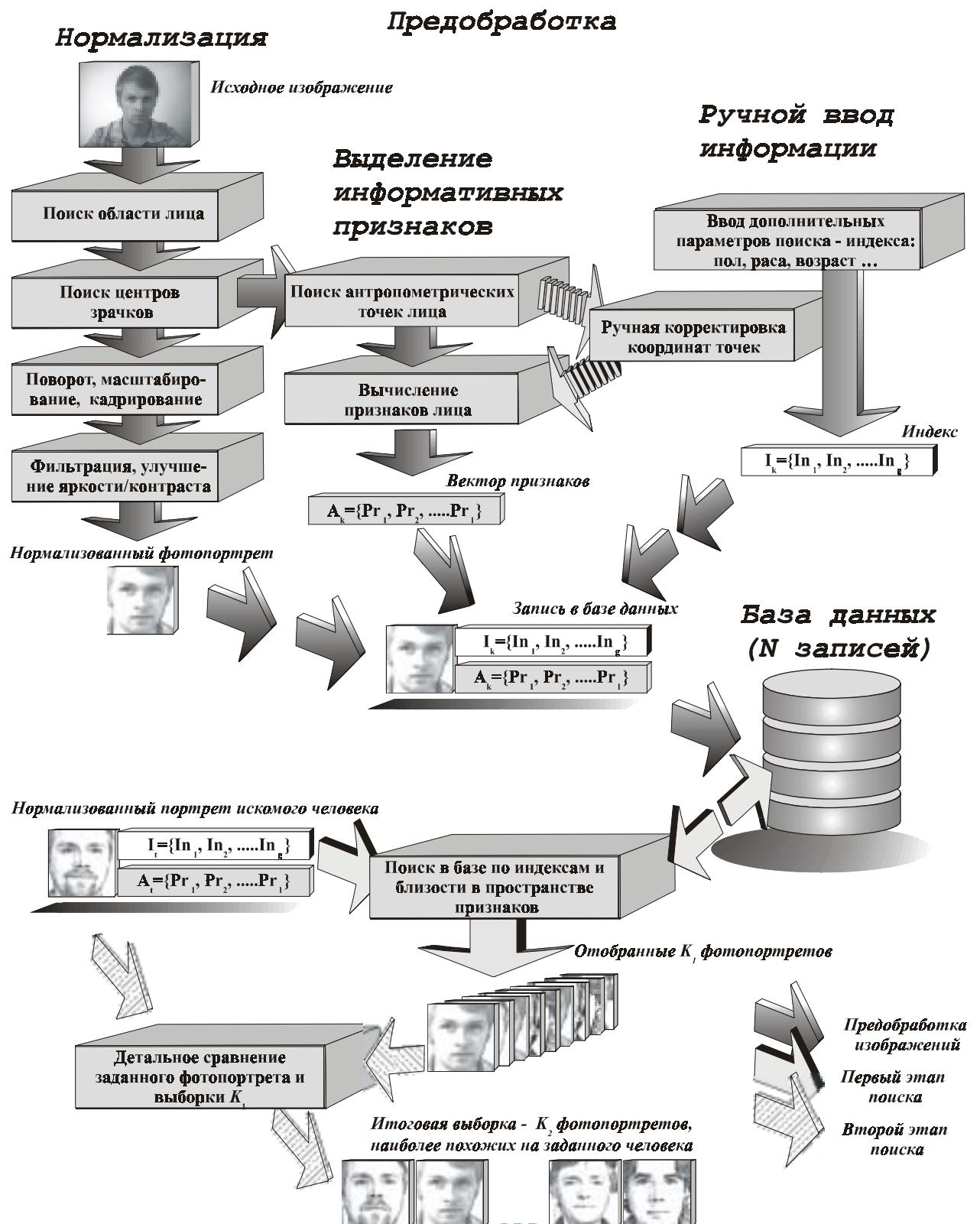


Рис. 3.1. Схема предварительной обработки и двухэтапного поиска фотопортретов, похожих на заданного человека, в большой базе данных

Все вышеназванные меры в таком же порядке применяются и к исходным фотопортретам, по которым производится поиск.

3.1.2 Первичный отбор портретов

Выбор признаков влияет на точность распознавания и является критическим этапом построения системы распознавания. Хорн в [46, стр. 334] отмечает: «... методы классификации образов не должны использоваться слепо, результаты хороши лишь настолько, насколько хороши признаки, выбранные для измерений. Никакие ухищрения в алгоритме решения не могут компенсировать неудачного выбора признаков». Так, например, на множестве из 40 точек только отрезков можно построить 780, не говоря уже про их комбинации, а также параметры всех возможных многоугольников (периметр, площадь), построенных на основе указанного множества точек. Поэтому в первоначальных экспериментах использовался набор признаков, апробированный в работах [49, 65]. Общее количество признаков – 30 (табл. 3.1). Половина из них – евклидовы расстояния между парами антропометрических точек лица. Остальные являются средними арифметическими пар расстояний между симметричными относительно вертикальной оси антропометрическими точками лица. Все признаки нормализованы делением на расстояние между зрачками. Общее количество точек лица, использовавшихся для получения признаков, - 33 (рис. 2.1).

Анализ этого набора признаков приведён в следующем разделе. Для подтверждения возможности эффективного поиска при помощи геометрических признаков были выполнены два эксперимента.

В обоих случаях использовались изображения из базы данных ORL фирмы Olivetti [107, 141]. Первый эксперимент проводился на изображениях пяти человек, которые рассматривались как пять классов по пять изображений в каждом. Общее количество снимков было 25. Целью эксперимента было найти пять портретов, ближайших к некоторому заданному в 30-мерном признаковом пространстве. В процессе проведения эксперимента каждый из портретов был искомым и поочерёдно сравнивался с каждым из оставшихся 24, по принципу «каждый со всеми». Результаты представлены в табл. 3.2. В именах фотопортретов $S_{i,j}$ закодированы номер класса i и номер снимка в классе j .

Из табл. 3.2 видно, что всего для двух изображений - $S5_10$ и $S10_10$ (строки выделены серым цветом) – пять ближайших портретов не содержат фотографий искомого человека. Помимо этого, только в одном случае ($S10_8$) ближайший портрет не принадлежал искомому человеку.

Таблица 3.1

Геометрические признаки лица человека

Номер признака	Расстояния между точками, используемыми для вычисления l -го признака	Номер признака	Расстояния между точками, используемыми для вычисления l -го признака
1	2	1	2
Pr ₁	(EL5,ER5)	Pr ₁₆	((ER1,ER2) + (EL1,EL2))/2
Pr ₂	(NL2,NR2)	Pr ₁₇	((ER6,ER4) + (EL6,EL4))/2
Pr ₃	(LL2,LR2)	Pr ₁₈	((EL5,NL2) + (ER5,NR2))/2
Pr ₄	(FL1,FR1)	Pr ₁₉	((NL2,LL2) + (NR2,LR2))/2
Pr ₅	(LL1,FL1)	Pr ₂₀	((LL2,FL1) + (LR2,FR1))/2
Pr ₆	(LR1,FR1)	Pr ₂₁	((ER2,NR1) + (EL2,NL1))/2
Pr ₇	(LL1,LR1)	Pr ₂₂	((ER1,BR1) + (EL1,BL1))/2
Pr ₈	(LM,LU)	Pr ₂₃	((BR1,BR3) + (BL1,BL3))/2
Pr ₉	(LM,LD)	Pr ₂₄	((EL2,BL3) + (ER2,BR3))/2
Pr ₁₀	(LD,TL)	Pr ₂₅	((EL6,BL4) + (ER6,BR4))/2
Pr ₁₁	(NN,LU)	Pr ₂₆	((BL4,BL5) + (BR4,BR5))/2
Pr ₁₂	(NL1,NR1)	Pr ₂₇	((BL5,BL3) + (BR5,BR3))/2
Pr ₁₃	(BL3,BR3)	Pr ₂₈	((BL5,BL1) + (BR5,BR1))/2
Pr ₁₄	(EL2,NN)	Pr ₂₉	((ER2,LR1) + (EL2,LL1))/2
Pr ₁₅	(ER2,NN)	Pr ₃₀	((EL1,EL5) + (ER1,ER5))/2

Таблица 3.2

Результаты эксперимента №1

Искомое изображение	Изображения, ближайшие к искомому				
	1	2	3	4	5
S5_6	S5_9	S5_2	S5_8	S9_4	S9_8
S5_10	S7_5	S7_4	S7_1	S6_5	S6_4
S5_9	S5_6	S5_8	S5_2	S10_5	S9_4
S5_2	S5_6	S5_9	S5_10	S7_5	S9_8
S5_8	S5_9	S5_6	S5_2	S10_5	S10_6
S6_5	S6_4	S6_7	S6_6	S10_10	S10_8
S6_3	S6_6	S6_5	S6_4	S6_7	S10_10
S6_4	S6_5	S6_6	S6_7	S9_8	S9_9
S6_6	S6_7	S6_3	S6_5	S6_4	S10_10
S6_7	S6_5	S6_6	S6_4	S6_3	S10_10
S7_5	S7_1	S7_4	S5_10	S7_8	S7_7
S7_7	S7_8	S7_1	S7_5	S7_4	S5_10
S7_8	S7_7	S7_1	S7_5	S7_4	S5_10
S7_1	S7_5	S7_8	S5_10	S7_4	S7_7
S7_4	S7_5	S5_10	S7_1	S7_8	S7_7
S9_9	S9_8	S9_4	S9_6	S9_1	S10_5
S9_8	S9_9	S9_4	S9_6	S9_1	S6_4
S9_6	S9_9	S9_4	S9_1	S9_8	S10_1
S9_1	S9_4	S9_6	S10_8	S9_9	S10_5
S9_4	S9_1	S9_9	S9_6	S9_8	S10_8
S10_5	S10_6	S10_8	S9_4	S9_1	S9_9
S10_10	S6_5	S6_6	S6_4	S6_7	S6_3
S10_6	S10_5	S10_8	S9_1	S9_4	S9_6
S10_8	S9_1	S10_5	S10_1	S10_6	S9_4
S10_1	S10_8	S9_4	S9_6	S9_1	S10_5

Во втором эксперименте количество классов было увеличено в два раза. Общее число портретов - 50. Изображения из первого опыта также использовались. Среди изображений были две пары похожих друг на друга людей (рис.3.2).

Изображения одного человека представляли собой один кластер в 30-мерном признаковом пространстве. Вычисление средних арифметических значений признаков позволило определить центры каждого из кластеров, а также расстояния от каждого портрета до центра его кластера. Для каждого класса были выбраны наиболее близкий и удалённый портреты по отношению к центру кластера, а также один из оставшихся. В табл. 3.3 названные портреты обозначены соответственно литерами N, F и M . Для каждого из них искались ближайшие пять портретов, аналогично первому опыту.

Из результатов опыта видно, что лишь для трёх тестовых изображений из 30 (*S5_10, S10_10* и *S17_3*) отобранные фото не содержат портретов искомых людей. Следует отметить, что названные изображения наиболее удалены от центров своих кластеров.

Если считать положительным результатом распознавания присутствие в отобранном множестве хотя бы одного портрета искомого человека, то в этом случае коэффициент распознавания составит 27/30, или 90%.

На основании вышеописанных результатов можно сделать вывод, что геометрический подход можно использовать для первоначального поиска в базе данных и отбора $k_1 = (\ln N)^2$ наиболее похожих изображений. Если считать, что для вычисления евклидового расстояния между двумя точками в 30-мерном пространстве необходимо произвести 30 арифметических и алгебраических операций (это предельный случай, использование более простых метрик требует меньших вычислительных затрат), то максимальное время поиска на первом этапе (без использования индексов) может быть рассчитано как $T = N^*90/t$, где t - количество операций в секунду (FLOPS) для определённой ЭВМ. Так как для персонального компьютера на базе процессора Pentium-III-500 $t \approx 20000$, то при $N = 10^5$ ($k_1 = 132$) $T = 450\text{c}$ (7,5 мин). Для обслуживания баз данных с большим количеством записей потребуются более мощные компьютеры.

3.1.3 Окончательный отбор портретов

Множество предварительно отобранных фотографий k_1 подвергалось более тщательному сравнению с исходным (рис.3.1). Оно основывалось на предположении, что два изображения, отображающие один объект с одного ракурса при похожих условиях, должны содержать большое количество пикселей с близкими значениями. Иными словами, если представить

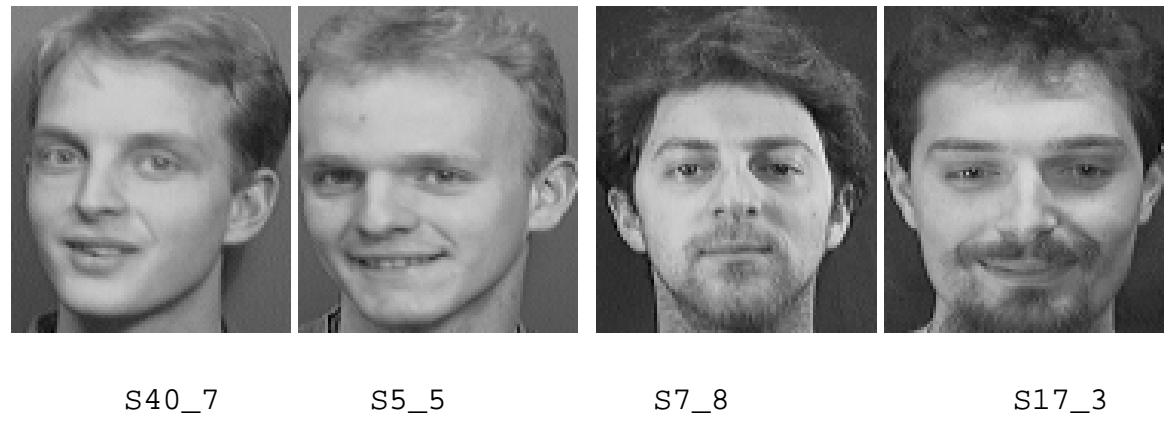


Рис. 3.2. Пары фотопортретов похожих людей в базе данных, на которой проводились эксперименты



Искомый портрет S27_3 и его центральная часть лица

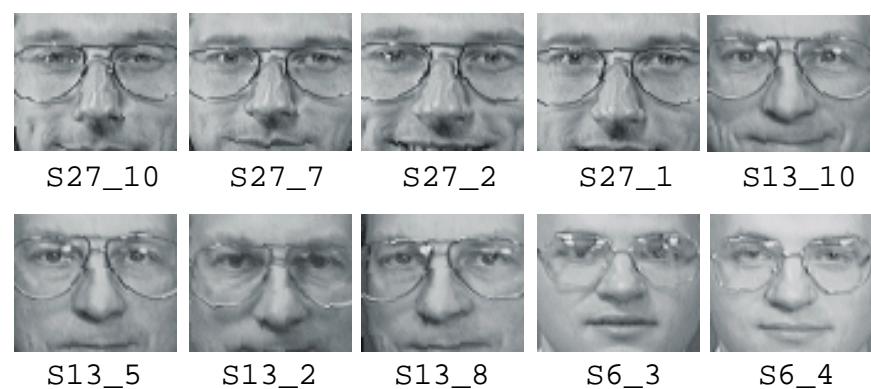


Рис.3.3. Пример результата отбора 10 ближайших портретов (последовательность C_1) в базе данных на основании только значений меры D . Искомый портрет - S27_3 (вверху), отобранные - расположены в порядке убывания близости, начиная с левого верхнего портрета S27_10 и заканчивая S6_4

Таблица 3.3

Результаты эксперимента №2

Тип портрета	Искомое изображение	Изображения, ближайшие к искомому				
		1	2	3	4	5
N	S5_6	S5_9	S5_2	S17_3	S40_6	S27_3
F	S5_10	S7_5	S7_4	S7_1	S6_5	S8_8
M	S5_9	S17_3	S5_6	S5_8	S5_2	S10_5
N	S6_5	S6_4	S6_7	S40_10	S40_7	S8_8
F	S6_3	S6_6	S6_5	S6_4	S6_7	S10_10
M	S6_4	S6_5	S8_8	S40_10	S40_7	S40_2
N	S7_5	S7_1	S7_4	S5_10	S7_8	S17_4
F	S7_7	S7_8	S7_1	S7_5	S7_4	S5_10
M	S7_8	S7_7	S7_1	S7_5	S7_4	S5_10
N	S9_9	S9_8	S9_4	S9_6	S9_1	S10_5
F	S9_8	S9_9	S9_4	S9_6	S40_6	S9_1
M	S9_6	S9_9	S9_4	S9_1	S9_8	S10_1
N	S10_5	S10_6	S10_8	S9_4	S9_1	S9_9
F	S10_10	S6_5	S6_6	S6_4	S8_5	S8_6
M	S10_6	S10_5	S10_8	S9_1	S9_4	S9_6
N	S8_6	S8_5	S8_8	S8_2	S17_6	S8_1
F	S8_8	S6_4	S8_5	S17_4	S8_8	S40_2
M	S8_2	S8_1	S8_6	S40_6	S17_6	S17_7
N	S13_5	S13_8	S13_10	S6_7	S6_5	S17_7
F	S13_3	S13_5	S13_10	S13_8	S6_7	S17_6
M	S13_8	S13_5	S13_10	S10_10	S6_5	S6_6
N	S17_6	S17_7	S17_9	S8_6	S17_4	S40_7
F	S17_3	S5_9	S40_5	S5_6	S5_2	S40_2
M	S17_4	S40_2	S8_8	S17_6	S17_9	S8_6
N	S27_3	S27_2	S27_1	S27_10	S40_6	S9_4
F	S27_7	S40_6	S27_3	S27_2	S5_6	S17_3
M	S27_1	S27_3	S27_2	S27_7	S9_4	S10_1
N	S40_10	S40_7	S6_5	S6_4	S40_2	S40_5
F	S40_6	S40_2	S9_8	S8_2	S27_3	S8_1
M	S40_5	S40_10	S40_2	S6_4	S6_8	S8_8

изображение в виде дискретной поверхности в трёхмерном пространстве: $A = \{(i,j, a_{ij})\}$, $0 \leq i,j \leq N$, то два пикселя, принадлежащие разным изображениям, но отображающие один и тот же элемент снимаемого объекта, должны располагаться сравнительно близко друг к другу в указанном пространстве. Простейшими вариантами измерения расстояния между двумя изображениями могут быть названы коэффициент корреляции и среднеквадратичное отклонение пикселей A от пикселей B .

В работах [36, 143, 146] была предложена мера различия изображений D и продемонстрировано её преимущество в точности оценки перед названными мерами, а также метрикой Хаусдорфа [36, 93, 154]:

$$D(A, B) = \frac{1}{2G} \left\{ \left(\frac{1}{N^2} \sum_{i,j=1}^N (f_2(a_{ij}, B) + f_2(b_{ij}, A))^E \right)^{1/E} \right\}, \quad (3.1)$$

где G – наибольшее значение яркости, f_2 – либо метрика d_{chess} либо d_{city} между пикселями двух изображений (метрика Евклида также приемлема, но требует больших вычислений), E может быть произвольным числом (наиболее универсальным является $E = 2$). Для упрощенных, более быстрых оценок можно использовать $E=1$.

Результаты экспериментальной проверки применения функции D для сравнения фотопортретов опубликованы в работе [144]. При этом использовалась шахматная метрика для вычисления расстояния между двумя пикселями:

$$d(a_{ij}, b_{lm}) = \max \{ |i-l|, |j-m|, |Ga_{ij} - Gb_{lm}| \}, \quad (3.2)$$

где i,j и l,m – пространственные координаты пикселей изображений A и B , соответственно, а Ga_{ij} , Gb_{lm} – их значения яркости. В качестве $f_2(a_{ij}, B)$ была использована следующая функция:

$$f_2(a_{ij}, B) = d(a_{ij}, B_W) = \min_{(lm) \in W} \{d(a_{ij}, b_{lm})\}, \quad (3.3)$$

где W – область поиска в виде квадратного окна размером $(2w+1) \times (2w+1)$ пикселей с центром в пикселе b_{ij} . В экспериментах использовался коэффициент $w = 4$. Аналогично вычислялось и $f_2(b_{ij}, A)$. Иными словами, для каждого пикселя a_{ij} искался ближайший по значению и расположению пиксель b_{lm} в пределах окна W . При этом перебор пикселей b_{lm} производился по спирали, начиная от центра b_{ij} . Алгоритм останавливался, если он находил ми-

нимально возможное $d(a_{ij}, b_{lm})$. Иначе $d(a_{ij}, b_{lm})$ вычислялось для всех $b_{lm} \in W$ и выбиралось минимальное из них.

Эксперименты проводились на 50 портретах вышеназванной базы данных ORL, к которым было добавлено 20 изображений двух человек, полученных на протяжении 30 лет. Из нормализованных портретов вырезались прямоугольники размером 55×66 пикселей (как показано на рис.3.3 для S27_3), содержащие центральную часть лица человека.

На рис. 3.3 приведена последовательность C_1 портретов людей, близких к S27_3, отобранных с помощью меры D из всей базы. Следует отметить, что все оставшиеся изображения искомого человека S27 занимают первые четыре позиции найденных изображений, а все остальные изображения - портреты людей в очках.

Если проводить поиск того же портрета S27_3 с помощью указанного выше геометрического метода, то результатом будет упорядоченное множество C_2 (рис.3.4). Однако после анализа признаков и сокращения их числа до 28 (исследования подробно рассмотрены в разделе 3.1.4), ближайшие 10 портретов к S27_3 составляют множество C_3 (рис.3.5).

После применения меры D к множеству C_3 оно преобразуется в множество C_4 , на первых позициях которого стоят все оставшиеся четыре портрета искомого человека (рис. 3.6).

Благодаря отбору наиболее информативных признаков коэффициент распознавания улучшился и составил 98,5%. Введение второго этапа поиска улучшило результаты поиска и в итоге лишь в двух случаях из 70 ближайшим к заданному портрету было названо изображение, принадлежащее другому классу.

Несмотря на очевидное преимущество в надёжности сравнения портретов мерой D , скорость её вычисления почти на два порядка ниже, чем определение близости геометрическим подходом.

Таким образом, на втором этапе поиска из k_1 изображений отбирается k_2 , где $\ln k_1 < k_2 < (\ln k_1)^2$ локально наиболее близких к изображению искомого человека. При $100\ 000 < N < 1\ 000\ 000$ $k_2 = 5$.

Из k_2 отобранных портретов окончательный отбор производится оператором, так как только человек может выделить небольшую общую деталь, отличающую одного человека от других (шрам, родинка).

3.2 Формирование признакового пространства

При проектировании систем автоматического распознавания возникает задача минимизации признакового пространства. Подход,

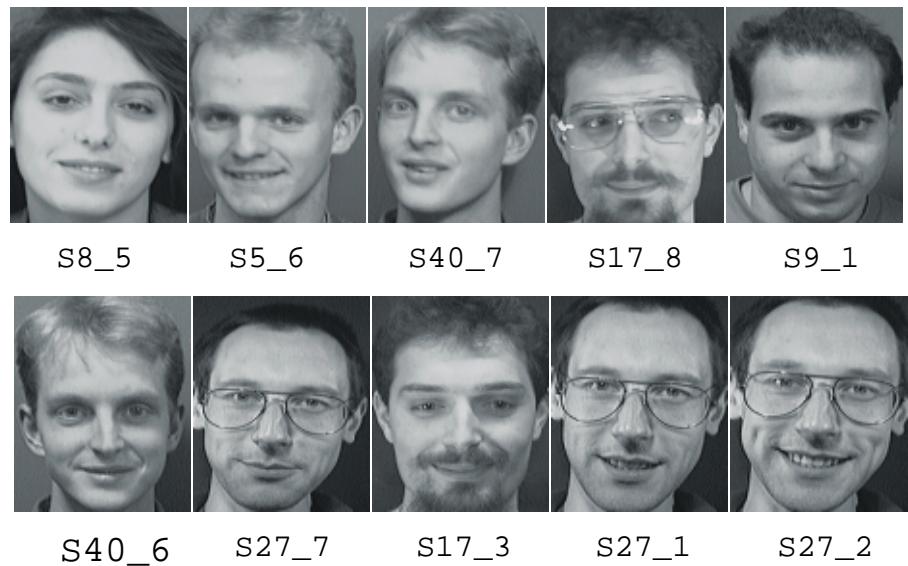


Рис.3.4. Пример результата поиска 10 ближайших портретов (последовательность C_2) в базе данных с помощью набора признаков, описанных в разделе 3.1.2. Искомый портрет S27_3 (рис.3.3), отобранные расположены в порядке убывания близости, начиная с левого верхнего портрета S8_5 и заканчивая S27_2

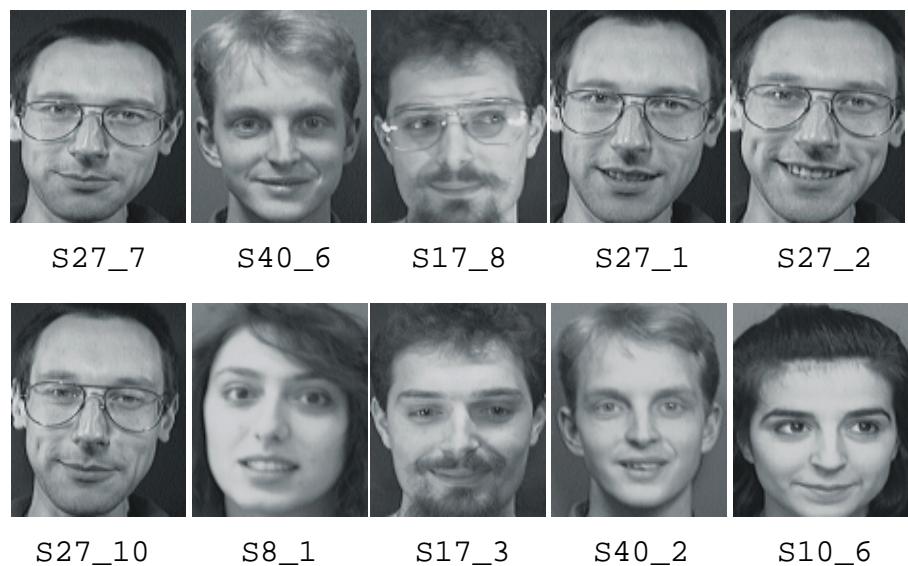


Рис.3.5. Пример результата поиска 10 ближайших портретов (последовательность C_3) в базе данных с помощью усечённого набора признаков, описанных в разделе 3.1.4. Искомый портрет S27_3 (рис.3.3), отобранные расположены в порядке убывания близости, начиная с левого верхнего портрета S27_7 и заканчивая S10_6

рассмотренный в данном разделе, является универсальным и позволяет отбирать наиболее информативные признаки при решении широкого круга задач, связанных с распознаванием. Его возможности демонстрируются на задаче идентификации человека по фотопортрету. Эксперименты проводились на базе данных, описанной в предыдущем разделе. Общее количество тестировавшихся портретов – 70 полутонаовых изображений разных размеров.

3.2.1 Оценка признаков на основе статистических данных

Были использованы два способа оценки признаков. Первый, основанный на статистических данных, заключался в анализе отношения $F = D_d / D_m$. Величины D_d и D_m определялись следующим образом. Для элементов каждого i -го кластера вычислялись среднее арифметическое M_{li} признака Pr_l - и его дисперсия - D_{li} :

$$M_{li} = \frac{\sum_{j=1}^{g_i} Pr_{lji}}{g_i}, \quad D_{li} = \sqrt{\frac{\sum_{j=1}^{g_i} (M_{li} - Pr_{lji})^2}{g_i - 1}}, \quad (3.4)$$

где i – номер кластера; g_i – количество элементов i -го кластера, т. е. количество портретов i -го человека; l - номер признака. Затем определялись среднее M_{lm} и дисперсия D_{lm} от величин M_{li} :

$$M_{lm} = \frac{\sum_{i=1}^k M_{li}}{k}, \quad D_{lm} = \sqrt{\frac{\sum_{i=1}^k (M_{lm} - M_{li})^2}{k - 1}}. \quad (3.5)$$

Здесь k – количество кластеров, т.е. человек, а также среднее M_{ld} и дисперсия D_{ld} величин D_{li} :

$$M_{ld} = \frac{\sum_{i=1}^k D_{li}}{k}, \quad D_{ld} = \sqrt{\frac{\sum_{i=1}^k (M_{ld} - D_{li})^2}{k - 1}}. \quad (3.6)$$

Геометрический смысл отношения $F_l = D_{ld} / D_{lm}$ следующий: чем больше для данного Pr_l дисперсия величин M_{li} , тем шире диапазон колебания признака при описании различных людей, тем, следовательно, выше ценность данного признака для их различения. Чем меньше дисперсия величин D_{li} , тем меньше колебание признака Pr_l для портретов одного и того же человека, т.е. тем кучнее будут располагаться в признаковом пространстве элементы одного кластера. Следовательно, чем больше D_{lm} и одновременно чем



Рис.3.6. Пример результата поиска 10 ближайших портретов (последовательность C_4) в базе данных с помощью усечённого набора признаков и меры D . Искомый портрет $S27_3$ (рис.3.3), отобранные расположены в порядке убывания близости, начиная с левого верхнего портрета $S27_10$ и заканчивая $S17_8$

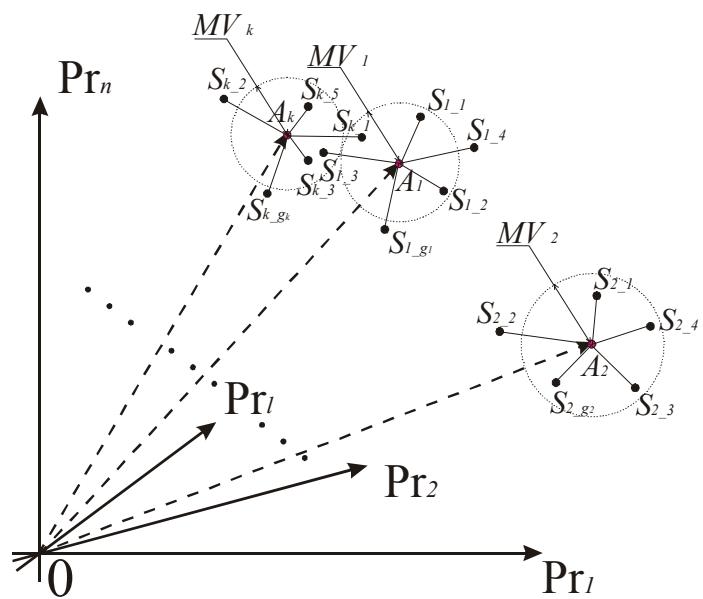


Рис. 3.7. Пример расположения кластеров в n -мерном признаковом пространстве

меньше D_{ld} , тем лучше признак выделяет образы одного кластера, а величина $F_l = D_{ld}/D_{lm}$ для такого признака будет стремиться к 0. Такая оценка признака Pr_l не зависит от значений других признаков и от изменения размерности признакового пространства, т.е. любой признак, претендующий на включение в итоговый набор, может быть оценен один раз и сравнен с остальными признаками без пересчета их оценок.

Для того чтобы проверить эффективность описанной выше оценки ранжирования признаков, использовался второй способ. Он основан на оценке изменений, происходящих во взаимном расположении образов в признаковом пространстве. Множество портретов одного человека составляет в нем один кластер. Для повышения качества распознавания необходимо выбрать такое пространство признаков, в котором расстояние между любыми образами одного кластера будет меньше, чем расстояние между любыми двумя образами разных кластеров. Чем дальше будут разнесены в пространстве кластеры друг от друга, тем выше вероятность того, что при увеличении числа кластеров на несколько порядков результаты распознавания будут успешными. Для проверки важности признака необходимо знать, как он влияет на расположение образов в пространстве. Для этого, определив средние значения признаков для образов каждого кластера и приняв их за координаты центра кластера A_i (рис. 3.7), вычислялись средние арифметические расстояния между центром A_i и образом S_{i-j} . Иными словами, определялся средний радиус MV_i n -мерной сферы в признаковом пространстве, в среднем описывающей данный кластер:

$$A_i = \left(\frac{\sum_{j=1}^{r_i} Pr_{1ij}}{g_i}, \frac{\sum_{j=1}^{r_i} Pr_{2ij}}{g_i}, \frac{\sum_{j=1}^{r_i} Pr_{3ij}}{g_i}, \dots, \frac{\sum_{j=1}^{r_i} Pr_{nij}}{g_i} \right), \quad (3.7)$$

где Pr_{nij} – значение признака Pr_n для j -го образа i -го кластера.

$$\vec{V}_{ij} = \overrightarrow{OA_i} - \overrightarrow{OS_{i-j}}, \quad (3.8)$$

$$MV_i = \frac{\sum_{i,j} |\vec{V}_{ij}|}{g_i}, \quad (3.9)$$

$$T = \frac{\sum_{i=1}^k MV_i}{k}. \quad (3.10)$$

Вычисление первой оценки $F_l=D_{ld}/D_{lm}$ было выполнено один раз для всех признаков. Радиусы кластеров MV_i пересчитывались каждый раз после изменения признакового пространства.

3.2.2 Экспериментальная проверка подхода

Цель экспериментов состояла в подтверждении или опровержении правильности первоначальной оценки $F_l=D_{ld}/D_{lm}$ ранжирования признаков. Идея заключалась в поочерёдном исключении признаков из базового набора, анализе изменений положения кластеров в признаковом пространстве и в оценке на этой основе признаков.

Эксперименты выполнялись следующим образом: поочерёдно каждый из 70 портретов становился тестируемым, а оставшиеся 69 - базой данных, на которой проводился поиск, т.е. фактически распознавание выполнялось 70 раз, но для краткости изложения ниже под словом «распознавание» будем подразумевать вышеописанную процедуру. При этом в выбранном признаковом пространстве искали пять наиболее близких к тестируемому изображению портретов. Основной оценкой результатов был коэффициент распознавания, показывающий, сколько раз из 70 в пятёрке ближайших был хотя бы один портрет искомого человека.

Эксперимент состоял из пяти опытов. В первом опыте была вычислена величина F_l для каждого из признаков (рис. 3.8). Затем проводилось распознавание с уменьшенным набором признаков. Поочерёдно исключался каждый признак. Кроме коэффициента распознавания каждый раз вычислялось среднее арифметическое радиусов кластеров T . На рис. 3.9 показано количество корректных распознаваний K , на рис. 3.14 – значения T при поочерёдном исключении признака Pr_l . Из обоих графиков видно, что при не использовании признака Pr_4 результаты распознавания улучшаются. Данный признак является нормированным расстоянием между точками FL1 и FR1 (см. рис. 2.1) и косвенно описывает ширину подбородка человека. Его исключение из базового набора приводит к существенному уменьшению средних радиусов кластеров (на 10%), а также к повышению коэффициента распознавания до $K=69/70$. Однако несмотря на то, что на втором месте по возрастанию величины T и на третьем по индексу $F_l=D_{ld}/D_{lm}$ стоит признак Pr_3 , его отсутствие в признаковом наборе приводит к уменьшению коэффициента распознавания до $66/70$. Возможное объяснение данному факту: при исключении Pr_3 , несмотря на уплотнение элементов кластеров относительно их центров, расстояния между кластерами тоже существенно изменились, и не в лучшую сторону, т.е. кластеры стали ближе друг к другу.

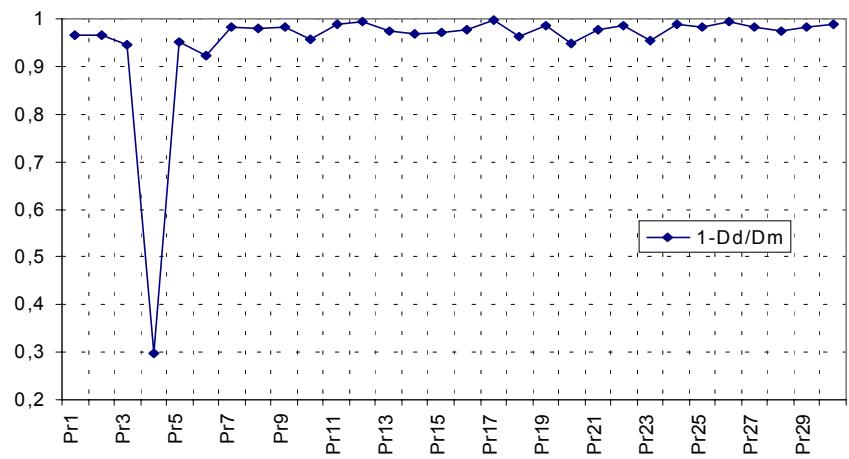


Рис. 3.8. Значение величины $f(Pr_l)=1-F(Pr_l)$ для всех признаков, $l \in [1, 30]$. Из графика видно явное несоответствие признака Pr_4 данному набору

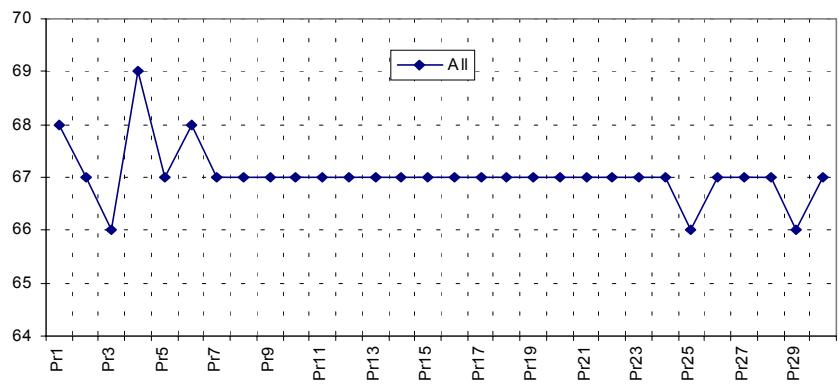


Рис. 3.9. Результаты распознавания в первом эксперименте после поочерёдного исключения Pr_l из начального множества признаков

Последующие опыты были аналогичны первому, разница заключалась в том, что начальное количество признаков было не n , а $(n-1)$, т.е. 29 вместо 30. Из базисного набора были исключены признаки во втором опыте – Pr_4 , в третьем – Pr_3 , в четвёртом – Pr_1 и в пятом – Pr_6 . Далее воспроизводилась процедура первого эксперимента, т.е. исключался l -й признак и фиксировались изменения в расположении кластеров.

Результаты второго опыта приведены на рис. 3.10. В трёх случаях из 29 коэффициент распознавания K улучшился на 1/70, в остальных случаях улучшение составило + 2/70. Несмотря на то, что согласно ранжированию по T признак Pr_{10} занимает шестое место, при его удалении из базисного набора достигается наибольшее изменение минимального и максимального радиусов для отдельных кластеров (см. рис. 3.8). Физический смысл признака Pr_{10} : расстояние между точками LD и TL – высота подбородка.

Основным результатом третьего опыта является подтверждение неизменности признака Pr_3 (рис. 3.11). Несмотря на то, что по предварительным результатам он являлся вторым (после Pr_4) кандидатом на исключение, его исключение приводит к значительному ухудшению результатов распознавания. Физический смысл этого признака – ширина лица человека на уровне линии рта. По сравнению с результатами при использовании полного набора из n признаков только в одном случае – при одновременном исключении Pr_3 и Pr_4 – результат улучшился на 1/70 и в двух остался неизменным. Во всех остальных случаях исключение признака Pr_3 вместе с каким-либо другим приводило к ухудшению результатов распознавания.

Изменения величин T для подавляющего большинства признаков незначительны – в пределах 2-3 %, лишь при исключении Pr_4 разница в величинах при использовании и неиспользовании признака порядка 15% (см. рис. 3.8).

В опыте 4 исключали Pr_1 (восьмой по значению T) признак. Можно констатировать, что при дополнительных исключениях признаков существенного улучшения результатов распознавания не наблюдается. Лишь в двух случаях - при неиспользовании Pr_4 и Pr_6 – был получен коэффициент 69/70, в остальных - 68/70 и 67/70 (рис. 3.12). Улучшение величины T наблюдалось лишь при дополнительном удалении признака Pr_4 (см. рис. 3.8). Изменение величины T оставалось в пределах 5% относительно результатов распознавания с 29 признаками. Физический смысл признака Pr_1 - ширина лица на уровне глаз.

При удалении признака Pr_6 – расстояния между правым уголком рта и точкой пересечения вертикали, проходящей через правый уголок глаза, с нижним краем лица – коэффициент распознавания составил 68/70, а при

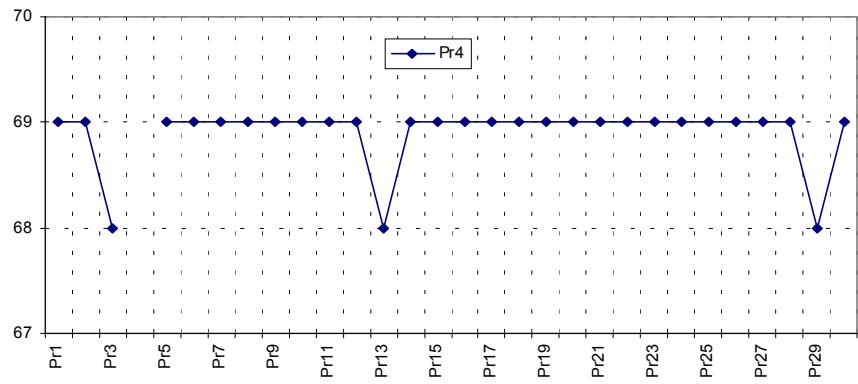


Рис. 3.10. Результаты распознавания во втором эксперименте – поочерёдное исключение Pr_1 при уже исключённом из базового набора Pr_4

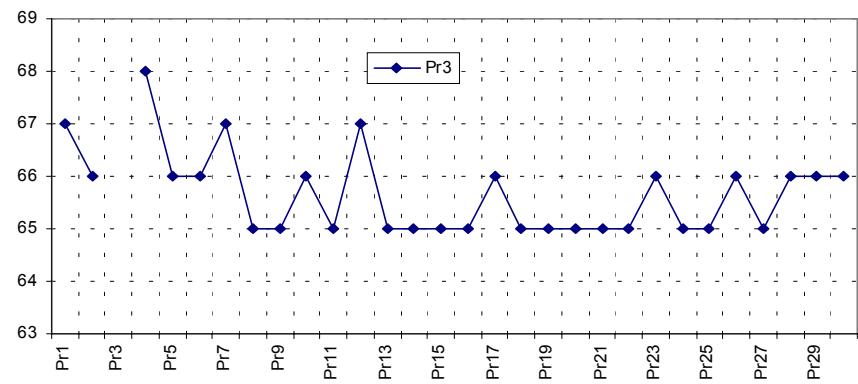


Рис. 3.11. Результаты распознавания в третьем опыте – исключённый признак Pr_3

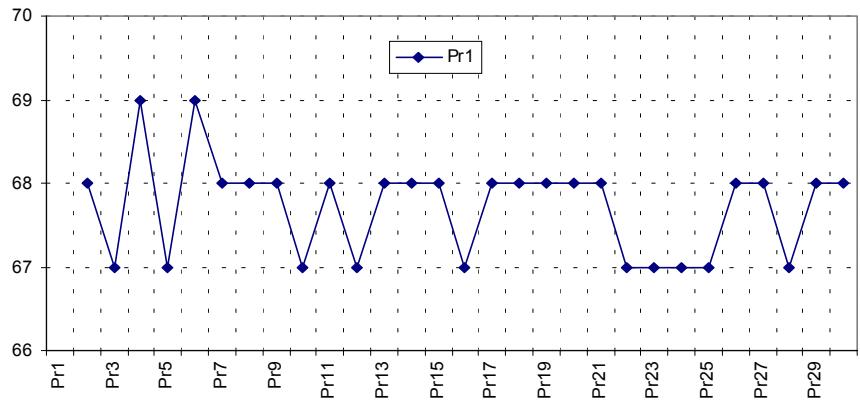


Рис. 3.12. Результаты распознавания в четвертом эксперименте после исключения Pr_1 и Pr_l

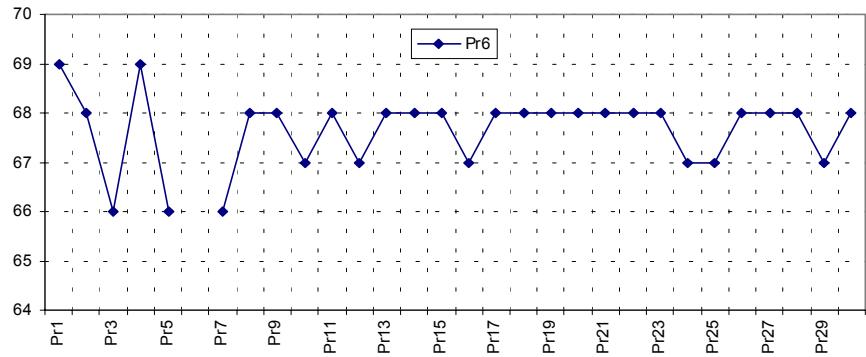


Рис. 3.13. Пятый опыт. Результаты распознавания при одновременном исключении Pr_6 и Pr_l из базового набора признаков

исключении симметричного ему признака Pr_5 – расстояния между такими же точками левой стороны лица – 67/70. Интересно также, что зависящий от Pr_5 и Pr_6 признак $Pr_{20} = (Pr_5+Pr_6)/2$ являлся в большей степени приводящим к изменениям средних радиусов, но в меньшей степени изменяющим результаты распознавания (рис. 3.8). Явное улучшение результатов наблюдалось при дополнительном удалении Pr_1 и Pr_4 - коэффициент распознавания составил $K=69/70$ (рис. 3.13). Итоговый график изменения средних радиусов T кластеров при исключении признака Pr_l во время выполнения вышеописанных опытов показан на рис.3.14.

Примером улучшения поиска на основе усечённого набора признаков является преобразование последовательности найденных портретов C_2 в C_3 (рис. 3.4, 3.5).

Вывод: Применение описанной оценки F_l позволяет сократить набор признаков, оставив наиболее информативные из них.

3.3 Сравнение фотопортретов методом эластичных экспоненциальных деформаций

Решение задачи паспортного контроля требует использования методов распознавания, которые базируются на сравнении лишь двух изображений А и В. Основная сложность её решения заключается в нахождении достаточного количества общих или различающих признаков на сравниваемых изображениях для уверенного ответа на вопрос об идентичности объектов. Определить необходимый минимум таких признаков, основываясь лишь на паре фотографий и часто не имея возможности использовать какую-либо априорную информацию об изображенных на них объектах, представляется сильно затруднительным. Поэтому исследователи, работающие в данном направлении, предложили методы, базирующиеся на деформации одного изображения с целью превращения его в другое и оценке «усилий», необходимых для её выполнения [13, 83, 84, 181].

В общем виде оценить необходимые усилия по деформации одного изображения (деформируемого, B) в другое (эталон, A) можно, допуская лишь три варианта преобразования:

1. Изменение яркостей пикселей B до уровня яркостей соответствующих им пикселей A .
2. Изменение координат пикселей (деформация решётки изображения) с целью уменьшения расстояния в пространственно-яркостном континууме (3D) между пикселями эталона и совмещаемого с ним изображения, близкими по уровню яркости.

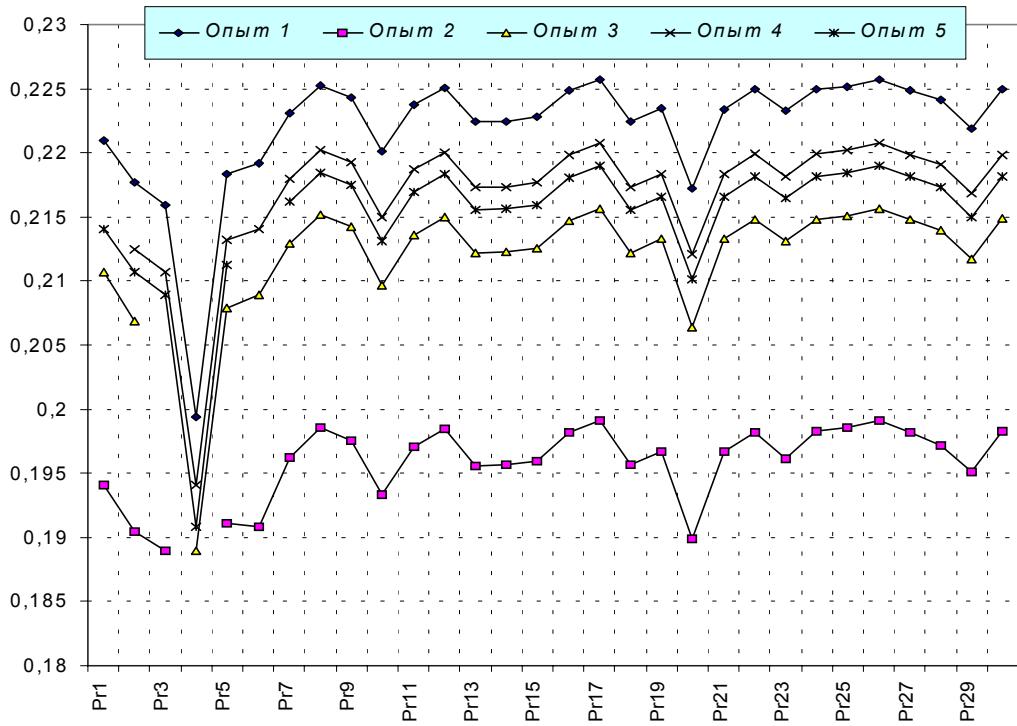


Рис. 3.14. Изменения средних радиусов T кластеров на протяжении пяти опытов: 1-й опыт – из базового набора признаков был поочерёдно исключён каждый признак Pr_i ; 2-й опыт – исключали Pr_i при уже исключённом Pr_4 ; 3-й опыт – исключали Pr_i при уже исключённом Pr_3 ; 4-й опыт – исключали Pr_i при уже исключённом Pr_1 ; 5-й опыт – исключали Pr_i при уже исключённом Pr_6

3. Изменение как значений яркостей пикселей, так и решётки изображения – координат пикселей [110].

В первом и третьем случаях имеет место более общее преобразование, нежели деформация. Так как нас интересует не сам процесс получения эталонной картинки из совмещаемой с ним, т.е. не видеоряд превращения одного изображения в другое, а оценка требуемых для этого усилий, то в первом случае простейший вариант такой оценки можно получить, например, по формуле

$$D = \frac{\sum_{i=1}^n \sum_{j=1}^m |a_{ij} - b_{ij}|}{Gnm}, \quad (3.11)$$

где a_{ij} , b_{ij} – яркости пикселей с координатами i, j эталона и совмещаемого изображения, G – количество уровней яркости, n , m – количество столбцов и строк в изображениях. Можно использовать коэффициент корреляции между A и B :

$$D = \frac{\sum_{i=1}^n \sum_{j=1}^m a_{ij} \times b_{ij}}{\sqrt{\sum_{i=1}^n \sum_{j=1}^m a_{ij}^2 \times \sum_{i=1}^n \sum_{j=1}^m b_{ij}^2}}, \quad (3.12)$$

а также любую формулу, операндами которой являются лишь значения яркостей пикселей и величины, определяющие количество самих пикселей.

Получить объективную оценку близости изображений при преобразованиях, затрагивающих как значения яркости пикселей, так и решётку изображения (третий случай), можно лишь следующим образом. Сначала для каждого пикселя деформируемого изображения находится минимальное из всех возможных изменение, как его уровня яркости, так и координат, которое приводит к получению соответствующего пикселя эталона, а затем полученные по всем пикселям изображения данные некоторым образом обобщаются в единую величину. Для этих целей может успешно применяться мера D , дающая требуемую оценку (раздел 3.1.3). Указанная мера учитывает как разницу в яркостях пикселей, так и расстояние между ними в матрице изображения. Сравнение нормализованных изображений лиц путём оценки деформаций изображений с применением некоторой модели искажения поверхности было проведено Настаром и Мохаддом [119], что также относится к третьему виду преобразований.

Однако несмотря на относительную простоту формул (3.1, 3.11, 3.12), описанные выше подходы имеют существенные недостатки. В частности,

за счёт возможного изменения яркости в процессе деформации наблюдается потеря чувствительности итоговой оценки к изменению формы объекта. При решении задачи паспортного контроля был использован метод эластичных экспоненциальных деформаций, описанный в [13]. Он был адаптирован в соответствии со спецификой сравнения цифровых фотопортретов.

3.3.1. Суть подхода. Деформируемое изображение, дискретная решётка которого представляется в виде модели эластичной мембраны, должно быть совмещено с эталоном. В соответствии с этой моделью любой элемент решётки i может быть перемещён на позицию элемента j , однако при этом изменяются и координаты соседних элементов. Характер перемещения, а также определение того, какие именно из соседей должны переместиться, зависят от параметра "пластичности" среды s :

$$d_k = d_{ij} e^{-sd_{ik}}, \quad (3.13)$$

где d_{ij} – расстояние между пикселями i и j , k – пиксель, лежащий на расстоянии d_{ik} от пикселя i , d_k – величина сдвига элемента решётки k в направлении, задаваемом вектором \bar{ij} . Совмещение производится путём многократного изменения дискретной решётки изображения, "подгоняемого" к эталону. На каждой итерации случайным образом выбирается элемент решётки i и вектор его перемещения, вычисляется характер изменения позиций других элементов решётки. Чем «жёстче» мембрана, тем большее количество элементов решётки меняют свои координаты при каждой деформации, и наоборот, чем «мягче», тем меньшее количество соседей затрагивает перемещение одного выбранного элемента.

Деформация оценивается с точки зрения минимизации затраченной энергии и улучшения функции близости сравниваемых изображений. Функция близости может быть выбрана исходя из специфики задачи, в том числе и из формул (3.1), (3.11) и (3.12). Если в результате произошедших изменений общая оценка близости изображений улучшается, то в следующей итерации участвует полученное изображение. В противном случае оно отбрасывается и используется предыдущая деформация. Процесс останавливается, если определённое количество итераций не улучшает итоговую оценку либо было выполнено максимально разрешённое количество деформаций.

В первую очередь при адаптации метода к задаче контроля документов ставилась цель снижения итогового количества итераций (в [13] для сравнения изображений требовалось порядка 20 – 50 тыс. итераций в

зависимости от их размера), а также снижения влияния случайности на результирующую оценку. Иными словами, на модель деформирования были наложены требования: уменьшить количество не приводящих к положительному результату деформаций и снизить зависимость результата сравнения от последовательности выполнения деформаций.

Спецификой решаемой задачи можно назвать тот факт, что на сравниваемых изображениях имеются по паре точек (центры зрачков), которые при любых деформациях не должны изменять своего положения. Кроме того, элементы решётки, находящиеся в непосредственной близости к названным точкам, не должны сильно деформироваться, а форма зрачков должна сохраняться:

$$d_k = d_{ij} e^{-sd_{ik} f(l)_k f(r)_k}, \quad (3.14)$$

$$f(l)_k = e^{-sq(d_{lk}, R)}, \quad (3.15)$$

$$q(d_{lk}, R) = \begin{cases} 1, & d_{lk} \geq R; \\ R - d_{lk}, & d_{lk} < R. \end{cases}$$

Здесь R - радиус зрачка на изображении; l , r - центры левого и правого зрачков; d_{lk} - расстояние от пикселя k до центра зрачка l .

С помощью этих ограничений удалось исключить случаи недопустимых искажений лица типа «превращение глаза в бровь».

Исходя из антропометрических данных [8, 14], были определены области изображения, которые с высокой степенью вероятности содержат четыре черты лица: левую и правую брови, нос, рот. Учитывая тот факт, что при любом наклоне/подъёме/повороте головы на незначительный угол (в рамках допустимых значений, указанных в разделе 2.1) упомянутые черты лица остаются в пределах областей определённых в соответствии с табл. 3.4, максимальная длина вектора деформации \bar{y} при условии, что i принадлежит указанным областям, была ограничена величиной, равной половине высоты этой области (прямоугольника). Следует отметить, что на изображениях одного человека форма брови, например, существенно не изменяется (при соблюдении ограничений на выражение лица и ракурс - раздел 2.1), а меняется лишь её положение относительно вертикальной координаты центров зрачков (уровня глаз) либо вертикальной оси лица, чего нельзя утверждать по отношению, например, к носу. Это потребовало замены модели эластичной мембраны с постоянным для всех областей лица параметром

пластичности s на модель мембранны с переменным s , зависящим от того, к какой области изображения принадлежит выбранный элемент решётки.

Учитывалась также и форма черт лица при деформации. При очередной итерации, если пиксель i не принадлежит перечисленным выше областям, то $d_{ik} = \sqrt{x_{ik}^2 + y_{ik}^2}$, где x_{ik}, y_{ik} – расстояния в пикселях между столбцами и рядами элементов решётки i и k . В противном случае, если i находится в указанных выше областях, расстояние до затрагиваемых текущей деформацией пикселей рассчитывается по формуле $d_{ik} = \sqrt{(\alpha x_{ik})^2 + (\beta y_{ik})^2}$ (при $i \in B_r$ или $i \in B_l$ в экспериментах использовались $\alpha=1, \beta=2$, при $i \in N - \alpha=2, \beta=1$, при $i \in M - \alpha=1, \beta=1$). Таким образом достигается неравномерное смещение пикселей области лица в выбранном направлении.

Таблица 3.4

Координаты окон поиска антропометрических точек

Область лица	Обозначение	Координаты левого верхнего угла*	Координаты правого нижнего угла
Левая бровь	B_r	(-30,25)	(25,2)
Правая бровь	B_l	(25,25)	(80,2)
Нос	N	(0,-15)	(50,-50)
Рот	M	(-5,-45)	(55,-75)

* Прим: Координаты даны в пикселях, начало координат находится в центре левого зрачка, ось ОY направлена вверх, а расстояние между зрачками равняется 50 пикселям.

Для уменьшения зависимости результата сравнения изображений от условий освещённости человека в момент съёмки и контрастности снимков при оценке каждой деформации в проведённых экспериментах использовался модифицированный коэффициент ранговой корреляции [12] в следующем виде:

$$v_r = \frac{1}{nm} \sum_i^n \sum_j^m (K_A(a_{ij}) - K_B(b_{ij}))^2, \quad (3.16)$$

где $K_A(a_{ij}) = \sum_{l=1}^{a_{ij}-1} H_A(l) + \frac{1}{2}(H_A(a_{ij}) + 1),$ (3.17)

H_A – гистограмма изображения А. $K_B(i)$ вычислялся для изображения В аналогично $K_A(i).$

Кроме того, учитывалась степень искажения решётки деформируемого изображения (величина U_2 из [13]):

$$\mathcal{V}_d = \frac{1}{N_d} \sum_{k \in A_d} [(x_{k+1} - x_{k+1} - 1) + (y_{k+1} - y_{k+1}) + (x_{k+n} - x_k) + (y_{k+n} - y_k - 1)]. \quad (3.18)$$

Здесь A_d – область деформации изображения A ; N_d – количество пикселей области A_d ; x_k, y_k – координаты пикселя k после деформации; n – количество пикселей в строке сравниваемых изображений.

Таким образом, итоговая оценка различия двух изображений может быть представлена в виде

$$\mathcal{U} = \mathcal{U}_r + \alpha \mathcal{V}_d, \quad (3.19)$$

где α – коэффициент, определяющий вес \mathcal{V}_d в итоговой оценке. В экспериментах применялось $\alpha = 0.1$.

Из выражений (3.16) – (3.19) видно, что $\mathcal{U} = 0$ в случае, если изображения A и B идентичны. Следовательно, целью отбора деформаций является минимизация величины \mathcal{U} . При сравнении фотопортретов одного человека, чем меньше данная оценка, тем лучше. В случае подачи на вход алгоритма портретов разных людей желательно, чтобы \mathcal{U} принимала как можно большее значение.

На рис. 3.15 показан пример работы алгоритма сравнения двух нормализованных фотопортретов, основанного на описанном методе. Для возможности визуальной оценки произведённых деформаций разрывы в итоговом изображении были устранены с помощью метода билинейной интерполяции. Результат работы всего комплекса алгоритмов приведён на рис. 3.16 (в данном случае интерполяция разрывов не производилась).

3.3.2. Экспериментальная проверка. Эффективность предлагаемого метода проверялась в экспериментах на реальных изображениях паспортов, сканированных при помощи «Считывателя паспортов» (мод. 7001, производитель НПП «Регула», Республика Беларусь), и фотопортретах, полученных с видеокамеры. Все изображения были полуточновыми с 256 уровнями яркости и удовлетворяли ограничениям, сформулированным в разделе 2.1. Обработка и сравнение изображений производились на ПЭВМ класса Pentium II-600.

Из-за специфики (природы) метода эластичных экспоненциальных деформаций результаты распознавания с его помощью сильно зависят от следующих факторов:

- метода оценки очередной деформации (обычная корреляция, городская метрика, ранговая корреляция, мера D и т.д.).

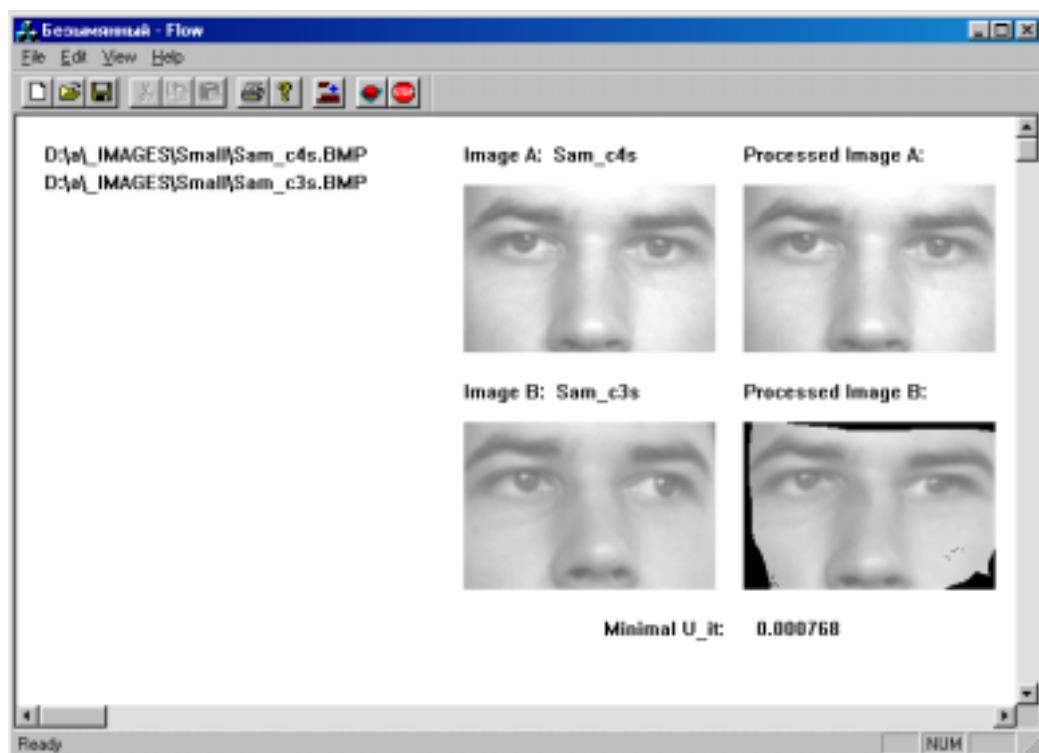


Рис.3.15. Пример эластичного деформирования нормализованного фотопортрета с целью максимальной подгонки к эталону. Исходные изображения – слева. Эталон – справа вверху, результат деформации – справа внизу. Разрывы в изображении интерполировались

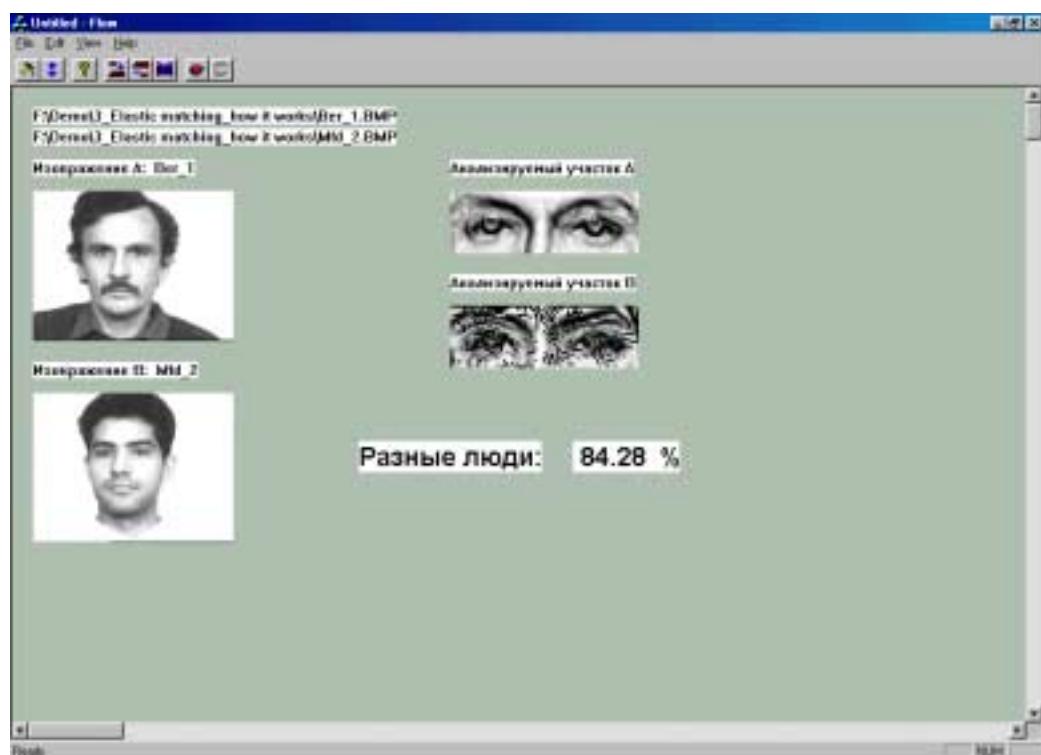


Рис.3.16. Пример сравнения портретов людей по нормализованным областям лица. Исходные изображения – слева. Эталон – справа вверху, результат деформации – справа внизу. Интерполяция разрывов не производилась

- ограничений, налагаемых на модель деформируемой решётки изображения (без ограничений - по Ковалеву, с зафиксированными центрами зрачков, с дополнительными ограничениями на направления и размер векторов смещения точки, с изменяющимися значениями параметра пластичности s в зависимости от области лица, к которой принадлежит пиксель i).
- формы, размера, частей сравниваемых изображений (в том числе и масштаба – расстояния между зрачками ED, см. рис. 2.16, 3.18).
- вариантов предварительной обработки изображений (с начальными значениями яркости пикселей, после применения операторов Собеля [98], Дериша [126] или комбинированного подхода «High Boosting», учитывавшего как информацию об исходных значениях пикселей, так и значения градиента [91], см. рис. 2.21, 3.18).

Кроме того, результат также может зависеть и от изменения шага квантования – количества уровней яркости изображения.

В связи с вышесказанным, экспериментально проверить все варианты самого алгоритма и вида представления входных данных на значимой выборке изображений не представлялось возможным. Поэтому параметры алгоритма, отобранные в процессе отдельных тестов и указанные выше, не изменились, а эксперименты проводились с целью выбора вариантов нормализации для наилучшего представления портретов, а также оценки временных и качественных характеристик (коэффициента распознавания) алгоритма. Было проведено два эксперимента.

Эксперимент №1. В разд. 3.2.1. было отмечено, что время работы алгоритма может быть различным в зависимости от того, остановился процесс деформации сам либо выполнил максимально разрешённое количество итераций M . Кроме того, из-за использования случайных величин последовательность выполнения деформаций будет уникальна как при повторном сравнении одной и той же пары изображений, так и при подаче на вход алгоритма других портретов. Очевидно, что время работы алгоритма будет зависеть от M , за исключением случаев досрочной остановки – при сравнении сильно отличающихся изображений, а размер поданных на вход портретов ($S = m \times n$) существенно влияет на время выполнения одной итерации. Диаграмма зависимости времени сравнения двух изображений $t(M, S)$ от указанных выше факторов приведена на рис. 3.17. Из рисунка видно, что зависимость $t(M)$ имеет линейный характер, однако скорость роста функции для изображений различных размеров $t = aM + b$ прямо пропорциональна S : чем больше S , тем больше a . Характер функции $t(S)$ меняется от

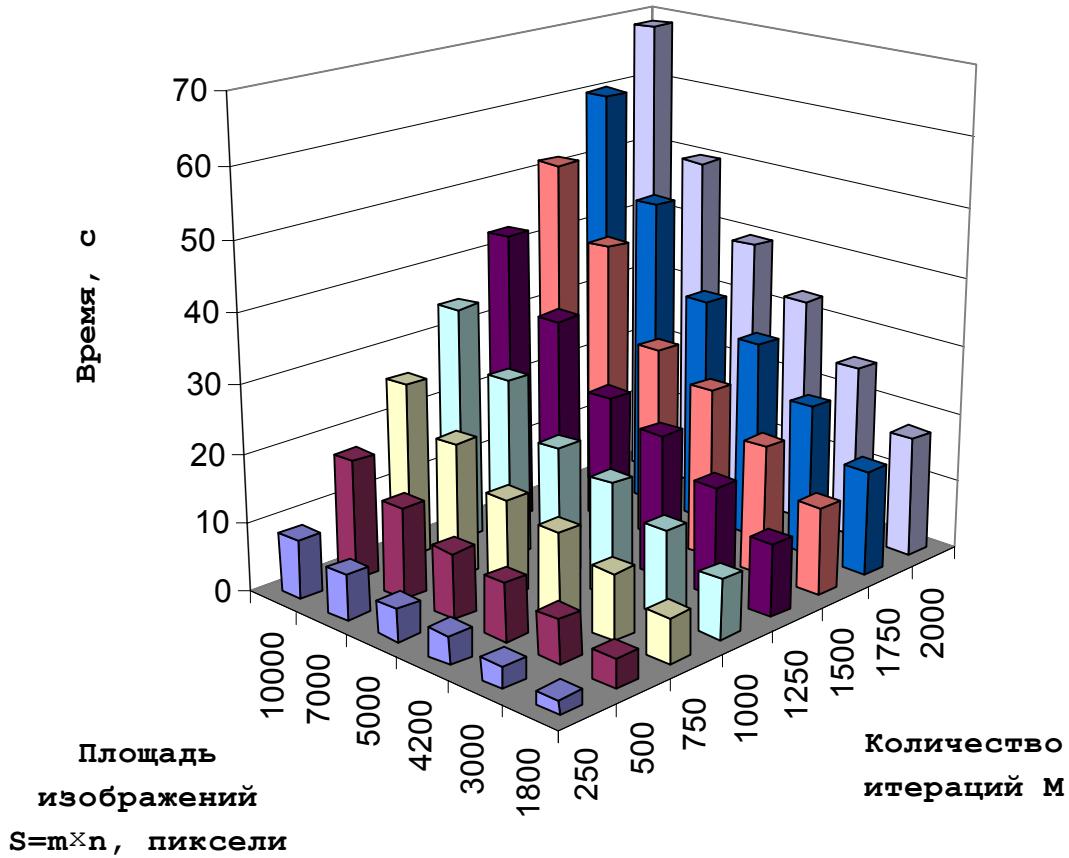


Рис. 3.17. Диаграмма зависимости времени работы алгоритма сравнения фотопортретов на базе эластичных экспоненциальных деформаций от размеров сравниваемых нормализованных изображений S и максимально разрешённого количества итераций M

линейного (при малых M) до параболического (при $M=2000$). Таким образом, время сравнения пары фотопортретов может колебаться от 2 до 68,5 с.

Эксперимент №2. Целью второго эксперимента было определение лучшей формы представления исходных для алгоритма данных с точки зрения как времени работы, так и точности результата верификации. Размеры, масштаб и форма, одним словом типоразмеры $U=\{U_1, \dots, U_l\}$, входных портретов были выбраны на основании предварительных тестов и указаны в табл. 3.5. Два оператора – Дериша и «High Boosting» – были отобраны в процессе анализа вариантов обработки исходных изображений для их нормализованного представления. Каждый из них применялся с тремя различными значениями своих параметров - α и f соответственно. Кроме того, для демонстрации эффекта от их применения был оставлен вариант представления нормализованных портретов с уровнями яркости исходных изображений (табл. 3.6). Таким образом вариантов обработки изображений $F=\{f_0, \dots, f_j\}$ было 8. Наглядное представление об использовавшихся вариантах нормализации даёт рис. 3.18.

Эксперимент производился на 17 парах изображений различных людей. Одно изображение из каждой пары было получено сканированием паспорта (рис. 3.18,а), второе – видеосъёмкой живого предъявителя указанного документа (рис. 3.18,ё). В процессе эксперимента на каждом из этапов предполагалось, что каждый из 17 человек пытается пройти процесс верификации как со своим документом, так и с оставшимися 16 паспортами. Таким образом, этап состоял из $C = 17^2 = 289$ сравнений паспортных изображений с «живыми». Так как общее число вариантов нормализации было $V = 10 \times 8 = 80$, то и этапов было столько же. Сравнение только «живых» либо только паспортных фотопортретов между собой в данном эксперименте не производилось. Максимальное число итераций M равнялось 1000.

Основные результаты эксперимента приведены на рис. 3.19 – 3.21. Для выдачи итогов верификации в удобной для человека форме производилось деление величины U по порогу T_{pc} на два класса – «один человек» и «разные люди». Кроме того, использовались два дополнительных порога – T_{min} и T_{max} – для обеспечения возможности принятия решений «точно один человек» и «точно разные люди». Разница между величинами U и T_{pc} оценивалась в процентах с учётом знака и выдавалась на экран как вероятность результата верификации. На протяжении всего эксперимента значения порогов были постоянны: $T_{min}=0.32$, $T_{pc}=0.4$, $T_{max}=0.5$.

На рис.3.19 приведена диаграмма зависимости среднего времени затраченного на полную обработку и сравнение пары изображений. Иными

Таблица 3.5

Расшифровка параметров типоразмеров нормализованных портретов

Номер типо-разме-ра	Параметры*			Номер типо-разме-ра	Параметры*		
	ED	$M \times n$	(EL _x , EL _y)		ED	$m \times n$	(EL _x , EL _y)
U ₁	30	60×45	(15,15)	U ₆	40	80×60	(20,20)
U ₂	30	60×35	(15,15)	U ₇	40	80×50	(20,20)
U ₃	30	30×60	(15,15)	U ₈	40	40×80	(20,20)
U ₄	30	60×30	(15,15)	U ₉	40	60×40	(10,20)
U ₅	30	60×25	(15,15)	U ₁₀	40	80×30	(20,15)

* Прим: ED - расстояние между зрачками, $m \times n$ – размер изображения, (EL_x, EL_y) – координаты центра левого зрачка. Все величины измерены в пикселях.

Таблица 3.6

Расшифровка параметров операторов предварительной обработки

Номер варианта	Оператор	Параметры*	Номер варианта	Оператор	Параметры*
f ₀	Без обработки	–	f ₄	“High Boosting”	$f=9.5$
f ₁	Дериша	$\alpha = 1.0$	f ₅	“High Boosting”	$f=9.8$
f ₂	Дериша	$\alpha = 2.0$	f ₆	“High Boosting”	$f=10.2$
f ₃	Дериша	$\alpha = 3.0$	f ₇	“High Boosting”	$f=10.5$

* Прим: Для каждого из операторов указаны абсолютные значения соответствующих ему параметров для “High Boosting” – f (см. формулу 2.16), для оператора Дериша – α [89].

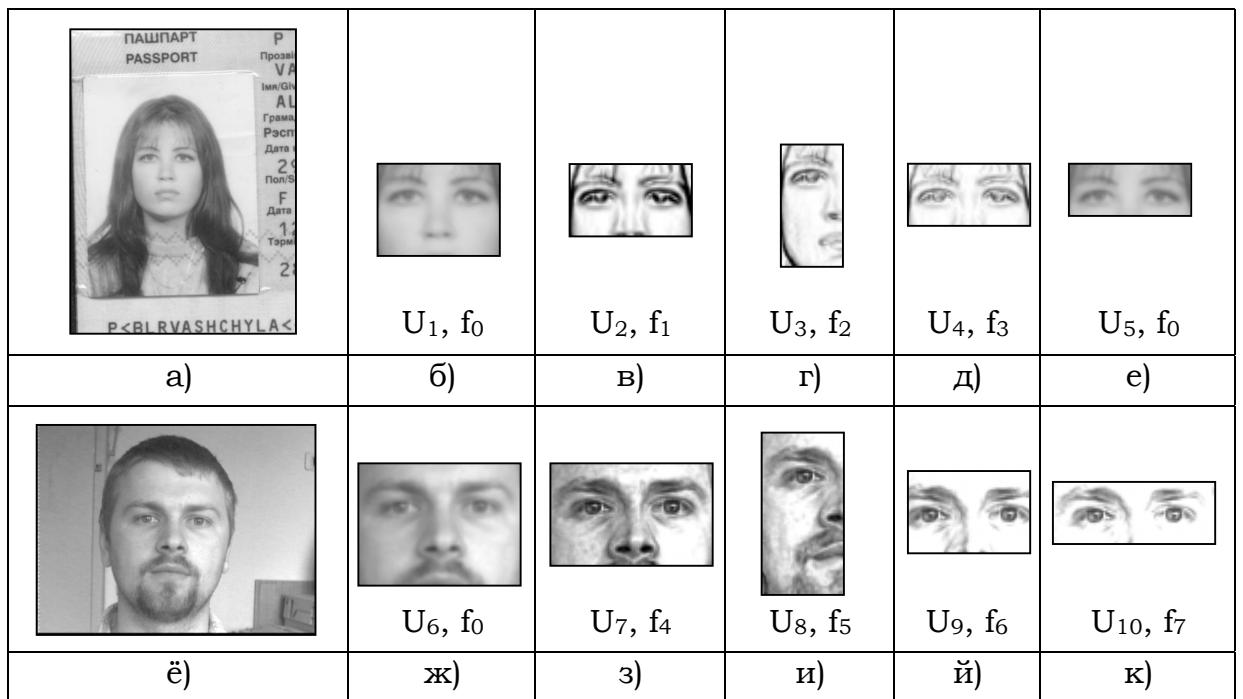


Рис. 3.18. Пример вариантов (типоразмеров) нормализованных частей лица $\mathbf{U}=\{U_1, \dots, U_{10}\}$ и параметров операторов предварительной обработки $\mathbf{F}=\{f_1, \dots, f_7\}$, использовавшихся при тестировании алгоритма сравнения на базе эластичных экспоненциальных деформаций. а), ё) исходные изображения. Расшифровка параметров для б)-е) и ж)-к) приведена в табл. 3.5 и 3.6

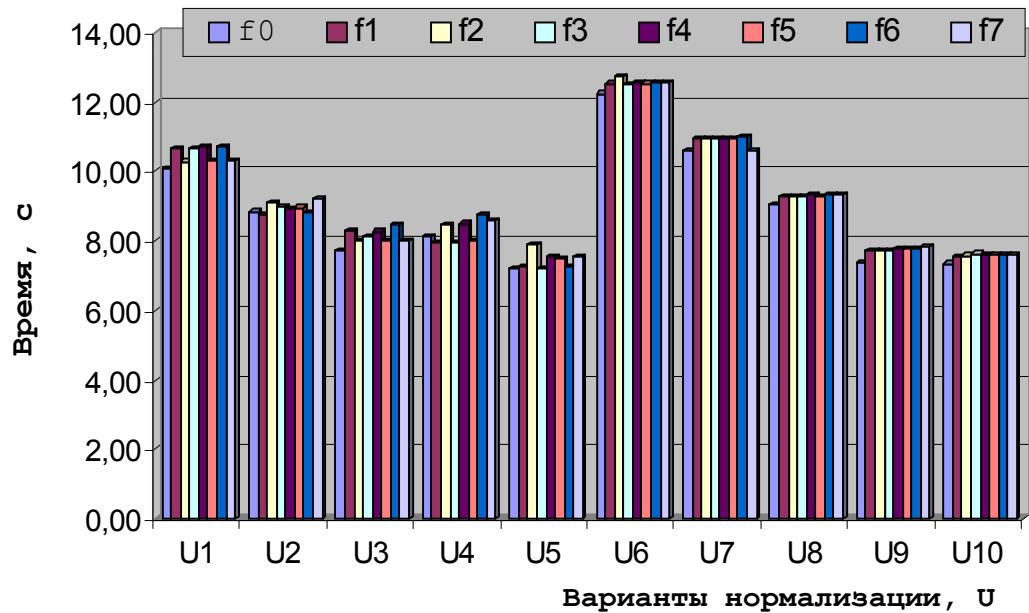


Рис. 3.19. Среднее время работы комплекса алгоритмов, затрачиваемое на предварительную обработку, нормализацию и сравнение пары изображений, в зависимости от \mathbf{U} и \mathbf{F} . Данные приведены для ПЭВМ типа Pentium II-600

словами, в данном эксперименте оценивалось время работы всех выносимых на защиту алгоритмов в комплексе: от подачи на вход двух изображений вида рис. 3.18,а,ё до выдачи на дисплей результата верификации (рис. 3.16, 4.4). На указанной диаграмме видна явная зависимость времени работы системы от размера нормализованных изображений U_1-U_{10} , а также тот факт, что основное время работы занимает алгоритм сравнения портретов методом эластичных экспоненциальных деформаций. Интерес вызывает тот факт, что при U_6-U_{10} время верификации практически не зависит от варианта функции предобработки f_j , за исключением f_0 , а при U_1-U_5 столь согласованных результатов не наблюдается. Это можно объяснить большим количеством досрочного завершения работы алгоритма сравнения. При f_0 время в среднем на 0.1 – 0.3 с меньше, но скорее это также вызвано более частыми случаями получения результата за меньшее количество итераций, нежели пропуском одного шага предварительной обработки.

Как уже отмечалось, целью эксперимента было получение объективной информации о наилучшей форме представления входных изображений для алгоритма сравнения. Для этого, чтобы результаты эксперимента не зависели ни от каких порогов, был введён усреднённый коэффициент R , который вычислялся следующим образом:

$$R = \begin{cases} \frac{1}{C} \sum_{i=1}^C (1 - U_i), & \text{если сравниваются портреты одного человека;} \\ \frac{1}{C} \sum_{i=1}^C U_i, & \text{если сравниваются портреты разных людей.} \end{cases} \quad (3.20)$$

Таким образом, чем выше R для определённой комбинации параметров **U** и **F**, тем лучше. Значения R в различных опытах колебались от 1.22 до 1.45, поэтому для наглядности на рис. 3.20 приведены значения коэффициента $R'=R-1.2$. Из результатов, указанных на диаграмме, следует, что наилучшими способами предварительной обработки являются f_2, f_3 – оператор Дериша с $\alpha=2$ и $\alpha=3$ соответственно. Более информативными частями исходных фотопортретов в данном случае следует признать U_3 и U_8 – вертикально направленные фрагменты области лица, различающиеся масштабом. Самым неудачным типоразмером можно назвать U_9 : при всех f_j $R'(U_9)$ имеет самые низкие значения. Это говорит о том, что наиболее информативные черты лица, отличающие людей друг от друга, остались «за кадром» либо потерялись на фоне большинства пикселей, не несущих важной для распознавания информации.

Данные выводы подтверждает и рис. 3.21, на котором изображена зависимость коэффициента распознавания K , выраженная в процентах. Он

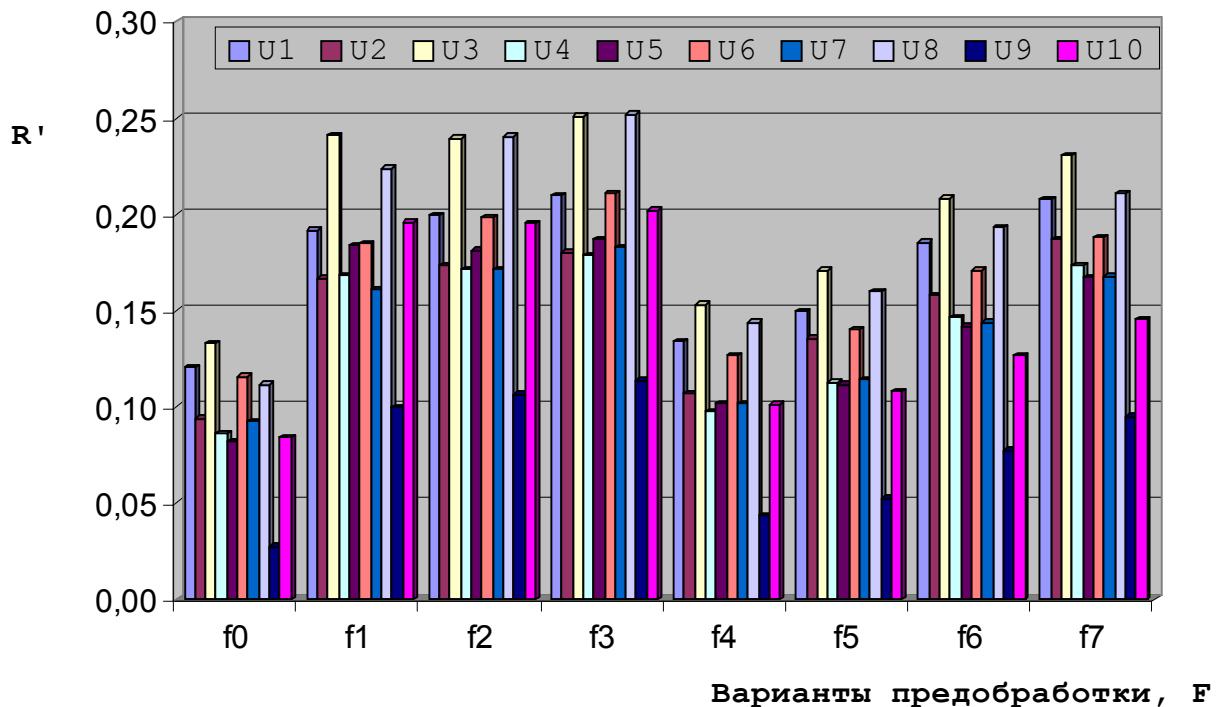


Рис. 3.20. Диаграмма зависимости коэффициента R' от параметров нормализации исходных изображений – частей лица $\mathcal{U}=\{U_1, \dots, U_8\}$ и операторов предварительной обработки $\mathcal{F}=\{f_1, \dots, f_7\}$

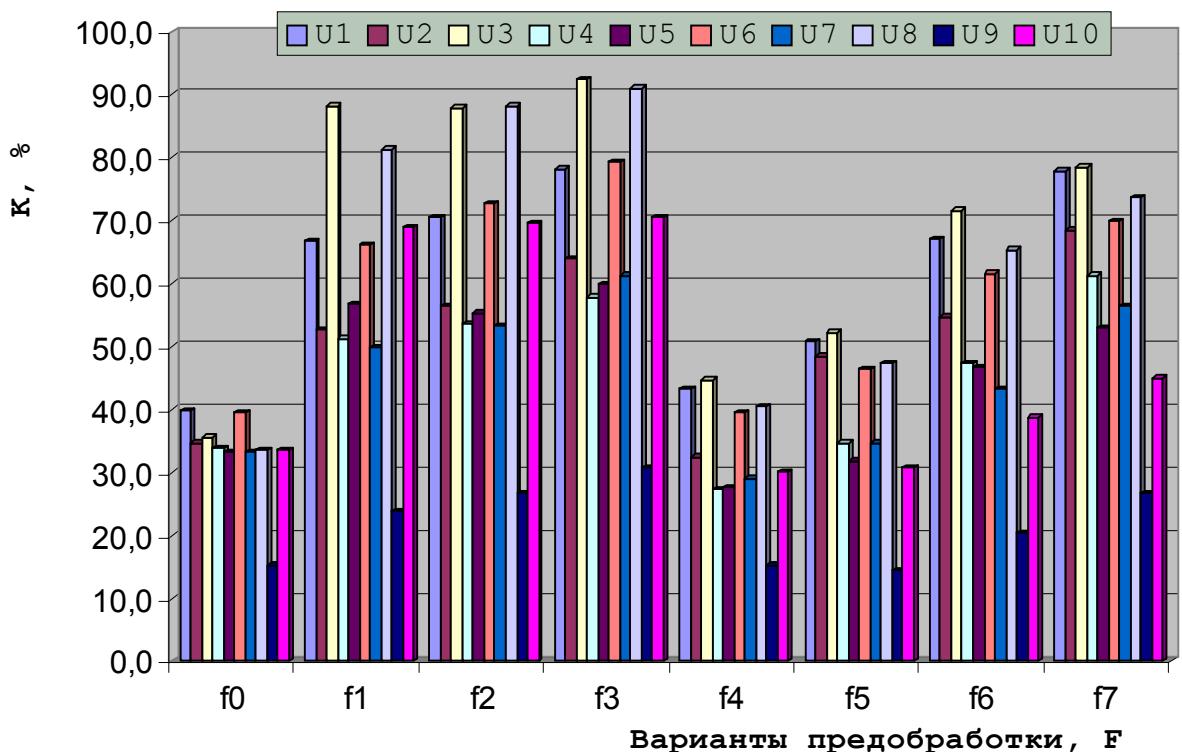


Рис. 3.21. Диаграмма зависимости коэффициента распознавания K от параметров нормализации исходных изображений – частей лица $\mathcal{U}=\{U_1, \dots, U_8\}$ и операторов предварительной обработки $\mathcal{F}=\{f_1, \dots, f_7\}$

рассчитывался по формуле $K=100 \times (C-ErI-ErII)/C$, где ErI,ErII – количество ошибок первого и второго рода соответственно. К сожалению следует отметить, что ErII превалировало над ErI. Очень низкие результаты при отдельных f_j и U_i объясняются фиксированными порогами T_{min} , T_{pc} , T_{max} , значения которых были подобраны эмпирически лишь для одного из указанных вариантов предобработки. В то же время абсолютные значения ψ существенно изменяются для различных f_j и U_i .

В целом эксперимент демонстрирует возможность получения коэффициента распознавания $K>90\%$ с помощью описанного подхода. В данном случае максимальное значение K составило 92,5%.

Выводы: С помощью предложенных ограничений, налагаемых на эластичную модель решётки изображений, удалось улучшить временные характеристики работы алгоритма в 5-7 раз, а также снизить роль фактора случайности порядка выполнения деформаций в итоговой оценке, одновременно повысив её стабильность.

Уменьшение времени сравнения нормализованных фотопортретов было достигнуто также и за счёт учета факта симметрии лица человека и уменьшения в два раза области деформации.

Экспериментальные исследования, проведённые в соответствии с указанным выше модифицированным методом эластичных экспоненциальных деформаций, показали его зависимость от разницы яркостных характеристик сравниваемых изображений. Несмотря на применение ранговой корреляции в качестве функции оценки близости изображений, указанная зависимость вносила существенные корректиры в результат сравнения.

3.4 Выводы

В третьей главе описаны новые алгоритмы решения задач поиска фотопортрета в базе данных и контроля удостоверений личности. Произведена экспериментальная проверка их эффективности. На основании приведённых в главе данных можно сформулировать следующие результаты:

1. Экспериментальные исследования показывают, что поиск фотопортретов в базе данных должен выполняться по принципу «от грубого к точному». Предложенная двухуровневая технология поиска фотопортретов, базирующаяся на инвариантных к условиям съёмки геометрических признаках и геометрической мере сходства изображений лица, была проверена на экспериментальной базе данных из 70 фотопортретов. Коэффициент распознавания составил 98,5%. (п.3.1).

2. Применение статистической оценки $F_l = D_{ld}/D_{lm}$ позволяет сократить признаковое пространство и отобрать наиболее информативные признаки. При удалении из признакового набора двух из тридцати признаков с наименьшей значимостью по F_l коэффициент распознавания увеличился с 90 до 98,5 % (п.3.2).

3. Алгоритм сравнения фотопортретов, разработанный на основе метода эластичных деформаций, позволил существенно уменьшить количество итераций, требуемых для получения достоверных оценок близости сравниваемых изображений, по сравнению с алгоритмом, описанным в [13] (с 20-30 тысяч до 1 тысячи). Коэффициент распознавания составил 92,5% при сравнении 289 пар реальных изображений людей. При этом минимальная разница во времени получения изображений, относящихся к одному классу, составляла более полутора лет.

Экспериментальные исследования показали зависимость результата работы алгоритма от разницы яркостных характеристик сравниваемых изображений, т.е. качества предварительной обработки входных изображений. Несмотря на применение ранговой корреляции в качестве функции оценки близости изображений, указанная зависимость вносила существенные корректиры в результат сравнения. К недостаткам алгоритма также можно отнести время его работы - более 10 с на сравнение одной пары изображений размером 100×100 пикселей на PC Pentium II – 600 (п.3.4).

Результаты описанных в данной главе исследований опубликованы в работах [26, 27, 28, 30, 143, 144, 146, 147, 148, 149] и отчётах о НИР [20, 22, 31, 34].

ГЛАВА 4 ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ КОМПЛЕКСА АЛГОРИТМОВ ИДЕНТИФИКАЦИИ ЧЕЛОВЕКА ПО ФОТОПОРТРЕТУ

В процессе поиска возможных способов решения задач, связанных с автоматической идентификацией человека по фотопортрету, выносимые на защиту алгоритмы и методики были лично реализованы автором в виде законченных программных модулей. Описываемые ниже программы были предназначены в первую очередь для экспериментальной проверки эффективности предложенных алгоритмов, отладки и демонстрации их результатов. Кроме того, разработанный комплекс алгоритмов был реализован в виде отдельного модуля динамически подключаемой библиотеки, готового к коммерческому использованию в составе пакетов программ заинтересованных организаций.

4.1 Экспериментальный комплекс программ

Для экспериментальной проверки предлагаемых алгоритмов были разработаны следующие программы:

- идентификации человека по геометрическим признакам лица, вычисленным на основе полученных/выделенных вручную координат антропометрических точек лица;
- обнаружения области лица на фотопортрете и получения координат центров зрачков;
- предварительной обработки и нормализации фотопортретов;
- поиска антропометрических точек лица с помощью фильтров Габора;
- сравнения фотопортретов на основе метода эластичных экспоненциальных деформаций.

Все программы, за исключением первой, были реализованы на языке C++ в среде программирования Microsoft Visual Studio™ v. 6.0 и предназначены для работы под управлением операционной системы (ОС) Microsoft Windows™ 95-98. Программа, реализующая алгоритм идентификации человека по геометрическим признакам лица, была написана для системы MS-DOS v.6.1 в среде Borland C++ v.3.1.

Общий вид программы поиска координат области лица и центров зрачков на фотопортрете приведён на рис. 4.1. На нём показана обработанная в соответствии с рассмотренными в главе 2 алгоритмами центральная область лица исходного изображения A''_F (справа) с интегральной

проекцией и найденным месторасположением глаз. Слева приводится исходное изображение с нанесёнными на него результатами поиска.

Экспериментальная отладка на большом множестве фотографий алгоритмов поиска областей лица и зрачков требовала возможности настройки параметров непосредственно в процессе эксперимента. Соответствующее меню программы изображено на рис. 4.2.

Обнаружение областей лица на фотографиях рассматривается как отдельная достаточно сложная задача. Так как в рамках данной работы был решён более простой случай этой задачи – поиск области лица на фотопортрете, т.е. лица только одного человека, то предусматривалось ручное ограничение области поиска в случае подачи на вход группового снимка (группа параметров Offsets for cutting, рис. 4.2).

Для нахождения оптимального варианта преобразования и последующей бинаризации изображения были перепробованы различные варианты операторов нахождения краёв и улучшения характеристик изображения (группы: Preprocessing, Thresholds, Service, Binarization, рис. 4.2). В процессе разработки алгоритма были исследованы варианты применения операторов симметрии, дистанционных преобразований и некоторых иных локальных признаков.

Подбор оптимального количества уровней пирамиды $Mz=\{A''_1, A''_2, \dots, A''_t\}$ и эталона области лица C также производился эмпирически.

Для поиска заданного фрагмента интегральной проекции (типа «локальный экстремум на интервале l после впадины глубиной не менее r и последующим пиком выше порога t ») производился анализ полученных функций с учётом задаваемых переменных величин в группах Integral Projection Draw Parameters и Analysis of Horizontal Integral Projection. Преобразование Хоха регулировалось одним единственным параметром Threshold for Loc Max of Possible Centers (рис. 4.2).

Таким образом, в процессе работы над решением поставленных задач автором были исследованы на множестве, состоящем из более двухсот фотографий, около сотни вариантов алгоритмов с различными параметрами.

В разделе 3.3 описан алгоритм верификации личности с помощью метода эластичных экспоненциальных деформаций. Было указано, что он является многоитерационным по своей сути, к тому же требующим значительных вычислений при каждой итерации. Если дополнительно отвлекать часть системных ресурсов компьютера на визуализацию результатов каждой итерации, то время сравнения изображений существенно возрастает. Поэтому для демонстрационных целей, а также для визуальной проверки



Рис. 4.1. Пример результата работы экспериментального ПО “Pas” для поиска области лица и координат центров зрачков на фотопортрете

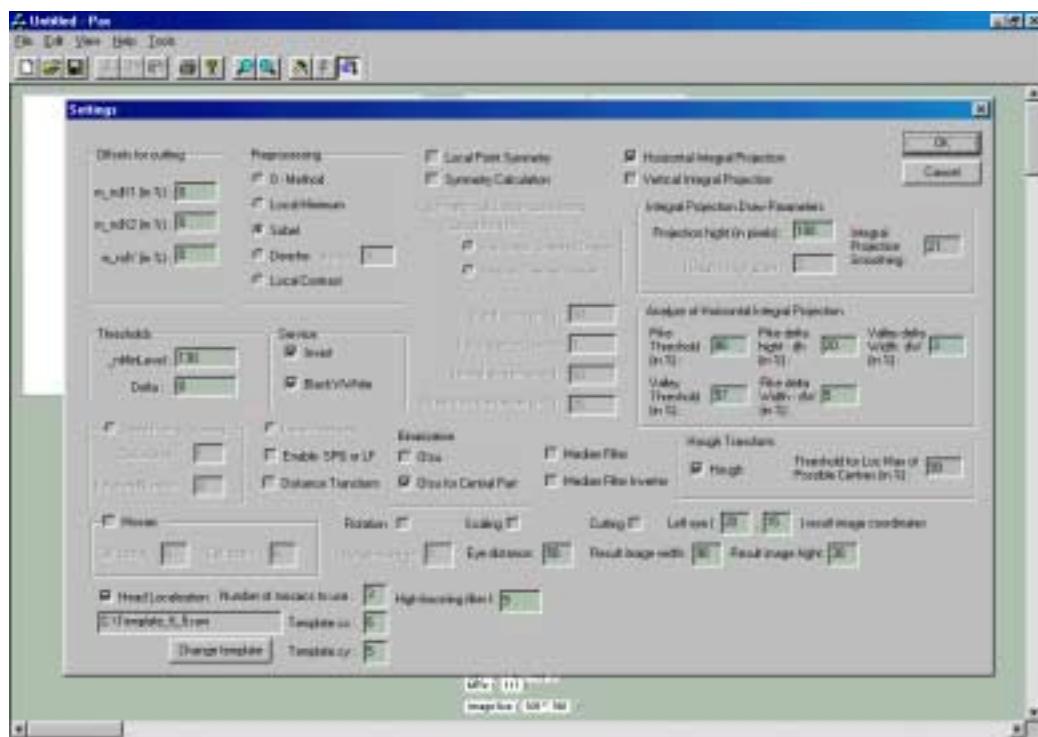


Рис. 4.2. Меню изменения параметров обработки изображения программы “Pas”

правильности работы программы были подготовлены несколько специальных вариантов программы сравнения нормализованных фотопортретов с помощью эластичных экспоненциальных деформаций. Разница между ними описана в п. 3.3.1, а их внешний вид приведён на рис. 3.15, 3.16. В версиях же, использовавшихся в экспериментальных целях, данный метод скрыт от пользователя и выдаёт лишь результаты своей работы.

4.2 Программная реализация комплекса алгоритмов в виде динамически подключаемой библиотеки

Комплекс разработанных алгоритмов, решающий задачу верификации человека по цифровым фотопортретам, был программно реализован в виде стандартного для ОС Microsoft Windows™ модуля динамически подключаемой библиотеки (Dynamic Link Library - DLL) – “IECFace22D.dll”. Его общая структура приведена на рис. 4.3.

Условно IECFace22D.dll можно разделить на два подмодуля – нормализации и, непосредственно, распознавания. Кроме того, в нём имеются не относящиеся к названным подмодулям блоки, отвечающие за связь с системными функциями ввода-вывода: блок чтения изображений, блок ручной коррекции координат обнаруженных точек лица и блок сохранения и визуализации результатов.

Подмодуль предварительной обработки и нормализации состоит из следующих блоков:

- определения координат области лица на фотопортрете;
- определения координат центров зрачков;
- поворота и масштабирования исходного изображения;
- изменения яркостных характеристик изображения;
- определения координат антропометрических точек лица;
- вычисления геометрических признаков лица на основе полученных антропометрических точек;
- коррекции вычисленных геометрических признаков лица с учётом информации о ракурсе съёмки.

Поступившие на вход фотопортреты преобразовываются из поддерживаемых программой форматов файлов (“BMP”, “TIF”) в блоке чтения изображения во внутреннее представление изображения – класс Cimage. Последующая обработка – от определения области лица до блока изменения яркостных характеристик – производится поочерёдно для исходных портретов. На вход блока определения антропометрических точек лица могут поступать как исходное, так и уже нормализованное изображения – в зависимости от варианта используемого алгоритма. В случае необходимости

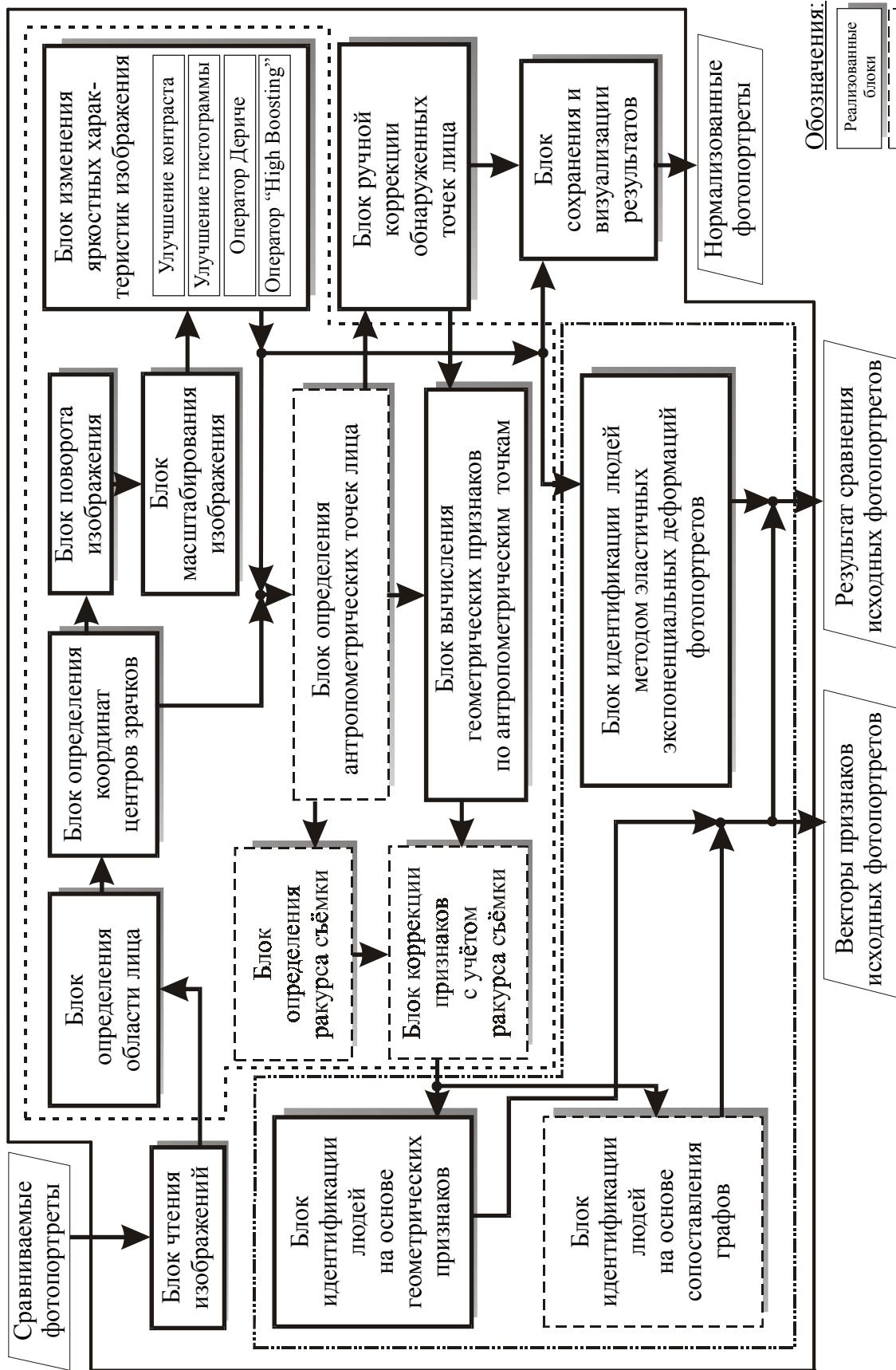


Рис. 4.3. Структурная схема комплекса алгоритмов, реализованного в виде модуля динамически подключаемой библиотеки – **TECHFace2D.dll**

оператор может подкорректировать положение точек с помощью функций соответствующего блока. Пример его работы приведён на рис. 4.4. На основе полученных координат точек и обобщённой модели головы можно приблизительно оценить её положение в пространстве и, соответственно, скорректировать геометрические признаки лица. Указанная процедура может быть реализована в соответствующем блоке определения положения человека перед камерой (ракурса) в момент съёмки.

Подмодуль распознавания состоит из блоков, реализующих принципиально различающиеся алгоритмы:

- на базе метода эластичных экспоненциальных деформаций;
- на основе геометрических признаков лица.

В отличие от последнего алгоритм сравнения фотопортретов с помощью эластичных экспоненциальных деформаций не может выдавать признаковое описание сравниваемых лиц. Его результатом является только оценка близости представленных портретов. Векторы признаков блока распознавания на основе антропометрических точек лица подробно рассмотрены в п.п.3.1.2, 3.2. Алгоритм распознавания – учитывающий как геометрические, так и дополнительные, характеризующие соответствующие элементы лица признаки – путём сопоставления описывающих лица графов, может быть реализован в виде отдельного блока.

Связь модулей основной программы (или программы-клиента) и библиотеки (“IECFace22D.dll” – программы-сервера) производится при помощи следующих пяти функций:

- `extern "C" __declspec(dllexport) int Compare_Fnames (const char* _FName_A, const char* _FName_B);`
- `extern "C" __declspec(dllexport) int* GetFaceArea (LPBYTE a_Image, CSize m_size);`
- `extern "C" __declspec(dllexport) int* Compare_Arrays (const char* _FName, int Im_flag);`
- `extern "C" __declspec(dllexport) int Compare_Arrays (LPBYTE a_ImageA, CSize m_sizeA, LPBYTE a_ImageB, CSize m_sizeB);`
- `extern "C" __declspec(dllexport) int Init_params(int* lp_params);`

Взаимодействие программ начинается с получения программой-клиентом двух сравниваемых цифровых фотопортретов. Исходные изображения сохраняются на диске и их имена при помощи функции **Compare_FNames()** передаются в программу-сервер. Предусмотрено несколько сценариев работы программ. Если клиенту кроме результата сравнения портретов никакой другой информации не требуется, то указанная функция возвращает результат – вероятность подобия/различия людей на фотопортретах – либо в случае ошибки – код ошибки. Если программа-клиент предполагает визуализацию найденных областей лица и центров

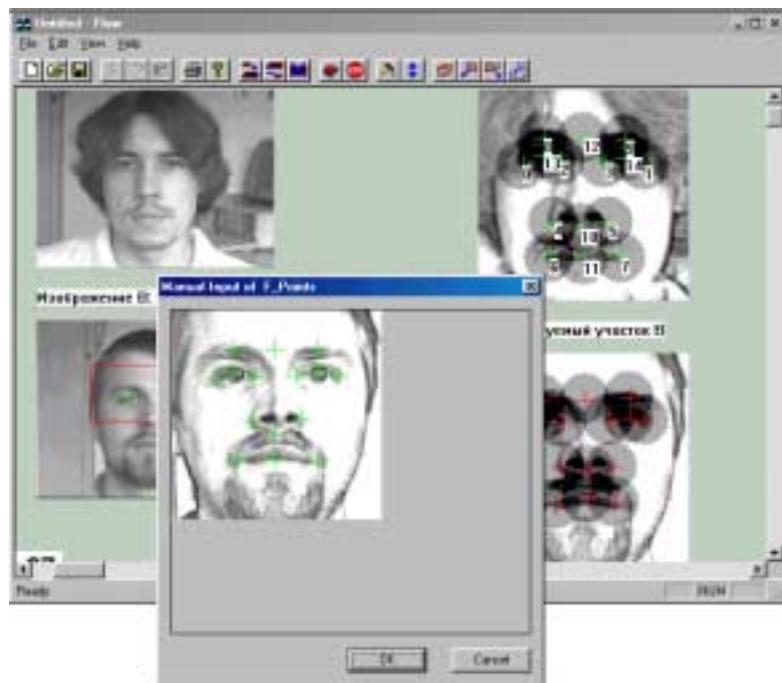


Рис.4.4. Пример результата работы блока ручной коррекции координат антропометрических точек лица. На рисунке показано основное и вспомогательные окна программы

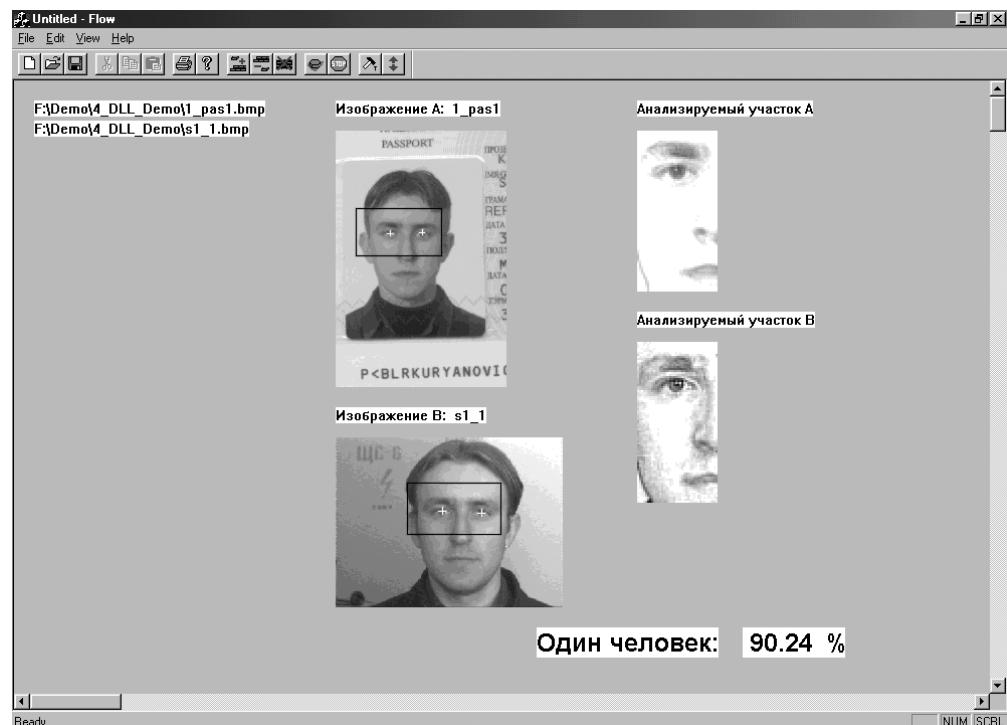


Рис.4.5. Пример результата работы комплекса алгоритмов для сравнения людей по фотопортретам. Слева – исходные изображения с найденными центрами зрачков. Справа – части нормализованных портретов, на основании сравнения которых было сделано заключение, что с правдоподобностью 90,24% на обоих портретах изображён один человек

зрачков, то она использует функцию **GetFaceArea()** и на выходе получает указатель на массив FaceCrd:

```
FaceCrd[0] = rectFace.left;
FaceCrd[1] = rectFace.top;
FaceCrd[2] = rectFace.right;
FaceCrd[3] = rectFace.bottom;
FaceCrd[4] = rectTemp.left;
FaceCrd[5] = rectTemp.top;
FaceCrd[6] = rectTemp.right;
FaceCrd[7] = rectTemp.bottom;
FaceCrd[8] = Radius_Eye,
```

где FaceCrd[0-3] – координаты соответственно верхнего левого и правого нижнего углов прямоугольника, ограничивающего область лица, а FaceCrd[4-5] и FaceCrd[6-7] – координаты центров левого и правого зрачков соответственно. Параметр FaceCrd[8] передаёт информацию о размере найденных зрачков на фотопортрете.

Для уменьшения времени, затрачиваемого на нормализацию портретов при проведении экспериментов, была предусмотрена возможность сохранения обработанных изображений на НЖМД (накопитель на жёстком магнитном диске). При сравнении нескольких сотен изображений по принципу «каждый с каждым» время выполнения эксперимента уменьшалось на порядок. Этот эффект достигался при помощи функций **GetNormFace()** и **Compare_Arrays()**, которые позволяют выполнять нормализацию и сравнение портретов раздельно, т.е. сперва проверялось наличие уже нормализованных сравниваемых портретов на диске и в положительном случае запускалось сравнение **Compare_Arrays()**. Только в том случае, если уже подготовленных портретов с указанными параметрами нормализации не существовало, вызовом **GetNormFace()** запускался соответствующий модуль.

Так как результат распознавания слабоконтрастных изображений сильно зависит от вариантов предобработки, то автором была написана также программа-клиент (рис. 4.5) и проведён ряд экспериментов по поиску оптимальных размеров нормализованных изображений, более инвариантных к изменениям частей лица, параметров предобработки и других, указанных в главах 2-3. На рис.4.6 изображено окно программы-клиента для изменения параметров алгоритмов программы-сервера. В частности, можно выделить четыре группы настроек: Operators – для указания параметров предобработки; Normalization – выбор размера и точки привязки нормализованных изображений; Gauss & Gabor Filters – для проведения экспериментов по автоматическому выделению точек лица; и безымянная группа параметров, отвечающих за работу метода эластичных экспоненциальных деформаций, – пороги принятия решения и указание количества итераций. Для передачи настроек предназначена функция **InitParams()**.

4.3 Экспериментальная проверка разработанных алгоритмов и сравнение с аналогами

Для сравнительного анализа разработанного комплекса алгоритмов было оценено время работы модуля “IECFace22D.dll”, затрачиваемое на выполнение трёх этапов: обнаружения области лица на фотопортрете, полной предварительной обработки и нормализации исходных изображений и всего процесса верификации от чтения портретов до выдачи результатов сравнения оператору.

Для исследования эффективности алгоритма поиска области лица на фотопортрете были проведены два эксперимента. В обоих из них определялась зависимость времени работы алгоритма от размеров исходных изображений. Соответственно были подготовлены два упорядоченных по возрастанию размеров фотопортретов множества. Минимальный размер фотопортретов, использовавшихся в экспериментах, был 314×235 пикселей, максимальный – 1047×783 . Разница между портретами заключалась в масштабе изображений искомых областей лица. В первом случае соотношение площади области лица – S_1 к общей площади снимка S_2 было фиксированным – $S_1/S_2 = 8,4\%$, во втором – переменным. При подготовке тестовых фотопортретов меньшие по размеру получались простым усечением наибольшего снимка без применения масштабирования $S_1/S_2 = k\%$. Рис. 4.7 демонстрирует описанные зависимости. Можно отметить близкий к линейному рост функции и практический отсутствие разницы во времени обработки изображений в обоих экспериментах. Для нахождения области лица на изображении минимального размера на ПЭВМ класса Pentium II-600 потребовалось менее 0,25 с, а для максимального – 2,5 с. Для сравнения можно указать, что поиск области лица на изображении 400×300 пикселей с помощью программы *FaceItTM* фирмы *Visionics* [134] занимает от 0,3 до 1 с.

В главе 1 было отмечено, что существует много работ по данной тематике, однако большинство из них посвящены решению задачи обнаружения лиц на групповых изображениях. Эксперименты проводились авторами на ЭВМ различных типов и классов и, что наиболее существенно, – на различных множествах тестовых изображений. Названные причины затрудняют объективную оценку эффективности того или иного алгоритма применительно к случаю верификации по фотопортретам. Однако для сравнения можно привести следующие данные.

Так, в работе [92] говорится о 86% достоверности обнаружения лиц при общем времени обработки одного изображения менее 5 с на ЭВМ “SUN Sparc 20 workstation”.

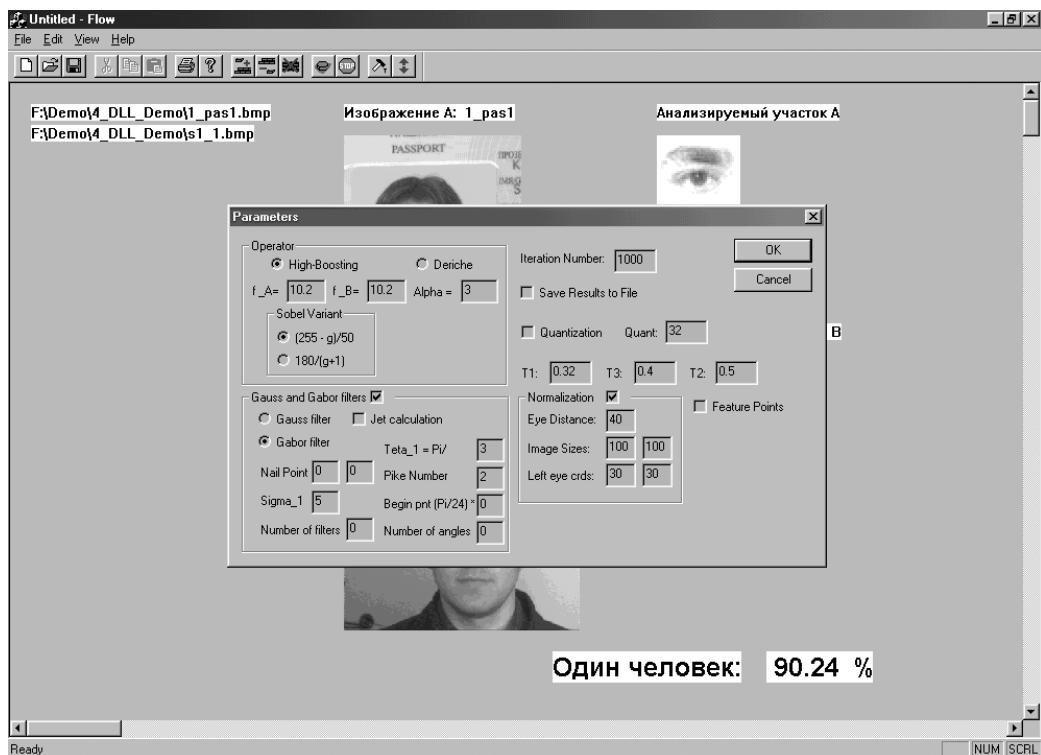


Рис. 4.6. Меню изменения параметров сравнения изображений программы "Flow" – клиента модуля **IECFace22D.dll**

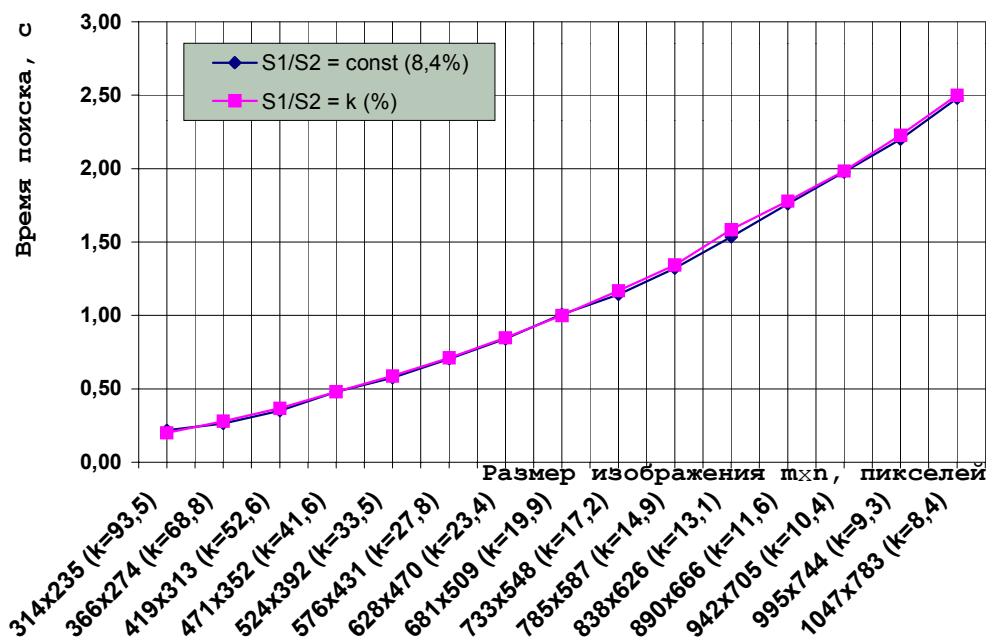


Рис. 4.7. График зависимости времени работы алгоритма поиска лица на фотопортрете от размеров исходного изображения. В первом случае отношение площади лица к площади всего изображения $S1/S2$ было постоянным (8,4%). Во втором было переменным: $S1/S2 = k\%$, k указано для каждого изображения в скобках, после размеров

Наиболее основательным исследованием в данной области можно назвать работу [166]. В двух экспериментах Вебер и др. испытывали свой подход на базе, состоящей из более чем 5,5 тыс. и 32 тыс. изображений, и достигли 96 – 99%-ной точности обнаружения. Время обработки одного изображения, но без построения пирамиды, составляло около 0,3 с на ЭВМ Pentium II-400.

В работах [77, 78] описана система, построенная на нейронных сетях. Точность её распознавания составляет 84 – 86,5% в зависимости от качества исходных данных. Время обработки одной групповой фотографии с произвольным фоном от 1,3 до 2,9 с на DEC Alpha-333 MHz.

В главах 2 и 3 было показано, что процесс предобработки изображений предусматривает возможность получения на выходе различных вариантов нормализованных изображений, которые могут различаться, в том числе и размерами. Для оценки времени работы подмодуля предобработки и нормализации на описанных выше множествах фотопортретов были проведены несколько экспериментов, результаты которых отражены на рис. 4.8 и 4.9. На них изображены графики зависимости времени нормализации от размера исходных изображений для итоговых изображений размерами 100×100 и 70×40 пикселей при фиксированном и переменном соотношении S1/S2 соответственно. Для оценки затрат на поиск центров зрачков, масштабирование и прочие операции на указанных рисунках приведены также графики времени, затрачиваемого только на поиск области лица. Отсутствие плавности роста рассматриваемых функций объясняется дополнительными временными затратами алгоритма поиска центров зрачков, если количество областей поиска больше одной (приблизительно в 25% случаев). Из рисунков видно, что при обработке изображений с размерами, большими, чем 400×400 пикселей, временные затраты превышают 2 с.

Сравнение алгоритма с аналогами также затруднено по причинам указанным выше. Можно отметить, что программа “*ВидеОко*” [17] находит центры зрачков за 0.5 – 1 с. Для одного из лидеров рынка *FaceIt™* время полной обработки лица неизвестно, однако сообщается, что время формирования вектора признаков – 1 с при уже обнаруженной области лица и весьма приблизительном поиске местоположения глаз [134].

Эксперименты, описанные в п.3.3.2, демонстрируют среднее время отдельной процедуры верификации по паре изображений, равное около 8,5 с, что, несомненно, больше, чем у названных выше систем. Так, “*ВидеОко*” тратит в среднем на обработку изображения 256×256 и распознавание от 4 до 5 с (в случае трёх классов, состоящих из трёх портретов каждый) и до 8 с при увеличении количества сравнений вдвое. Фирма eTrue [135]

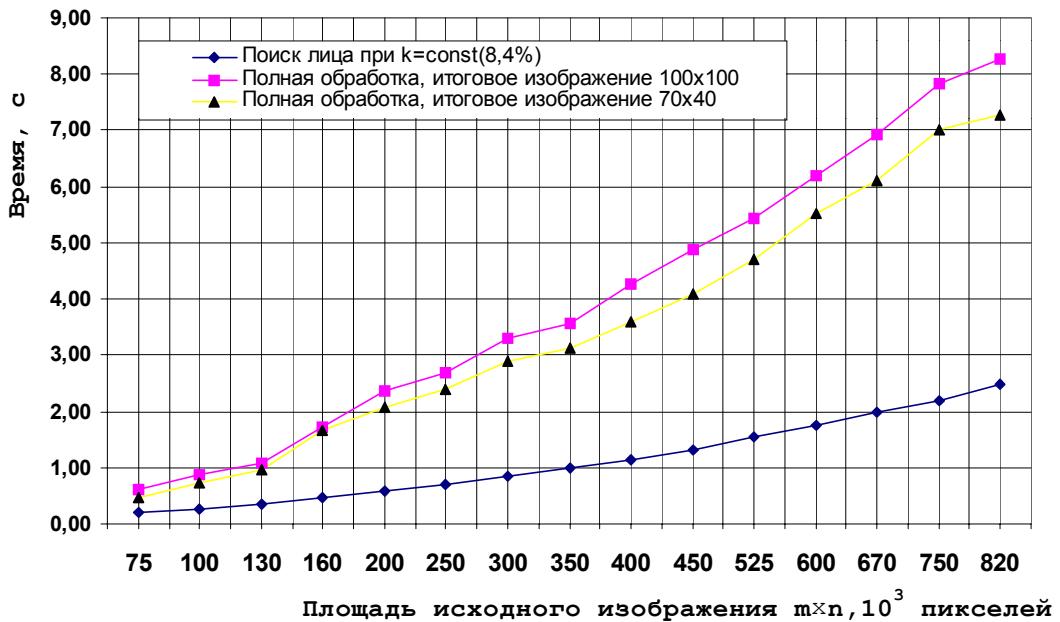


Рис. 4.8. График зависимости времени работы алгоритмов нормализации от размеров исходного и итогового изображений. Отношение площади лица к площади обрабатываемых изображений S_1/S_2 было постоянным (8,4%)

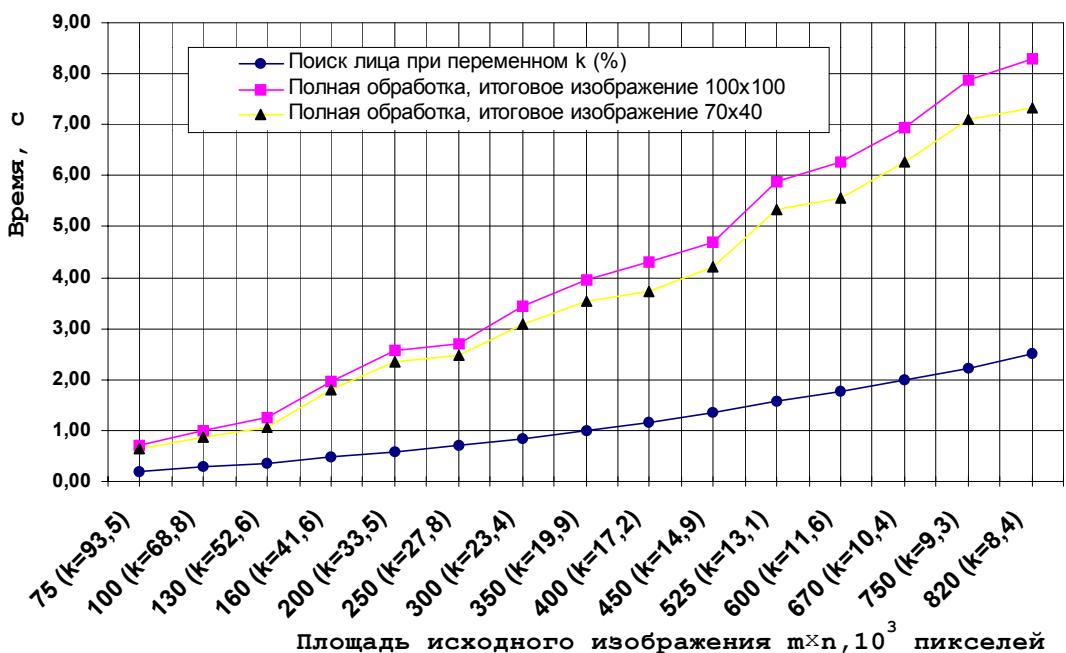


Рис. 4.9. График зависимости времени работы алгоритмов нормализации от размеров исходного и итогового изображений. Отношение площади лица к площади обрабатываемых изображений было переменным: $S_1/S_2 = k\%$, k указано для каждого изображения в скобках

утверждает, что её система TrueFaceID™ производит верификацию двух нормализованных изображений менее чем за 1 с. FaceCam™ТА – программно-аппаратный комплекс, комбинирующий несколько способов верификации (продукт компании Vision Sphere Technologies), обеспечивает пропускную способность от 8 до 10 человек в минуту, т.е. на одного человека уходит от 6 до 7,5 с. При этом производится нормализация лишь одного – поступившего с видеокамеры – изображения и его сравнение с уже имеющимися признаками лица идентифицируемого человека. Следует отметить, что алгоритм распознавания на основе геометрических признаков лица обеспечивает практически моментальное распознавание (менее 0.05 с – на сравнение пары векторов), однако в настоящий момент не все антропометрические точки лица могут быть выделены автоматически с требуемой точностью.

При всей очевидности несущественного проигрыша в скорости работы выносимого на защиту алгоритма верификации на основе эластичных экспоненциальных деформаций основным его преимуществом является устойчивость результатов при возрастных изменениях во внешности человека. В то время как по данным, опубликованным в отчёте по испытанию систем распознавания на базе изображений FERET [128], коэффициент распознавания тестировавшихся систем (в том числе и FaceIt™, TrueFaceID™) резко падал – до 60%, если на вход подавались изображения людей, сделанные с временной разницей в полгода, разработанная система демонстрировала коэффициент распознавания $K = 92,5\%$ на изображениях, сделанных с заранее большей разницей во времени (от года до двадцати).

В заключение следует заметить, что при очень агрессивной рекламной компании названных разработчиков информация о применении названных продуктов в системах паспортного контроля ни в широкую печать, ни в специализированные источники не поступала.

4.4 Использование результатов диссертационной работы

Следует заметить, что для разработки полноценной системы автоматического идентификации человека по фотопортрету, готовой к коммерческому использованию, требуются большие трудозатраты на системное и прикладное программирование. Несмотря на то, что задача построения готовой коммерческой системы выходит за рамки настоящей работы, комплекс алгоритмов, выносимых на защиту был реализован в виде стандартного модуля динамически подключаемой библиотеки (DLL), готового для использования в коммерческих целях в составе программных пакетов заинтересованных организаций. В частности, он был передан НПП «Регула» для

использования в пакете специализированного программного обеспечения пограничного контроля [66] – «Автоматизированное рабочее место пограничника» (АРМП) «Контролер». Данный пакет программ обеспечивает «автоматический ввод в компьютер данных паспорта, формирование при необходимости базы паспортных данных, а также управление системами сравнения фотографических и живых изображений владельца и поиска фотографий в базах данных» [47, 48].

Схема взаимодействия модулей приведена на рис. 4.10. Из рисунка видно, что за получение изображений с периферийных устройств ПЭВМ – видеокамеры и «Считывателя паспортов» (мод. 7001, производитель НПП «Регула») – отвечает программное обеспечение фирмы – (АРМП) «Контролер». После того как изображения сохранены в стандартном для ОС Microsoft Windows™ формате “BMP” на жёсткий диск, управление передаётся описанному в п.4.2. модулю IECFace22D.dll с указанием имён файлов сравниваемых фотопортретов. В нём поочерёдно производится полная предварительная обработка и нормализация поступивших изображений посредством описанных в главе 2 алгоритмов. Промежуточным выходом модуля после данного этапа являются координаты обнаруженных областей лица и центров зрачков на фотопортрете. Они передаются в вызывающую программу, которая выделяет соответствующие области на изображениях для возможности визуального контроля оператором. Нормализованные изображения могут быть сохранены в формате “BMP” на жёсткий диск для потенциального использования другими программами пакета.

Окончательным результатом работы модуля IECFace22D.dll является оценка вероятности принадлежности сравниваемых портретов одному человеку, которая и передаётся клиентской программе при возвращении управления. За отображение результата на экране и дальнейшую обработку полученной информации несёт ответственность ПО АРМП «Контролер».

Результаты диссертационной работы, описанные в [30, 144, 146, 149], а именно проведённый анализ особенностей построения автоматических систем идентификации человека по цифровым фотопортретам, алгоритмы обнаружения зрачков и сравнения фотопортретов с помощью эластичных деформаций, использовались при чтении курсов лекций в Белорусском государственном университете информатики и радиоэлектроники и в Белорусском государственном технологическом университете.

Акты внедрения и использования результатов работы приведены в приложении 2.

Программное обеспечение автоматизированного паспортного контроля

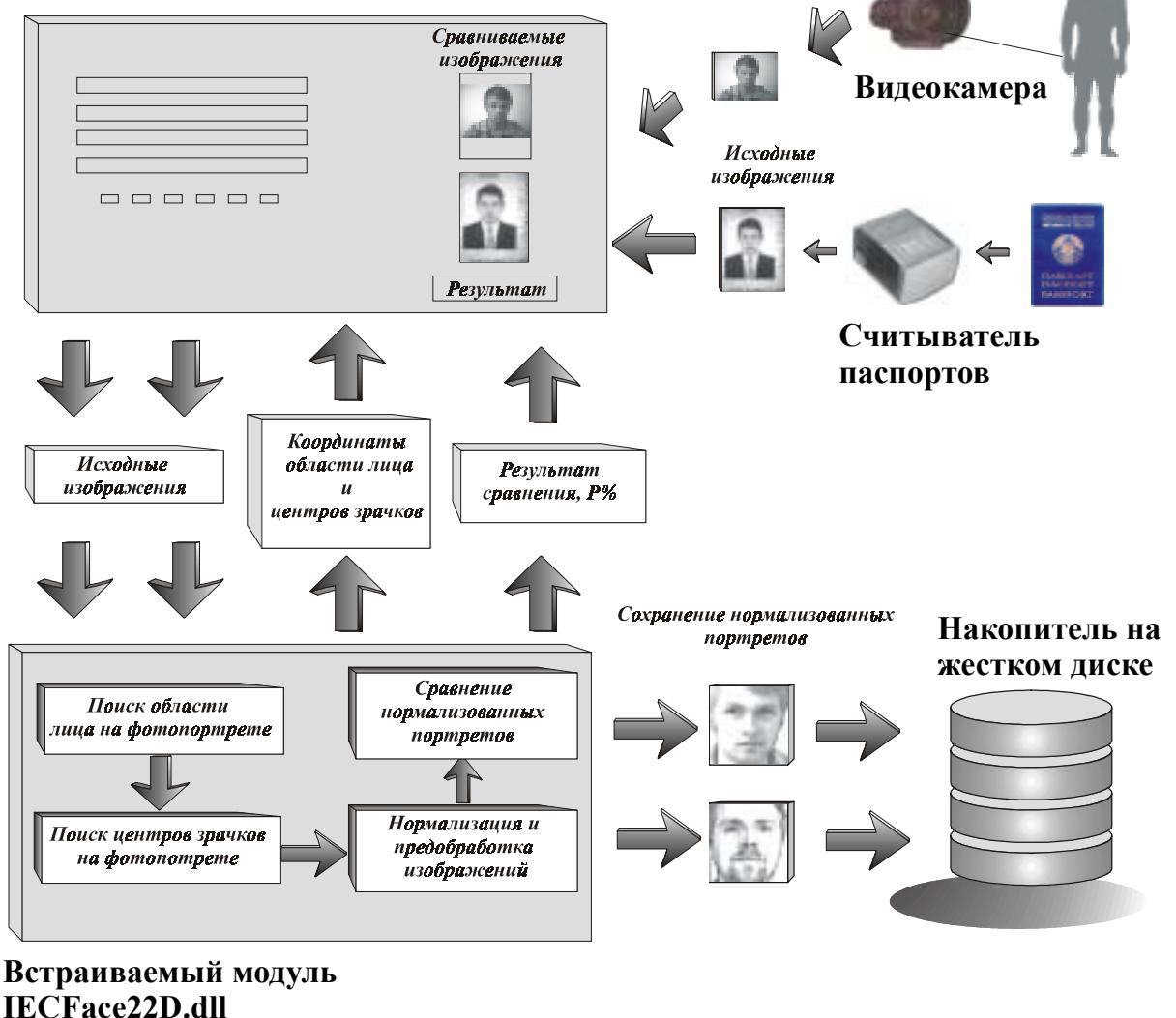


Рис.4.10. Схема взаимодействия ПО АРМП «Контролер» НПП «Регула» [40] с комплексом алгоритмов обработки и распознавания фотопортретов, оформленных в виде динамически подключаемой библиотеки (DDL)

4.5 Выводы

В четвёртой главе описана программная реализация разработанных алгоритмов в виде стандартного для ОС Microsoft Windows™ 95-98 модуля динамически подключаемой библиотеки. Проведено экспериментальное исследование эффективности комплекса разработанных алгоритмов на собственной базе данных (376 портретов 50 людей) и базе ORL (400 портретов 40 людей), выполнено сравнение его характеристик с аналогами.

На основании приведённых в главе данных можно сформулировать следующие выводы:

1. Алгоритм определения координат области лица на фотопортрете и алгоритм нахождения центров зрачков устойчиво работают на слабоконтрастных изображениях в режиме реального времени (общее время предобработки – от 0.5с). Быстродействие алгоритмов может быть увеличено за счёт более эффективной программной реализации в процессе доработки разработанного программного обеспечения до коммерческого уровня (п.4.3).

2. Алгоритм идентификации человека с помощью эластичных экспоненциальных деформаций может использоваться в режиме реального времени в задачах паспортного контроля. Основным его недостатком является принципиальная невозможность формирования вектора признаков портрета (т.е. повторного использования результатов работы алгоритма) и существенного снижения вычислительной стоимости из-за итерационной природы алгоритма. Однако в отличие от известных аналогов, алгоритм показал хорошие результаты при сравнении фотопортретов, сделанных с разницей во времени от года до 30 лет – коэффициент распознавания составил 92,5% (п.4.3).

Результаты исследований, описанных в данной главе, опубликованы в работах [148, 149] и отчёте о НИР [22].

ЗАКЛЮЧЕНИЕ

Основные результаты, полученные в диссертационной работе, можно сформулировать следующим образом:

1. Разработаны алгоритмы поиска области лица на фотопортретах и локализации центров зрачков глаз на найденной области лица. Эти алгоритмы, в отличие от ранее опубликованных в литературе, позволяют находить центры зрачков как на контрастных, так и на слабоконтрастных изображениях в режиме реального времени и не требуют предварительного обучения и настройки [148, 149, 150].

2. Исследованы различные алгоритмы предварительной обработки цифровых фотопортретов. Показано, какие алгоритмы поворота цифровых изображений лучше использовать в системах распознавания (минимизируя искажения исходной информации), а какие – для лучшего визуального восприятия человеком. Разработан алгоритм повышения контраста цифровых фотопортретов и их нормализации. Он отличается возможностью выравнивания яркостных характеристик как слабоконтрастных (полученных сканированием документа) фотопортретов, так и изображений человека, поступающих с видеокамеры [20, 33, 139, 145].

3. Исследована задача автоматического поиска основных антропометрических точек лица с помощью фильтра Габора. Проанализировано 5040 комбинаций параметров фильтра Габора для её решения. Впервые экспериментально определены параметры фильтра Габора и нормализации фотопортретов, позволяющие находить восемь основных антропометрических точек лица с точностью до 2 пикселей [27, 29, 150].

4. Предложены и экспериментально исследованы информативные признаки лица на основе комбинаций расстояний между определёнными антропометрическими точками лица. Их особенностями является робастность к возрастным изменениям внешности человека, освещённости, макияжу и незначительным вариациям ракурса съёмки, инвариантность к масштабу, яркости и контрасту. Использование этих признаков позволило получить точность распознавания 98,5% на экспериментальной базе данных из 70 фотопортретов, сделанных с разницей во времени в несколько лет при разных ракурсах съёмки и условиях освещённости. Предложена и исследована двухуровневая технология поиска фотопортретов людей, похожих на заданного человека, в базах данных, основанная на использовании геометрических признаков лица человека [22, 26, 28, 30, 31, 34, 143, 144, 147].

5. Разработан итерационный алгоритм идентификации человека по цифровым фотопортретам, основанный на методе эластичных деформаций изображения. Этот алгоритм, использующий не более тысячи итераций, позволил обеспечить коэффициент распознавания 92,5% на изображениях, полученных сканированием паспортов и других удостоверений личности, изготовленных в разные годы [148, 149].

Результаты, описанные в диссертации, опубликованы в 14 печатных работах и 6 отчётах о НИР. Использование совокупности новых результатов, полученных в рамках проведённых исследований, обеспечивает решение важной прикладной задачи идентификации человека по цифровым фотопортретам в режиме реального времени. Разработанные алгоритмы, реализованные в виде динамически подключаемой библиотеки программ, были внедрены в специализированную систему пограничного контроля АРМП «Контролер» (НПП «Регула», г. Минск), разработанную по заказу Государственного комитета пограничных войск Республики Беларусь.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Абламейко С.В., Лагуновский Д.М. Обработка изображений: технология, методы, применение. – Минск: Ин-т техн. кибернетики НАН Беларуси, 1999. – 300с.
2. Айвазян С.А., Енюков И.С., Мешалкин А.Д. Прикладная статистика. Статистическое оценивание зависимостей. – М.: ФиС, 1985. -484с.
3. Быстрые алгоритмы в цифровой обработке изображений/ под ред. Хуанга Т.С. – М.: РиС, 1984. –224с.
4. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов – М.: Наука, 1974. – 416с.
5. Вентцель Е.С. Теория вероятностей – М.: Высш. шк., 1999. -576с.
6. Верхаген К., Дёйн Р., Грун Ф. Распознавание образов: состояние и перспективы – М.: РиС, 1985. -104с.
7. Виниченко И.Ф., Житников В.С., Зинин А.М., Овсянникова М.Н., Снетков В.А. Криминалистическое описание внешности человека – М.: Щит-М, 1998. –198с.
8. Герасимов М.М. Основы восстановления лица по черепу. М.: Сов. наука, 1949. -187с.
9. Горелик А.Л., Скрипкин В.А. Методы распознавания – М.: Высшая школа, 1984. – 208с.
10. Дуда Р., Харт П. Распознавание образов и анализ сцен. – М.: Мир, 1976. – 511с.
11. Зинин А.М., Кирсанова Л.З. Криминалистическая фотопортретная экспертиза. – М.: МВД СССР ВНКЦ, 1991. - 88с.
12. Кендэл М. Ранговые корреляции. – М.: Статистика, 1975. -214с.
13. Ковалёв В.А. Метод эластичных экспоненциальных деформаций для совмещения цифровых изображений// Сб. науч. тр. «Цифровая обработка изображений».– Минск. – Ин-т техн. киберн. НАН Беларуси.– 1999.– С.147-156.
14. Куприянов В.В., Стовичек Г.В. Лицо человека: анатомия, мимика. М.: Медицина, 1988. - 268с.
15. Кухарев Г.А. Биометрические системы: Методы и средства идентификации личности человека. – СПб.: Политехника, 2001. – 240 с.
16. Мартинес Ф. Синтез изображений: принципы, аппаратное и программное обеспечение – М.: РиС, 1990. –192с.

17. Махалов С. Применение аппарата собственных векторов для фотопортретной экспертизы// Труды конференции Академии МВД «Автоматизация работ подразделений МВД». - Москва. – 1999. – С.192-197.
18. Патрик Э. Основы теории распознавания образов – М.: Сов. радио, 1980. – 408с.
19. Полевой Н.С. Криминалистическая кибернетика – М.: Изд-во МГУ, 1982. – 208с.
20. Предварительная обработка и анализ фотопортретов: Отчёт о НИР (заключительный)/ Ин-т техн. киберн. НАН Беларусь; Рук. Д.И. Самаль. - №ГР 19994336. – Минск., 2000. – 60с.
21. Прикладная статистика. Классификация и снижение размерности. Справ. изд. /Под. ред. Айвазяна С.А. – М.: ФиС, 1989. -507с.
22. Проблемно-ориентированное распознавание объектов полутоновых изображений: Отчёт о НИР (заключительный)/ Ин-т техн. киберн. НАН Беларусь; Рук. С.В. Абламейко. - №ГР 19991970. – Минск., 2001. – 124с.
23. Прэтт У. Цифровая обработка изображений/ в 2-х. кн. – М.: Мир, 1982. - 792с.
24. Пухначев Ю.В., Попов Ю.П. Математика без формул. – М.: Столетие, 1995. - 512с.
25. Разработка методов, инструментальных средств и интеллектуальных технологий формирования цифровых моделей местности на основе автоматизированного дешифрирования аэро- и космоснимков - Описание методов и алгоритмов автоматизированного дешифрирования снимков: Отчёт о НИР (эскизный проект)/ Ин-т техн. киберн. НАН Беларусь; Рук. С.В. Абламейко. - №ГР 2002145. – Минск., 2001. – 223с.
26. Самаль Д.И. Построение систем идентификации личности на основе антропометрических точек лица// Сб. науч. тр. «Цифровая обработка изображений». - Минск.- Ин-т техн. киберн. НАН Беларусь.- 1998.- С.72-78.
27. Самаль Д.И., Старовойтов В.В. Выбор признаков для распознавания на основе статистических данных// Сб. науч. тр. «Цифровая обработка изображений». - Минск. - Ин-т техн. киберн. НАН Беларусь. – 1999. - С.105-114.
28. Самаль Д.И., Старовойтов В.В. Методика автоматизированного распознавания людей по фотопортретам// Сб. науч. тр. «Цифровая обработка изображений». - Минск. - Ин-т техн. киберн. НАН Беларусь. – 1999. - С.81-85.
29. Самаль Д.И., Старовойтов В.В. Обнаружение антропометрических точек лица с помощью фильтров Габора// Сб. науч. тр. «Цифровая обра-

ботка изображений». - Минск. - Ин-т техн. киберн. НАН Беларуси. – 2001. - С.141-150.

30. Самаль Д.И., Старовойтов В.В. Подходы и методы распознавания людей по фотопортретам. – Минск, 1998. – 54с. – (Препринт/ НАН Беларуси Ин-т техн. киберн.; № 8).

31. Сегментация изображений с помощью ориентационных параметров и функционала энергии: Отчёт о НИР (промежуточный)/ Ин-т техн. киберн. НАН Беларуси; Рук. С.В. Абламейко. - №ГР 19991970. – Мн., 1999. – 50с.

32. Снетков В.А., Виниченко И.Ф., Житников В.С. и др. Криминалистическое описание внешности человека/ Под общ. ред. В.А. Снеткова – М.: МВД СССР ВНИИ, 1984. -127с.

33. Создание теоретических основ разработки, анализа и распознавания изображений с использованием методов искусственного интеллекта// Отчёт о НИР (промежуточный)/ Ин-т техн. киберн. НАН Беларуси; Рук. С.В. Абламейко. - №ГР 19962838. – Мн., 1999. – 28с.

34. Создание теоретических основ разработки, анализа и распознавания изображений с использованием методов искусственного интеллекта// Отчёт о НИР (заключительный)/ Ин-т техн. киберн. НАН Беларуси; Рук. С.В. Абламейко. - №ГР 19962838. – Мн., 2000. – 91с.

35. Средства контроля доступа// Иностранныя печать о техническом оснащении полиции капиталистических государств – М.: ВИНИТИ. - 1992. - № 4. - С.12-27.

36. Старовойтов В.В. Локальные геометрические методы цифровой обработки и анализа изображений. – Мн.: ИТК НАН Беларуси, 1997. – 284с.

37. Стёпин В.С., Савушкин А.В., Зотов А.Б. Криминалистическое отождествление человека по разноракурсным фотопортретам. Метод. рекомендации – М.: ЭКЦ МВД, 1992. – 24с.

38. Талеб М.А., Старовойтов В.В., О сегментации цветных изображений// Известия Национальной Академии Наук Беларуси. – Минск. – 2000. - № 1. - С.107-111.

39. Тиори Т., Фрай Дж. Проектирование структур баз данных/ В 2-х. кн. – М.: Мир, 1985. - 602с.

40. Файн В.С. Опознавание изображений. – М.: Наука, 1970. – 299с.

41. Ту Дж., Гонсалес Р. Принципы распознавания образов. – М.: Мир, 1978. – 411с.

42. Фомин Я.А., Тарловский Г.Р. Статистическая теория распознавания образов – М.: РиС, 1986. – 264с.

43. Фу К. Последовательные методы в распознавании образов и обучении машин – М.: Наука, 1971. – 256с.

44. Фуканага К. Введение в статистическую теорию распознавания образов – М.: Наука, 1979. – 368с.
45. Хансен Г., Хансен Дж. Базы данных: разработка и управление. – М.: Издательство БИНОМ, 1999. – 284с.
46. Хорн Б. К.П. Зрение роботов – М.: Мир, 1989. - 400с.
47. Шумский И.П. Автоматизированный анализ подлинности паспортов и денежных банкнот по их изображениям: Дис. канд. тех. наук: 05.13.01.- Минск. - 2001. - 145с.
48. Шумский И.П. Комплексная система получения и обработки изображений “VideoScope”// Сб. науч. тр. «Цифровая обработка изображений». - Минск. - Ин-т техн. киберн. НАН Беларуси. – 2000. - С.88-94.
49. Abay E., Akarum L. and Alpaydyn E. A comparative analysis of different feature sets for face recognition// Proceedings of Int. Symp. On Computer and Information Sciences. – Antalya. - 1997. - P. 220-227.
50. Achermann B. and H. Bunke Combination of face classifiers for person identification// Proceedings of Int. Conf. on Pattern Recognition. - 1996. -Vol.4. - P. 416-420.
51. Argyle E. Techniques for edge detection// Proceedings of the IEEE. - 1971. -Vol.59. - P. 285-287.
52. Atherton T.J. and Kerbyson D.J. Size invariant circle detection// Image and Vision Computing. - 1999.-Vol.17.- P.795-803.
53. Ben_Arie J. and Nandy D. A volumetric/iconic frequency domain representation for objects with application for pose invariant face recognition// IEEE Trans. on Pattern Analysis and Machine Intel.-1998.-Vol.20.- No.5.- P.449-457.
54. Bigun J., Choy K. and Olsson H. Evidence on skill differences of women and men concerning face recognition// Lecture Notes in Computer Science.- Vol.2091.- Springer.- 2001.- P. 44-51.
55. Bigün J., Hans du Buf J.M. N-folded symmetries by complex moments in Gabor space and their application to unsupervised texture segmentation// IEEE Trans. on Pattern Analysis and Machine Intel.- 1994.-Vol.16.- No.1.- P.80-87.
56. Brunelli R., Poggio T. Caricatural effects in automated face perception// Biological Cybernetic.- 1993.- No.69.- P.235-241.
57. Brunelli R. and Poggio T. Face recognition: features versus templates// IEEE Trans. on Pattern Analysis and Machine Intel.- 1993. - Vol.15. - No 10. - P. 1042-1052.

58. Buhmann J., Lades M. and Malsburg C. Size and distortion invariant object recognition by hierarchical graph matching// Proceedings of Int. Joint Conf. on Neural Networks.- 1990. - P. 411-416.
59. Burton A.M., Bruce V., Dench N. What's the difference between men and women? Evidence from facial measurements// Perception.- 1993.- No.22.- P.153-176.
60. Castleman K.R. Digital Image Processing, Prentice Hall Int., Englewood Cliffs, 1996. – P. 668.
61. Chang T. C., Huang T. S. and Novak C. Facial feature extraction from color images// Proceedings of IAPR.- 1994.- Vol. 2. - P. 39-43.
62. Chellappa R., Wilson C. L., Sirohey S. and Barnes C. S., Human and machine recognition of faces: a survey// Proceedings of IEEE.- 1995. - Vol. 83.- No 5. – P. 705 –739.
63. Choi C., Aizawa K., Harashima H. and Takebe T. Analysis and synthesis of facial image sequences in model-based image coding// IEEE Trans. on Comp. Science and Vision Technologies. - 1994. - Vol.4. - No 3. – P. 257-274.
64. Colmenarez A. J. and Huang T.S. Frontal view face detection// Proceedings of SPIE. - 1995. - Vol.2501. - Part 1. - P. 90-98
65. Cox I. J., Ghosn J. and Yianilos P. N. Feature-based face recognition using mixture distance// NEC Research Institute, Technical Report #95-09, 1995.- P.158.
66. Curran L.J. OCR and image processing enhance border security// Image Processing Europe. – 2000. – Jan./Feb. - P. 26- 29.
67. Daugmann J.G. High confidence visual recognition of persons by a test of statistical independence// IEEE Trans. on Pattern Analysis and Machine Intel.- 1993.-Vol.15.- No.11.- P.1148-1161.
68. Daugman J.G. Phenotypic versus genotypic approaches to face recognition// NATO ASI on Face Recognition, Stirling, UK.- 1997.- P.269-283.
69. DeMers D. and Cottrell G. W. Non-linear dimensionality reduction// Advances in Neural Information Processing Systems, Morgan Kaufmann Publishers.- San Mateo.- USA.- 1993.
70. Desk research// BEE Consortium, Ed. D. Polemi, Technical Report #IST-1999-20078. – 2001. – P. 115.
71. DeSilva L.C., Aizawa K. and Hatori M. Detection and tracing of facial features// Proceedings of SPIE. - 1995. - Vol. 2501. - Part 2. - P.1161-1172.
72. Duc B., Fischer S., Bigün J. Face authentication with Gabor information on deformable graphs// IEEE Trans. on Pattern Analysis and Machine Intel.- 1999.-Vol.8.- No.4.- P.504-516.

73. Dunn D., Higgins E. and Wakeley J. Texture segmentation using 2-D Gabor elementary functions// IEEE Trans. on Pattern Analysis and Machine Intel.- 1994.-Vol.16.- No.2.- P.130-149.
74. Eickeler S., Müller S. and Rigoll G. Improved Face Recognition Using Pseudo-2D Hidden Markov Models// Proceedings of Workshop on Advances in Facial Image Analysis and Recognition Technology in conjunction with 5th European Conference on Computer Vision. - Germany. – 1998. – P.125-134.
75. Esme B., Sankur B. and Anarim E. Facial feature extraction using genetic algorithms// Proceedings of European Signal Processing Conf. – Trieste.- 1996. – P. 1511-1514.
76. Essa M. and Pentland A. A Vision system for observing and extracting facial action parameters// M.I.T. Media Laboratory, Perceptual Computing Section, Technical Report # 247.-1994. - P. 76-83.
77. Feraud R., Bernier O.J., Viallet J-E., Collobert M A fast and accurate face detector based on neural networks// IEEE Trans. on Pattern Analysis and Machine Intel.- 2001.-Vol.23.- No.1.- P.42-52.
78. Feraud R., Berliner O., Viallet J., Collobert M. and Collobert D. A conditional mixture of neural networks for face detection, applied to locating and tracking an individual speaker// Proceedings of Int. Conf. on Computer Analysis of Images and Patterns. - Germany. - 1997. - Vol.1256. - P. 464-471.
79. Gabor D. Theory of communication// Journal of Institute of Electric Engineers (London). – 1946.– Vol.93.- No.3. – P.429-457.
80. Golomb L. A., Lawrence D. T. and Sejnowski T. J. SexNet: A neural network identifies sex from human faces// Advances in Neural Information Processing Systems, Morgan Kaufmann Publishers.- San Mateo.- USA.- 1991. - P.77-83.
81. Guo G.-D. and Zhang H.-J. Boosting for fast face recognition// Proceedings of Int. Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real Time Systems.- 2001. – P.96-100.
82. Gutta S. and Wechsler H. Face recognition using hybrid classifiers// Pattern Recognition, 1997. - Vol. 30. - No 4. - P.539-553.
83. Hassanien A.E. and Nakajima M. Image morphing with scattered data points based on snakes and thin plate spline transformation// Machine Graphics & Vision.- 1997. - Vol.6. - No. 3. – P. 341-451.
84. Hassanien A.E. and Nakajima M. An efficient cross-dissolve transformation with elastic body spline warping interpolation for facial image morphing// Machine Graphics & Vision.- 1998. - Vol.7. - Nos. 1/2. – P.397-408.

85. Hjelmas E. Biometric Systems: A Face Recognition Approach// Proceedings of the Norwegian Conference on Informatics. - 2000. – P. 89-98.
86. Intrator N., Reisfeld D. and Yeshurum Y. Face recognition using hybrid supervised/unsupervised neural network // Proceedings of IAPR.- 1994. - Vol.2. - P. 50-54.
87. Ioannou D., Huda W. and Laine A.F. Circle recognition through a 2D Hough Transform and radius histogramming// Image and Vision Computing. - 1999.- Vol.17.- P.15-26.
88. Jain A.K., Hong L. and Pankanti Sh. Biometric identification// Communications of the ACM. – 2000. – Vol.43. – N.2. – P.91 – 98.
89. Jain A.K. Fundamentals of Digital Image Processing, Prentice Hall, Englewood Cliffs.- 1989. – P. 570.
90. Jain A.K., Ratha N.K. and Lakshmanan S. Object detection using Gabor filters// Pattern recognition. – 1997. – Vol.30. – N.2. – P.295 – 309.
91. Jeng S-H., Liao H-Y., Lui Y-T. and Chern M-Y. An efficient approach for facial feature detection using geometrical face model// Proceedings Int. Conf. on Pattern Recognition. - 1996. - Vol. 4. - P. 426-430.
92. Jeng S-H., Liao H-Y., Han C., Chern M-Y and Lui Y-T., Facial feature detection using geometrical face model: an efficient approach// Pattern Recognition.- 1998.-Vol.31.- No.3.- P.273-282.
93. Jesorsky O., Kirchberg K.J. and Frischholz R.W. Robust face detection using the Hausdorff distance// Lecture Notes in Computer Science.- Vol.2091.- Springer.- 2001.- P. 90-95.
94. Kalocsai P., von der Malsburg C. and Horn J. Face recognition by statistical analysis of feature detectors// Image and Vision Computing. - 2000.- Vol.18.- P.273-278.
95. Kamel M.S., Shen H.C., Wong A.K.C. and Hong T.M. Campeanu R.I. Face recognition using perspective invariant features// Pattern Recognition Letters. - 1994.- No.15.- P.877-883.
96. Kanade T. Picture processing by computer complex and recognition of human faces// PhD thesis, Kyoto University, 1973. – P.178.
97. Katahara S. and Aoki M. Face part extraction windows based on bilateral symmetry of gradient direction// Proceedings of Int. Conf. on Computer Analysis of Images and Patterns. – Slovenia. - 1999.- P.489-497.
98. Klette R., Zamperoni P. Handbook of image processing operators// John Wiley & Sons.- Chichester.- 1996.- P.398.
99. Kohonen T. Self-organization and associative memory// Springer-Verlag. – Berlin. - 1988. – P. 364.

100. Kotropoulos C., Tefas A., Pitas I. Frontal face authentication using morphological elastic graph matching// IEEE Trans. on Image Processing.- 2000.-Vol.9.- No.4.- P.555-560.
101. Krüger N. An algorithm for the learning of weights in discrimination functions using a priori constraints// IEEE Trans. on Pattern Analysis and Machine Intel.- 1997.-Vol.19.- No.7.- P.764-768.
102. Kruizinga P. and Petkov N. Optical flow applied to person identification// Proceedings of Conf. on Massively Parallel Processing Applications and Development.- Netherlands.- 1994. - P. 871-878.
103. Kuo C., Huang R. and Ling T. Synthesizing lateral face from frontal facial image using anthropometric estimation// Proceedings of Int. Conf. on Image Proc. - 1997. - Vol.1. - P. 133-136.
104. Lades M., Vorbruggen J., Buhmann J., Lange J., Masburg C. and Wurtz R. Distortion invariant object recognition in the dynamic link architecture// IEEE Trans. on Computers.- 1993. - Vol. 42. - P. 300-311.
105. Lam K. M., Yan H. An analytic-to-holistic approach for face recognition based on a single frontal view// IEEE Trans. on Pattern Analysis and Machine Intel.- 1998.-Vol.20.- No.7.- P.673-686.
106. Lam K. M., Yan H. Locating and extracting the eye in human face images// Pattern Recognition.- 1996.-Vol.29.- No.5.- P.771-779.
107. Lawrence S., Giles C. L., Tsoi A. C. and Back A. D. Face recognition: a convolutional neural network approach// IEEE Trans. on Neural Networks.- Special Issue on Neural Networks and Pattern Recognition.- 1997.- P.97-113.
108. Lee T.S. Image representation using 2D Gabor wavelets// IEEE Trans. on Pattern Analysis and Machine Intel.- 1996.-Vol.18.- No.10.- P.959-971.
109. Lee Ch. H., Kim J.S., Park K.H. Automatic human face location in a complex background using motion and color information// Pattern Recognition.- 1996.-Vol.29.- No.11.- P.1877-1889.
110. Leroy B., Herlin I., Cohen L.D. Face identification by deformation measure// Proceedings Int. Conf. on Pattern Recognition. - 1996. - Vol. 5. - P.633-637.
111. Li S.Z. and Lu J. Face recognition using the nearest feature line method //IEEE Trans. on Neural Networks. - 1999. - Vol.10.- No.2.- P.439-443.
112. Li X., Roeder N. Face contour extraction from front-view images// Pattern Recognition. - 1995. - Vol.28.- No.8.- P.1167-1179.
113. Li Y. and Kobatake H. Extraction of facial sketch images and expression transformation based on faces// Proceedings of Int. Conf. on Image Proc. - 1995. - Vol. 3. - P. 520-524.

114. Matsuno K. and Tsuji S. Recognizing human facial expressions in a potential field // Proceedings of IAPR.- 1994.- Vol.2. - P. 44-49.
115. Moghaddam B. and Pentland A. Face recognition using view based and modular eigenspaces// Automatic Systems for the identification and inspection of humans.- SPIE.- 1994. - Vol. 2257. - P.1868-1876.
116. Moghaddam B., Nastar C. and Pentland A. A Bayesian similarity measure for direct image matching// Pattern recognition and Signal Analysis.- 1996. - P.350-358.
117. Morimoto C., Koons D., Amir A. and Flickner M. Pupil detection and tracking using multiple light source// Image and Vision Computing, special issue on Advances in Facial Image Analysis and Recognition Technology. - No. 4. – 2000. – P.331-335.
118. Nakamura O., Mathur S. and Minami T. Identification of human faces based on isodensity maps// Pattern Recognition.- 1991. - Vol. 24. - P. 263-272.
119. Nastar C. and Moghaddam B. Generalized Matching for recognition and retrieval in an image database// Proceedings of Int. Conf. on Communicating by Image and Multimedia.- France.- 1996.-P. 366-374.
120. Nelson L. Commercializing face recognition: how to judge fresh players and approaches// Advanced Imaging.- 1996. – P. 85-86.
121. Norton R.E. Biometric products: Revenues could hit \$2.5 billion in ten years// Biometric Digest. – 2001. – May. – P.11-12.
122. O'Toole A., Abdi H., K. Deffenbacher and D. Valentin Low-dimensional representation of faces in higher dimensions of face space// Journal Opt. Soc. Am. A.- 1993. - Vol. 10. - No 3. - P. 405-411.
123. Osuna E., Freund R., Girosi F. Training support vector machines: an application to face detection// Proceedings of Int. Conf. on Computer Vision and Pattern Recognition. – Puerto Rico, - 1997. - P. 256-264.
124. Otsu N., A threshold selection method from the gray-level histograms// IEEE Trans. on Syst., Man, Cybern.- 1979.-Vol.SMC-9.- No.1.- P.62-67.
125. Owen C. B. and Makedon F. Bottleneck-Free Separable Affine Image Warping// Proceedings of Int. Conf. on Image Processing.- 1997. – P.683-686.
126. Parker J. R. Practical Computer Vision Using C, Wiley & Sons, New York, 1993. – P. 476.
127. Pentland A., Moghaddam B. and Starner T. View-based and modular eigenspaces for face recognition// M.I.T. Media Laboratory, Perceptual Computing Section, Technical Report # 245.- 1994. - P. 84-91.
128. Phillips P.J., Moon H., Rizvi S.A. and Rauss P.J. The FERET evaluation methodology for face-recognition algorithms// National Institute of Standards and Technology, Technical Report #6264.- USA.- 1999. – P. 20.

129. Pitas I. and Venetsanopoulos A.N. Nonlinear digital filters, Kluwer Academic Publishers.- Boston.- 1990. – P.392.
130. Pitas I. Digital Image Processing Algorithms, Prentice Hall.- New York.- 1993. – P.361.
131. Quan Sh. and Chen D. Discrete Gabor transform// IEEE Trans. on Signal Processing.- 1993.-Vol.41.- No.7.- P.2429-2438.
132. Raducanu B., Grana M., Albizuri F. X., d'Anjou A. Face localization based on the morphological multiscale fingerprints// Pattern Recognition Letters.- 2001.-Vol.22.- No.3-4.- P.359-371.
133. Ratan A. L., Grimson W., Wells W.M. Object detection and localization by dynamic template warping// Int. Journal of Computer Vision.- 2000.- Vol.36.- No.2.- P.131-147.
134. Rogers W. Birmingham City Centre CCTGV installs Visionics' FaceIt// Biometric Digest. – 2001. – May. – P.3-4.
135. Rogers W. ETrue partners with Bio4 to combine facial biometrics// Biometric Digest. – 2001. – May. – P.4-5.
136. Romdhani S. Face recognition using principal components analysis// MSc thesis.- University of Glasgow.- 1997.- P.183.
137. Russ J.C. The mage Processing Handbook, CRC Press, Boca Raton, 1995. – P.674.
138. Saco H., Whitehouse M., Smith A. and Sutherland A. Real-time facial feature tracking based on matching techniques and its application// Proceedings of IAPR.- 1994. - Vol. 2. - P. 320-324.
139. Samal D.I., Taleb M.A. and Starovoitov V.V. Experiments with preprocessing of human portraits// Proceedings of Int. Conf. on Pattern Recognition and Image Processing. – 2001. – Minsk. - P. 15-20.
140. Samal A. and Iyengar P.A. Automatic recognition and analysys of human faces and facial expressions: a survey// Pattern Recognition. - 1992. - Vol. 25. - No 1. - P. 65-77.
141. Samaria F. and Young S. A HMM-based architecture for face identification// Image and Vision Computing. - 1994. - No 8. – P.537-543.
142. Smeraldi F., Carmona O. and Bigun J. Saccadic search with Gabor features applied to eye detection and real-time head tracking// Image and Vision Computing. - 2000. - No 18. - P. 323-329.
143. Starovoitov V. and Samal D. A distance based dissimilarity measure for color images// Proceedings of Int. Conf. on Signal and Systems in Intelligent Technologies.- Minsk.- 1998.- P. 298-306.

144. Starovoitov V. and Samal D. A geometric approach to face recognition// Proceedings of Workshop on Nonlinear Signal Image Processing.- Antalia, Turkey.- 1999.- P. 210-213.
145. Starovoitov V. and Samal D. Comparison of rotation algorithms for digital images// Proceedings of Int. Symposium on Optical Science, Engineering, and Instrumentation, Vision Geometry VIII. -1999. - SPIE Vol. 3811.- Denver, USA.- P. 346-356.
146. Starovoitov V. and Samal D. Experimental study of color image similarity// Machine Graphics & Vision, - 1998. - Vol. 7. - No. ½, - P. 445-462.
147. Starovoitov V. and Samal D. Geometric features for face recognition// Proceedings of Int. Conf. UkrOBRAZ'98.- Kyjiv.- Ukraine. -1998. - P. 109-110.
148. Starovoitov V.V. and Samal D. I. Human identification for document control// Proceedings of Int. Conf. on Pattern Recognition and Image Processing. – 2001.- Minsk. – P. 192-196.
149. Starovoitov V., Samal D. and Sankur B. Matching of faces in camera images and document photographs// Proceedings of Int. Conf. on Acoustic, Speech, and Signal Processing. – 2000. - vol. IV.- Istambul., P.2349-2352.
150. Starovoitov V., Samal D., Votsis G., and Kollias S. Face recognition by geometric features// Proceedings of Int. Conf. on Pattern Recognition and Image Processing. - Minsk.- 1999.- P. 39-43.
151. Sung K. K., Poggio T. Example-based learning for view-based human face detection// IEEE Trans. on Pattern Analysis and Machine Intel.- 1998.- Vol.20.- No.1.- P.39-51.
152. Swets D. L. and Weng J. Using discriminant eigenfaces for image retrieval // IEEE Trans. on Pattern Analysis and Machine Intel.- 1996. - Vol. 18. - No 8. - P. 831-836.
153. Tai B., Healy G. Modeling and classifying symmetries using a multiscale opponent color representation // IEEE Trans. on Pattern Analysis and Machine Intel.- 1998.-Vol.20.- No.11.- P.1224-1235.
154. Takacs B. Comparing face images using the modified Hausdorff distance// Pattern Recognition.- 1998.-Vol.31.- No.12.- P.1873-1881.
155. Takas B. and Wechsler H. Locating features using SOFM// Proceedings of IAPR.- 1994. - Vol. 2. - P. 55-60.
156. Tankus A., Yeshurum Y. and Intrator N. Face detection by direct convexity estimation// Pattern Recognition Letters.- 1997. - Vol. 18. - P. 913-922.
157. Tistarelli M. and Grossi E. Active vision-based face authentication// Image and Vision Computing. – 2000. - No 18. - P. 299-314.

158. Troje N.F. and Bülthoff H.H. Face recognition under varying pose: The role of texture and shape// Max-Planck-Institut fur biologische kybernetik, Technical Report #17.-1995. – P.128.
159. Troje N.F. and Bülthoff H.H. How is bilateral symmetry of human faces used for recognition of novel views// Max-Planck-Institut fur biologische kybernetik, Technical Report #38.-1996. – P.134.
160. Turk M. and Pentland A. Eigenfaces for recognition// Journal of Cognitive Neuroscience.- 1991. - № 3. - P. 71-86.
161. Umbaugh S.E. Computer Vision and Image Processing , Prentice Hall PTR.- Upper Saddle River.- 1999. – P. 504.
162. Valentin D. and Abdi H. Can linear autoassociator recognize faces from new orientations?// Journal Opt. Soc. Am. A.- 1996. - Vol.13. – P.522-530.
163. Valentin D., Abdi H., O'Toole A. J. and Cottrell G. W. Connectionist models of face processing: a survey// Pattern Recognition.- 1994. - Vol. 27. - P. 1209-1230.
164. de Vel O., Aeberhard S. Line-based face recognition under varying pose// IEEE Transactions on Pattern Analysis and Machine Intel.- 1999.- Vol.21.- No.10.- P.1081-1088.
165. Wang J., Tan T. A new face detection method based on shape information// Pattern Recognition Letters.- 2000.-Vol.21.- No.5.- P.463-471.
166. Weber F., Hernandez A.H. Face location by template matching with a quadratic discriminant function// Proceedings of Int. Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real Time Systems.- Greece. - 1999. – P.10-13.
167. Weng J., Ahuja N. and Huang T.S. Learning recognition and segmentation of 3-d objects from 2-d images// Proceedings of Int. Conf. on Computer Vision. – 1993. - P. 121-128.
168. Wieghardt J. and Loos H.S. Finding faces in cluttered still images with few examples// Proceedings of Int. Conf. on Artificial Neural Networks. – Austria.- 2001. – P.1026-1033.
169. Wiskott L. Phantom faces for analysis// Pattern Recognition. - 1997. - Vol.30. - No.6. - P.837-846.
170. Wiskott L. Phantom faces for analysis// Proceedings of Int. Conf. on Computer Analysis of Images and Patterns. – Germany. - 1997. - Vol.1256. - P. 480-487.
171. Wiskott L., Fellous J. M., Krüger N., von der Malsburg C. Face recognition by elastic bunch graph matching// IEEE Trans. on Pattern Analysis and Machine Intel.- 1997.-Vol.19.- No.7.- P.775-779.

172. Wu C. and Huang J. Human face profile recognition by computer// Pattern Recognition.- 1990. -Vol.23. - No. 3/4. – P. 255-259.
173. Wu H., Fukumoto T., Chen Q. and Yashida M. Active face observation system// Proceedings of Int. Conf. on Pattern Recognition. - 1996. - Vol. 4. - P. 441-445.
174. Wu H., Chen Q. and Yashida M. Facial feature extraction and face verification// Proceedings of Int. Conf. on Pattern Recognition. - 1996. - Vol. 4. - P. 484-488.
175. Yacoob Y. and Davis L. Computing Spatio-Temporal Representations of Human Faces// Proceedings of Int. Conf. on Computer Vision and Pattern Recognition. – Seatle. – USA. – 1994. – P. 70-75.
176. Yang G., Huang T., Human face detection in complex background// Pattern Recognition. – 1994. - Vol.27. - No.1.- P.53-63.
177. Yemez Y., Aydin T., Sankur B. and Amarim E. A new approach to 3-D face modeling// Proceedings of European Conference on Circuit Theory and Design.- Istambul.- 1995 - P.343-346.
178. Yin L. and Basu A. MPEG4 Face Modeling Using Fiducial Points// Proceedings of Int. Conf. on Image Proc. - 1997 – P. 109-112.
179. Yow K. and Cipolla R. Feature-based human face detection// Image and Vision Computing. - 1997. - No 15. - P. 713-735.
180. Yuen P. C. and Feng G. C. Automatic Eye Detection for Human Identification// Proceedings of Scandinavian Conf. on Image Analysis. - Finland. - 1997. - P. 1293-1300.
181. Zhang Y. A fuzzy approach to digital image warping //IEEE Computer Graphics and Applications.- 1996.- No.7.- P.34-41.
182. Zhao W., Chellappa R., Rosenfeld A. and Phillips P.J. Face recognition: a literature survey// National Institute of Standards and Technology, Technical Report #7478.- USA.- 2001. – P. 66.

ПРИЛОЖЕНИЕ 1

Дополнительные иллюстрации к главам 2, 3

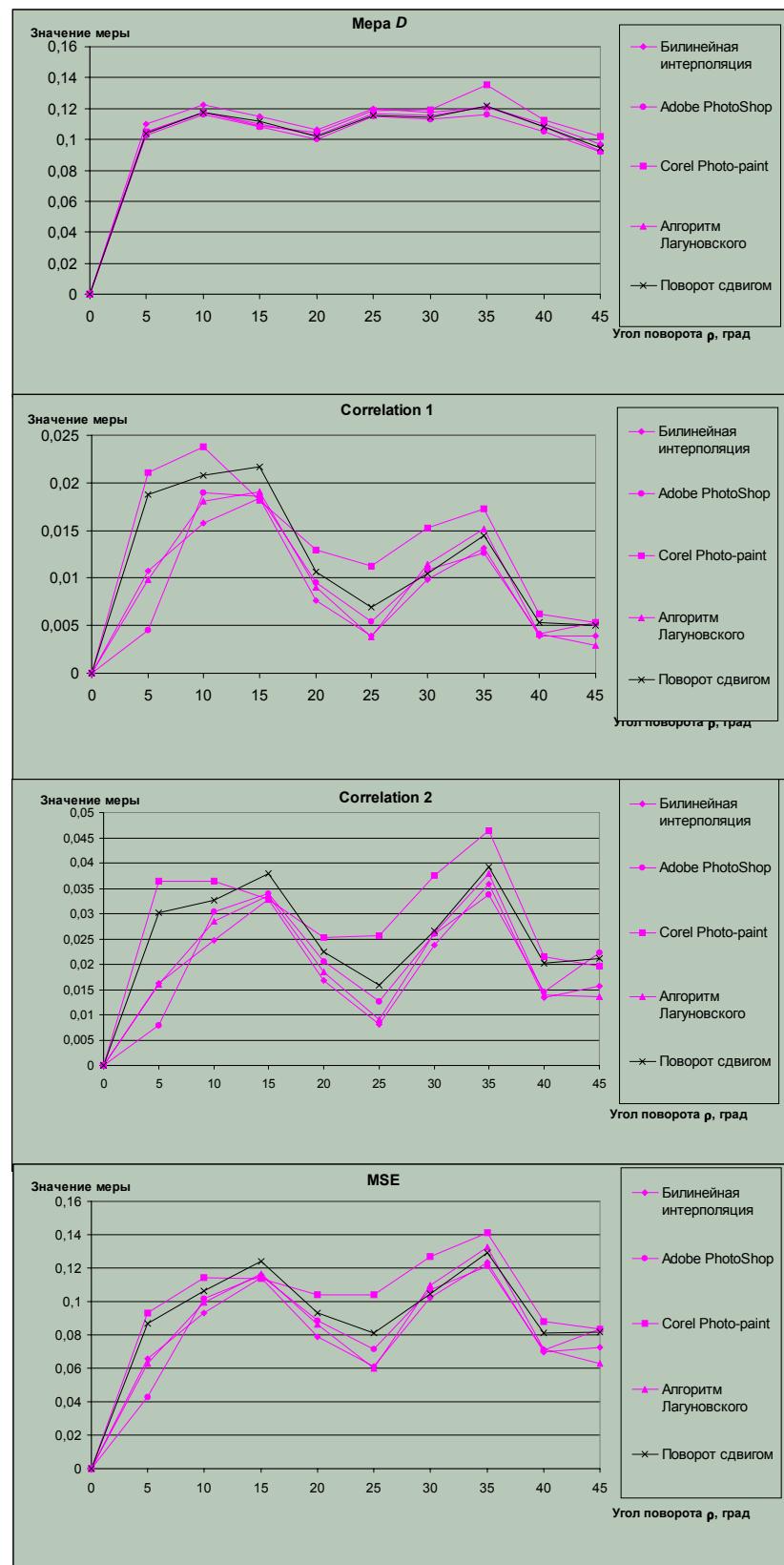


Рис. П.1.1. Результаты сравнения изображений \mathbf{G}^0 и \mathbf{G}_k^0 с помощью меры D, коэффициентов корреляции Cor1 и Cor2, оценки среднеквадратичной ошибки

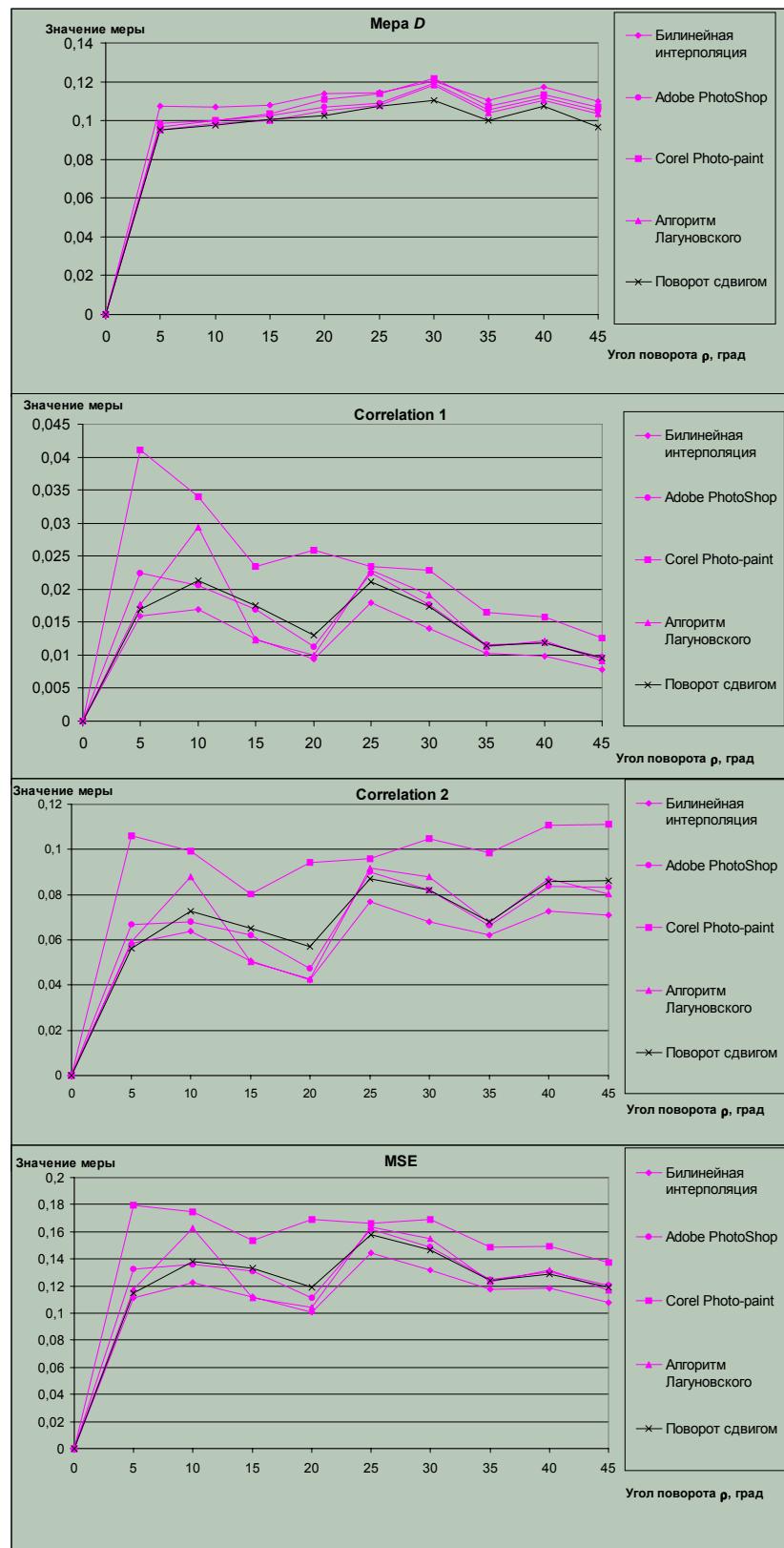


Рис. П.1.2. Результаты сравнения изображений N^P и N_k^P с помощью меры D, коэффициентов корреляции Cor1 и Cor2, оценки среднеквадратичной ошибки



Рис. П.1.3. Изменения лица человека на протяжении 30 лет на примере изображений двух классов S50 и S51 из экспериментальной базы данных



Рис. П.1.4. Пример возможного изменения внешности человека за период 2 года (изображения класса S_28 из экспериментальной базы данных)

ПРИЛОЖЕНИЕ 2

**Документы об использовании результатов
диссертационной работы**