МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Южно-Уральский государственный университет»
(национальный исследовательский университет)
Факультет «Экономика и предпринимательство»
Кафедра «Информационные системы»

ПРОЕКТ ПРОВЕРЕН
Рецензент, к.т.н., доцент
кафедры «Финансовый менеджмент»
_____ (М.К. Ветров)
«22» января 2016 г.

ДОПУСТИТЬ К ЗАЩИТЕ
Д.т.н., проф., зав. кафедрой
«Информационные системы»
_____ (В.В. Мокеев)
«25» января 2016 г

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ МЕТОДОВ ОБРАБОТКИ ИЗОБРАЖЕНИЙ В СИСТЕМАХ РАСПОЗНАВАНИЯ ЛИЦ

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ
(МАГИСТЕРСКОЙ ДИССЕРТАЦИИ)
ЮУрГУ–080500.68.2016.126 ПЗ ВКР

Руководитель проекта
д.т.н., проф., зав.
кафедрой «Информационные системы»
_____ (В.В. Мокеев)
«22» января 2016 г.

Автор магистерской диссертации
студент-магистрант группы ЭиП-344
_____ (А.И. Гришин)
«22» января 2016 г.

Нормоконтролер, к.т.н., доцент
кафедры «Информационные системы»
_____ (Е.В.Бунова)
«19» января 2016 г.

Челябинск
2016

КАЛЕНДАРНЫЙ ПЛАН

Наименование этапов выпускной квалификационной работы	Срок выполнения этапов работы	Отметка о выполнении руководителя
1 Анализ и общая характеристика отрасли машинного зрения и искусственного интеллекта	18.05.15 – 25.05.15	
2 Теоретические основы алгоритмов обработки изображений	18.05.15 – 1.06.15	
2.1 Метод МГК	25.05.15 – 28.05.15	
2.2 Метод ЛБШ	28.05.15 – 1.06.15	
3 Разработка и подготовка программного обеспечения исследования	10.09.15 – 10.10.15	
4 Проведение экспериментального исследования и фиксация результатов	10.11.15 – 24.11.15	
5 Анализ полученных результатов, сравнение, разработка рекомендаций	24.11.15 – 15.12.15	

Заведующий кафедрой _____/д.т.н, проф. В.В. Мокеев
(подпись)

Руководитель работы (проекта) _____/д.т.н, проф. В.В. Мокеев
(подпись)

Студент _____/А.И. Гришин
(подпись)

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ «ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ» (НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)
ФАКУЛЬТЕТ «ЭКОНОМИКА И ПРЕДПРИНИМАТЕЛЬСТВО»
КАФЕДРА «ИНФОРМАЦИОННЫЕ СИСТЕМЫ»
НАПРАВЛЕНИЕ 080500.68 – БИЗНЕС-ИНФОРМАТИКА
МАГИСТЕРСКАЯ ПРОГРАММА – ЭЛЕКТРОННЫЙ БИЗНЕС

УТВЕРЖДАЮ
Заведующий кафедрой,
д.т.н. профессор
_____Мокеев В.В.
«10» декабря 2015 г.

ЗАДАНИЕ
на выпускную квалификационную работу магистра

Гришина Антона Игоревича
(Ф.И.О. полностью)
ГруппаЭиП-344

1 Тема работы Исследование эффективности методов обработки изображений в системах распознавания лиц

утверждена приказом по университету от «30»_____декабря_ 2015 г. № 2630

2 Срок сдачи студентом законченной работы «16»__января_2016 г.

3 Исходные данные к работе (проекту):

4 Перечень вопросов, подлежащих разработке:

– рассмотреть различные подходы и алгоритмы к обработке изображений для распознавания лиц;

– определить варьируемые параметры методов выбранных для исследования и сравнения;

– разработать алгоритмы формирования тестового и обучающего наборов для проведения исследования;

– провести экспериментальное исследование выбранных методов на подготовленном исходном материале.

– проанализировать и сравнить результаты показанные различными методами, дать рекомендации по их использованию.

5 Иллюстративный материал (плакаты, альбомы, раздаточный материал, макеты, электронные носители и др.):

– 1 электронный диск с исходными материалами к дипломной работе;

– раздаточный материал к работе, содержащий презентацию печатного вида;

6 Дата выдачи задания«11» мая 2015 г.

Руководитель_____/д.т.н, проф. В.В. Мокеев
(подпись)

Задание принял к исполнению _____/А.И. Гришин
(подпись студента)

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Южно-Уральский государственный университет»
(национальный исследовательский университет)
Факультет «Экономика и предпринимательство»
Кафедра «Информационные системы»

ПРОЕКТ ПРОВЕРЕН
_____
_____
_____( )
«____»_____2016 г.

ДОПУСТИТЬ К ЗАЩИТЕ
Заведующий кафедрой, д.т.н.,
профессор
_____(В.В. Мокеев)
«____»_____2016 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
«ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ МЕТОДОВ ОБРАБОТКИ ИЗОБРАЖЕНИЙ В СИСТЕМАХ РАСПОЗНАВАНИЯ ЛИЦ»
ЮУрГУ-080700.68.2016.120.ПЗ ВКР

Руководитель,
Заведующий кафедрой, д.т.н.,
профессор
_____(В.В. Мокеев)
«____»_____2016 г.

Автор работы,
студент группы ЭиП - 344
_____(А. И. Гришин)
«____»_____2016г.

Нормоконтролер, к.т.н., доцент
_____(Е. В.Бунова)
_____«____»_____2016 г.

Челябинск 2016

АННОТАЦИЯ

Гришин Антон Игоревич, выпускная квалификационная работа на тему: «Исследование эффективности алгоритмов распознавания лиц в системах видеонаблюдения». – Челябинск: ЮУрГУ, 67 с.., 19 ил., 4 табл., 3 прил., библиогр. список – 36 наим.

В работе приведены различные подходы к классификации методов распознавания лиц, приведен обзор и краткое описание соответствующих методов. Кроме того рассматриваются существующие подходы к построению выборочной совокупности, эти методы могут быть использованы для формирования обучающего и тестового набора на основе имеющихся исходных данных.

Подробнее рассматриваются, выбранные для исследования, метод главных компонент и метод локальных бинарных шаблонов для решения задачи распознавания изображений. Описан принцип их действия и алгоритм работы.

Описываются результаты распознавания изображений лиц на изображениях баз данных ORL и FERET, выбранными методами с применением различных параметров. Результаты сравниваются между собой с целью определения наиболее эффективных вариантов. Рассматривается влияние использования различных параметров метода на эффективность распознавания, а также влияние степени нормализации исходного материала на эту эффективность.

ОГЛАВЛЕНИЕ

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ 2
ВВЕДЕНИЕ 3
Глава 1. Обзор существующих подходов и методов 7
Классификация методов распознавания лиц 7
1. 1. 1 Классификация по основным классам решаемых задач 7
1. 1. 2 Классификация по инвариантности методов к вариациям исходного изображения и способы её достижения 8
1. 1. 3 Классификация по алгоритмическим особенности методов распознавания 10
1. 1. 4 Классификация по используемым признакам объектов 15
Способы построения выборки 28
1. 2. 1 Перекрестная проверка 29
1. 2. 2 Вероятностные способы 31
1. 2. 3 Детерминированные способы 33
Глава 2. Теоретический базис: Исследуемые методы распознавания лиц 34
2. 1 Метод главных компонент 34
Алгоритм работы метода главных компонент. 37
2. 2 Метод локальных бинарных шаблонов 39
Алгоритм работы метода локальных бинарных шаблонов 40
Глава 3. Экспериментальное исследование на основе метода главных компонент и метода локальных бинарных шаблонов. 43
3. 1 Исходные данные: базы изображений лиц 43
3. 1. 1 База ORL 43
3. 1. 2 База FERET 43
3. 2 Формирование выборки 44
3. 2. 1 Создание обучающих и тестовых наборов 45
3. 2. 2 Наборы, сформированные для обучения и проверки классификаторов 46
3. 3 Программное обеспечение исследования 47
3. 3. 1 Программная платформа 47
3. 3. 2 Среда разработки 47
3. 3. 3 Язык программирования 48
3. 3. 4 Расширения для работы с изображениями 49
3. 4 Результаты, полученные в ходе исследования 50
3. 4. 1 Результаты, показанные на материале базы ORL 50
3. 4. 2 Результаты, показанные на материале базы FERET 54
3. 4. 3 Сравнение результатов методов для баз FERET и ORL 59
ЗАКЛЮЧЕНИЕ 62
БИБЛИОГРАФИЧЕСКИЙ СПИСОК 63
ПРИЛОЖЕНИЕ А. Формирования тестового и обучающего наборов 69
ПРИЛОЖЕНИЕ Б. Многопотоковая обработка выборки изображений 70
ПРИЛОЖЕНИЕ В. Вспомогательные функций формирования выборок 73

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

МГК – метод главных компонент;

PCA - Principal Component Analysis;

ЛБШ – локальный бинарный шаблон;

LBP – local binary pattern:

ЛДА – линейный дискриминантный анализ;

НС – нейронная сеть;

СНС – сверточная нейронная сеть;

МНС – многослойная нейронная сеть;

СММ – скрытые марковские модели;

SVM – Support Vector Machine

ORL – Olivetti Research Laboratory;

FERET – Face Recognition Technology;

DARPA – Defense Advanced Research Projects Agency;

CLR – Common Language Runtime;

OpenCV – Open Source Computer Vision Library.

ВВЕДЕНИЕ

Задача выделения лица человека в естественной или искусственной обстановке и последующей идентификации всегда находилась в ряду самых приоритетных задач для исследователей, работающих в области систем машинного зрения и искусственного интеллекта.

Во множестве исследований, проведенных в рамках решения данной задачи в течение нескольких десятилетий, получены самые различные результаты, как для отличающихся друг от друга методик, так и для методик, использующих одни и те же алгоритмы, так и для систем, основывающихся на одном и том же алгоритме.

Конечно, алгоритмы и методы, используемые для идентификации изображений лица, постоянно совершенствуются, однако, и среди уже существующих методов, сложно найти объективное и достоверное сравнение их эффективности и результатных показателей.

Для объективного сравнения различных методов распознавания изображений необходимо использовать единую методику проверки и аналогичную выборку изображений. Кроме того, сам способ разделения выборки на тестовую и обучающую части должен иметь минимальную зависимость от предъявляемого набора данных.

Для того, чтобы выделить наиболее эффективный метод идентификации по изображению лица необходимо применить математически обоснованный метод оценки того, насколько успешны исследуемые методы на независимом наборе данных. Кроме того, следует оценить влияния цветовой нормализации изображения и исключить влияние его геометрической нормализации. А также рассмотреть зависимость результатов от объема тестовой выборки для того, чтобы выделить наиболее эффективный метод идентификации минимизировать влияние внешних параметров на оценку результатов.

Объектом исследования являются системы распознавания лиц.

Предмет исследования

Любая система распознавания изображений выполняет ряд соответствующих её назначению функций, одной из таких функций является идентификация личности по изображению лица, поэтому предметом исследования стали методы обработки изображений с целью распознавания лиц присутствующих на изображении. Разумеется, результаты такой обработки во многом зависят от таких особенностей, как качество изображения (объем информации, которое оно содержит), условия освещенности, зашумленность и так далее. Это означает, что значительную роль в эффективности работы метода играют его параметры, обеспечивающие устойчивость к действию неблагоприятных факторов, обусловленных спецификой исходного материала.

Целью исследования проведенного в рамках данной работы является повышение эффективности и качества работы систем распознавания лиц, за счеты выбора наиболее эффективного метода распознавания, и подбора необходимого в конкретных условиях сочетания параметров этого метода.

Задачи исследования

Для достижения этой цели были поставлены следующие задачи, которые необходимо решить в рамках исследования:

1. Выбрать методы и алгоритмы, с помощью которых будет проводиться идентификация и распознавание лиц, выделенных на отдельном изображении.
2. Определить варианты задаваемых параметров для данных методов, которые могут оказать влияние на эффективность распознавания.
3. Подготовить широкую выборку изображений, в различных условиях освещенности на которых присутствуют изображения лиц с различных ракурсов, заранее соотнесенных с определенными личностями.
4. Определить методику, с помощью которой имеющаяся выборка будет наиболее объективным способом разделена на обучающий и тестовый наборы.
5. Провести обработку изображений с применением различных методов, задавая различные значения параметров.

6. Применить выбранные варианты обработки исходного материала для различных сочетаний изображений в обучающем и тестовом наборе.

7. Соотнести полученные результаты, между собой, тем самым сравнить эффективность.

Научная новизна

В работе рассматриваются некоторые аспекты использования алгоритмов идентификации личности по изображению лица, которые могут влиять на эффективность такого распознавания. Для достижения большей объективности были решены следующие вопросы:

1. Разработан алгоритм формирования тестовой и обучающей выборки из неоднородного исходного набора данных;
2. С использованием этого алгоритма проведено исследование и сравнение эффективности методов идентификации личности по изображению лица;
3. Выявлены особенности исходного набора данных, при которых определенный метод идентификации обладает большими преимуществами;
4. Определено, какие входные параметры методов в большей степени влияют на эффективность и точность распознавания методов.
5. На основании результатов исследования были сформулированы рекомендации по использованию методов распознавания лиц.

Практическая ценность

Сегодня системы распознавания лиц применяются в трех областях. Первая – борьба с криминалом. Фоторобот преступника сверяется с изображениями в базах данных либо производится съемка человеческого потока в людных местах, и лица людей в реальном времени сравниваются с лицами нарушителей, находящихся в розыске. Систему можно использовать и в борьбе с банковским мошенничеством.

Распознавание работает в системах контроля доступа. Обычно технология действует в связке с электронными пропусками: образец фото на них сравнивается с моделью, полученной в результате съемки человека, входящего в здание. В этом случае фиксация лица — самый удобный способ идентификации.

Лицо всегда доступно для съемки. Фотографирование происходит мгновенно, и процедура не требует от человека никаких дополнительных действий».

Третья сфера применения, быстро развивающаяся в последнее время, — маркетинг. Человек подходит к интерактивному рекламному стенду, система фиксирует его лицо, строит трехмерную модель и определяет пол и возраст. В результате клиент видит на стенде только ту рекламу, которая потенциально ему интересна. Есть и другая маркетинговая задача: усовершенствовать традиционные программы лояльности. Раньше понять, что организацию посетил постоянный клиент, можно было только на кассе, когда он достанет из бумажника скидочную карту. Теперь, поместив в зале магазина камеру и установив систему распознавания лиц, можно узнавать каждого.

Глава 1. Обзор существующих подходов и методов

Классификация методов распознавания лиц

1. 1. 1 Классификация по основным классам решаемых задач

Задачи распознавания человека по изображению лица делятся на три больших класса: поиск в больших базах данных, контроль доступа и контроль фотографий в документах. Они различаются по требованиям, предоставляемым к системам распознавания, и по способам решения, поэтому представляют собой отдельные классы. В рамках данной работы, актуальным является, первый класс задач. Поскольку любой архив, не столь важно видеоархив это или просто набор изображений, является в первую очередь базой данных.

Различны и требования, предъявляемые к ошибкам первого и второго рода для таких классов [8]. Ошибка первого рода – ситуация, когда объект заданного класса не распознаётся (пропускается) системой. Ошибка второго рода ситуация, когда объект заданного класса принимается за объект другого класса. В экспериментальной части работы, основным вопросом является факт успешного распознавания как таковой, то есть ситуация, когда система относит предъявляемый объект к тому классу, которому он соответствует, во всех остальных ситуациях исход испытания считается неудачей.

Следует также отметить различие понятий верификации и распознавания (идентификации) [20]. В задаче верификации неизвестный объект заявляет, что он принадлежит к некоторому известному системе классу. Система подтверждает или опровергает это. В процессе распознавания нужно отнести объект, принадлежащий неизвестному классу к одному из известных или сделать вывод, что этот объект не относится ни к одному из известных классов.

Поиск изображения в базах данных, большого объема

В этой ситуации предъявляются высокие требования к ошибке первого рода – система распознавания должна находить изображения соответствующие данному человеку, по возможности не пропустив ни одного такого изображения.

Допускается наличие в результирующей выборке небольшого количества других людей.

Часто в больших базах данных нужно найти такие изображения,которые больше всего похожи на то, что подается на вход [10]. Поиск должен производиться за приемлемый период времени. Один из вариантов заключается в том, что в базе данных хранятся некоторые заранее извлечённые наборы важных признаков, определяющих изображение с предельной однозначностью. В данной ситуации требования, предъявляемые к точности, не являются настолько критичными как в задачах контроля доступа или же контроля документов.

К данному классу, в первую очередь, следует отнести метод главных компонент (метод «собственных лиц») [3,20]. Коэффициенты, полученные при разложении входного изображения на главные компоненты, используются для сравнения изображений путём вычисления Евклидова расстояния или же – на основе метрики Махаланобиса с использованием распределения Гаусса [3].

1. 1. 2 Классификация по инвариантности методов к вариациям исходного изображения и способы её достижения

Качество, имеющееся у методов распознавания, инвариантные к таким вариациям как изображения как масштаб, смещение, условия освещённости и т.д. Помимо этого, существуют некоторые способы обработки изображения с целью достижения подобной инвариантности при использовании остальных методов.

1. 1. 2. 1 Изменения масштаба

Неокогнитронам и нейронным сетям высокого порядка присуща инвариантность к изменению масштаба [12,13]. Это означает, что указанные сети не требуют специального обучения на изображениях с различными масштабами объекта, то есть однократное предъявление изображения в качестве обучающего примера, позволяет впоследствии распознавать объект на изображениях с другими вариантами масштаба.

Предполагается, что в других методах этот параметр аппроксимируется при геометрической нормализации изображения, иначе в обучающий набор необходимо включать изображения, имеющие различные вариации масштаба.

1. 1. 2. 2 Изменение условий освещения

Для многих методов критичным является фактор изменения в освещении, при его воздействии результативность распознавания значительно снижается. Выравнивания гистограммы яркостей (световая нормализация) позволяет в какой-то степени решить данную проблему [3,14], хотя чаще всего предобработка не позволяет устранить этот фактор.

В некоторых работах используется алгоритм выделения линейных подпространств, позволяющий провести успешную классификацию при различных условиях освещения, но для достижения этого эффекта необходимо формировать обучающий набор из изображений, на которых присутствуют различные варианты освещения.

Существуют так же различные эвристические подходы. К примеру, в одном из вариантов метода главных компонент, первые две три компоненты, в основном, содержат информацию о различиях в освещении, по этой причине они не учитываются при сравнении[13].

В рамках данной работы будет рассмотрено влияние световой нормализации (выравнивания гистограммы яркостей) на эффективность распознавания исследуемых методов.

1. 1. 2. 3 Изменения ориентации, ракурса и смещению изображения

Во многих методах, необходима предварительная обработка, модифицирующая положение лица так, что оно приобретает стандартную ориентацию в пространстве (линия глаз выравнивается горизонтально, а ось симметрии – вертикально).

Среди методов распознавания изображений, не существует таких, на эффективность которых не влияют большие изменения ракурса. С этой проблемой позволяет справиться лишь обучение на изображениях лица в различных ракурсах.

Разрабатываются методы, позволяющие преодолеть этот недостаток, в их основе лежит синтез изображения в других ракурсах по одному примеру изображения в произвольном ракурсе.

При использовании различных по уровню геометрической нормализации баз изображений, можно оценить влияние этих факторов на исследуемые методы.

1. 1. 2. 4 Внутриклассовые различия

Алгоритмы распознавания работают таким образом, что каждое лицо отдельного человека представляет собой некоторый класс. Внутриклассовые различия, к примеру: очки, борода, прическа или эмоции затрудняют процесс распознавания.

В некоторой степени описанная проблема преодолевается выделением областей лица, неподверженных таким изменениям [3,5,10]. Так же это может решаться в процессе обучения.

1. 1. 3 Классификация по алгоритмическим особенности методов распознавания

1. 1. 3. 1 Способы сравнения изображений

Практически все многообразие алгоритмов распознавания, ориентирована на выделение основных признаков и определение их наличия у тестового примера. можно выделить три группы методов, различающихся способами сравнения изображений.

В первой группе методов набор признаков (в простейшем случае исходное изображение) представляет собой точку в пространстве признаков, где значение каждого признака (например, яркость отдельного пикселя, значение главной компоненты, коэффициент частотного преобразования и тому подобные) представляет собой координату вдоль некоторой оси пространства признаков. Процедура сравнения основывается на разделении пространства признаков на области, относящиеся к одинаковым классам.

Для этого, например, может быть вычислено расстояние от неизвестного образа до всех остальных образов при помощи какой–либо метрики [3,10]. Класс может быть представлен центром кластера, тогда расстояние от неизвестного

образа вычисляется до центров всех кластеров, как это делается в методе главных компонент [3]. Линейный дискриминант строит набор линейных разделяющих поверхностей , полагая, что классы линейно разделимы в пространстве признаков [13]. Разновидность нейронных сетей – многослойные персептроны позволяют строить разделяющие поверхности любой сложности.

На рисунке 1 показано схема классификации способов сравнения изображений

Рисунок 1 – Схема классификации способов сравнения изображений

Несмотря на то, что самые совершенные методы могут лучшим способом разбить исходное пространство на области, этого недостаточно, поскольку для реальных объектов требуется огромное количество разделяющих областей (и обучающих примеров), чтобы учесть всевозможные способы изменения изображений объектов. Это связано с тем, что даже незначительное, с человеческой точки зрения, изменение изображения (например, ракурс, освещение или наличие бороды), может дать положение в пространстве признаков, очень далеко лежащее от исходного. И система в этом случае может среагировать не на одинакового человека, а, например, на одинаковый ракурс, посчитав изображение другого человека в том же ракурсе наиболее похожим на неизвестное. Ни применение различных преобразований для начального представления изображения, ни предобработка не может решить эту проблему для общего случая.

Следующая группа методов частично преодолевает данный недостаток. На изображении находятся важные области лица (например, области глаз, бровей,

носа, губ). Затем с помощью первого способа производится сравнение каждой области, результат каждого такого сравнения вносит вклад в окончательный результат. К таким методам относится например сравнение по эталонам, различные способы анализа главных компонент по блокам изображений.

В третьей группе учитывается топологическое искажения изображения. В методе сравнения эластичных графов на изображение накладывается набор ключевых точек, связанных между собой дугами. На неизвестном изображении находятся точки, соответствующие исходным и затем измеряется суммарное изменение размеров дуг. В гибких контурных моделях извлекается контур очертаний лица, затем сравниваются формы контуров для разных изображений. В других исследованиях используются нейронные сети и скрытые Марковские модели для анализа соотношений расстояний между областями глаз, носа и рта.

В вышеописанных методах после нахождения ключевых областей или контуров остальная информация не учитывается, что снижает точность.

Вышеприведённые методы сравнивают только суммарное искажение, не пытаясь учесть его характер, в этом и заключается их недостаток. Характер искажения изображения несёт важную информацию для распознавания изображения, потому нижеприводимые методы имеют лучшую точность распознавания.

К таким методам относятся псевдодвумерные скрытые Марковские модели, свёрточные нейронные сети, когнитроны и неокогнитроны.

В алгоритмах технологии FaceIta корпорации Visionics изображение разбивается на блоки и анализируется взаимное расположение таких блоков. Эти методы для сравнения используют как характеристики участков (блоков) изображений (набор яркостей пикселей, коэффициентов частотных преобразований, главные компоненты блоков изображения и тому подобное), так и взаимное расположение таких участков. Причём характер искажений, и содержимое участков усваивается в процессе обучения. В процессе распознавания искажение производится не к каждому примеру тренировочного набора, а к обобщённой модели класса.

1. 1. 3. 2 Использование обучающего набора примеров

По характеру использования обучающего набора методы распознавания лиц можно разделить на два больших класса. В методах первого класса в процессе настройки не используют обучающие примеры. В таких методах все параметры задаются вручную и, как правило, подбираются путём различных экспериментов. Например, криминалистическая идентификация человека по лицу с применением ключевых точек.

В противоположность этому второй класс методов в различной степени для извлечения признаков использует анализ обучающей выборки, представляющей собой набор типичных объектов нужных классов. Поэтому такие методы вычислительно более трудоёмки и их развитие стало возможным с возрастанием мощи вычислительной техники. Это метод главных компонент, линейные дискриминантные методы, метод сравнения эластичных графов. Однако в них основной упор сделан на автоматизированное извлечение признаков, а не отношений между ними и определения их важности (взвешивание). Существуют исследования, где была предпринята попытка вычислить весовые параметры для эластичного сравнения графов. Линейный дискриминантный анализ использовался для анализа признаков и преобразования их такому виду, максимизирующему межклассовые и минимизирующему внутриклассовые различия.

1. 1. 3. 3 Полнота использования информации изображения

Методы распознавания по изображению лица можно разделить на три класса.

Первый класс использует исходное изображение только для начального выделения каких–либо ключевых точек, расстояний между такими точками, соотношения расстояний или других геометрических характеристик и в дальнейшем к исходному изображению (например полутоновому) не обращается. Такая информация достаточно достоверно характеризует конкретного человека, поэтому используется в криминалистике для опознания людей. Однако показано, что в целом геометрической информации недостаточно для распознавания человека по изображению лица.

Второй класс методов использует всё исходное изображение. Например, метод анализа главных компонент, линейный дискриминантный анализ оперируют исходным полутоновым изображением как многомерным вектором. К этому классу так же относятся различные нейронные сети, принимающие на вход изображение целиком. Многие источники сходятся во мнении, что для общих случаев распознавания требуется анализировать всю информацию, предоставленную на изображении (например, всё полутоновое изображение).

Третий класс представляет собой сочетание двух предыдущих. Здесь используются и геометрические характеристики, и элементы, например, полутонового изображения. Например, метод сравнения эластичных графов. Вершины графа характеризуют особые точки изображения, дуги – расстояния между этими точками, кроме этого каждой вершине приписан набор локальных характеристик изображения (джет). Другим примером является метод сравнения эталонов. На изображении лица выделяются особые области, такие как глаза, нос, рот, и изображение сравнивается по этим областям.

1. 1. 3. 4 Аналитические и эмпирические методы

Методы распознавания характеризуются так же различным соотношением в них аналитических и эмпирических способов решения задачи распознавания. Метод главных компонент и линейные дискриминанты, например, основаны на решении систем матричных уравнений и нахождений расстояний между образами в различных метриках. В методе сравнения эластичных графов ключевые точки отмечаются на обучающем наборе изображений лиц исходя из эмпирических соображений.

Особое место занимают нейросетевые методы. В процессе настройки сети используются аналитические методы градиентного спуска. Но такая настройка не даёт ни гарантии достижения минимума ошибки, ни оценки отклонения полученной ошибки от оптимальной. В отличие от статистических методов, решения, получаемые при помощи нейронных сетей, носят не вероятностный, а правдоподобный характер, определяемый обучающим набором.

Похожая ситуация со скрытыми Марковскими моделями. Настройка параметров каждой модели осуществляется математическими оптимизационными методами, но они не дают точного решения.

1. 1. 3. 5 Использование обобщенной модели объекта

Методы делятся на два класса по способу сравнения неизвестного изображения с тренировочным набором (или базой изображений).

Первый класс методов использует тренировочный набор для построения обобщённой модели каждого класса лица (или всех классов сразу). Это различные нейронные сети, линейные дискриминанты, скрытые Марковские модели.

Второй класс методов непосредственно сравнивает неизвестное изображение с каждым изображением из тренировочной выборки. Это метод главных компонент, сравнение эталонов, методы сравнения эластичных графов, анализ оптического потока.

1. 1. 3. 6 Потребность в предобработке изображений

Многие методы распознавания человека по изображению лица требуют большого числа предварительных операций, таких как центрирование области лица, приведение к эталонному масштабу и ориентации, выравнивание гистограммы яркости пикселей, без которых они будут работать плохо. К таким методам относятся, например, метод главных компонент [3] анализ антропометрических точек лица [21].

Существуют так же методы, требующие минимум предобработки. Это свёрточные нейронные сети [18], неокогнитроны [25].

1. 1. 4 Классификация по используемым признакам объектов

В начале 2000–х годов выделилось две базовых группы методов сравнения лиц: «холистические подходы», (или «глобальные подходы») и «признаковые подходы», (или «структурные подходы») [2]. В первом случае, лица рассматриваются как цельные изображения, сравнивающиеся между собой. Во втором случае, из изображения лица выделяются локальные признаки, такие как информация об абсолютном или же взаимном расположении глаз, носа, рта и т.

п. Обе группы имеют свои недостатки: холистические подходы показали в целом большую надежность в близких к идеальным условиям распознавания по сравнению с признаковыми подходами, однако признаковые подходы лучше себя зарекомендовали в ситуации, когда отдельные части лица почему–либо не видны (например, закрыты элементами одежды или очками.

1. 1. 4. 1 Структурные подходы.

Структурные подходы к распознавания объектов основаны на автоматической подстройке к свойствам обрабатываемых данных, что позволяет разработать распознающую систему с приемлемыми характеристиками. Ко второй категории относятся методы, характеризующиеся сокращением размерности данных.

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

- Метод опорных векторов;
- Искусственные нейронные сети;
- Скрытые Марковские модели;
- Метод гибкого сравнения на графах.

Метод опорных векторов.

Метод опорных векторов (Support Vector Machines, SVM) [19] используется для решения задач классификации и регрессионного анализа. Применение SVM к решению задачи заключается в поиске гиперплоскости в признаковом пространстве, отделяющей классы изображений «лицо» и «не лицо».

Основная идея метода — перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве. Две параллельных гиперплоскости строятся по обеим сторонам гиперплоскости, разделяющей наши классы. Разделяющей

гиперплоскостью будет гиперплоскость, максимизирующая расстояние до двух параллельных гиперплоскостей. Алгоритм работает в предположении, что чем больше разница или расстояние между этими параллельными гиперплоскостями, тем меньше будет средняя ошибка классификатора.

Недостатки: существенные временные затраты при настройке, необходимость большого объема памяти.

Многослойные нейронные сети

Архитектура многослойной нейронной сети (МНС) состоит из последовательно соединенных слоев, где нейрон каждого слоя своими входами связан со всеми нейронами предыдущего слоя, а выходами – следующего [12,13]. НС с двумя решающими слоями может с любой точностью аппроксимировать любую многомерную функцию. НС с одним решающим слоем способна формировать линейные разделяющие поверхности, что сильно сужает круг задач, ими решаемых, в частности, такая сеть не сможет решить задачу типа "исключающее или". НС с нелинейной функцией активации и двумя решающими слоями позволяет формировать любые выпуклые области в пространстве решений, а с тремя решающими слоями – области любой сложности, в том числе и невыпуклой. При этом МНС не теряет своей обобщающей способности. Обучаются МНС при помощи алгоритма обратного распространения ошибки, являющегося методом градиентного спуска в пространстве весов с целью минимизации суммарной ошибки сети. При этом ошибки (точнее, величины коррекции весов) распространяются в обратном направлении от входов к выходам, сквозь веса, соединяющие нейроны. Простейшее применение однослойной НС (называемой автоассоциативной памятью) заключается в обучении сети восстанавливать подаваемые изображения. Подавая на вход тестовое изображение и вычисляя качество реконструированного изображения, можно оценить, насколько сеть распознала входное изображение. Положительные свойства этого метода заключаются в том, что сеть может восстанавливать искаженные и зашумленные изображения, но для более серьезных целей он не подходит. МНС также используется для

непосредственной классификации изображений – на вход подается или само изображение в каком–либо виде, или набор ранее извлеченных ключевых характеристик изображения, на выходе нейрон с максимальной активностью указывает принадлежность к распознанному классу.

Нейрон с максимальной активностью (здесь первый) указывает принадлежность к распознанному классу. Если эта активность ниже некоторого порога, то считается, что поданный образ не относится ни к одному из известных классов. Процесс обучения устанавливает соответствие подаваемых на вход образов с принадлежностью к определенному классу. Это называется "обучением с учителем". В применении к распознаванию человека по изображению лица такой подход хорош для задач контроля доступа небольшой группы лиц. Он обеспечивает непосредственное сравнение сетью самих образов, но с увеличением числа классов время обучения и работы сети возрастает экспоненциально, и потому для таких задач, как поиск похожего человека в большой базе данных, требует извлечения компактного набора ключевых характеристик, на их основе можно производить поиск.

В частности, МНС может использоваться для классификации изображений лиц на основе таких характеристик, как расстояния между специфическими частями лица (нос, рот, глаза). Также существуют гибридные системы, например объединение с марковской моделью. В классической МНС межслойные нейронные соединения полносвязны, и изображение представлено в виде одномерного вектора, хотя оно двумерно. Архитектура сверточной НС направлена на преодоление этих недостатков. В ней используются локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов), общие веса (обеспечивают детектирование некоторых черт в любом месте изображения) и иерархическая организация с пространственными подвыборками (spatial subsampling). Сверточная НС (СНС) обеспечивает частичную устойчивость к изменениям масштаба, смещениям, поворотам, искажениям. Архитектура СНС состоит из многих слоев, каждый из них имеет несколько плоскостей, причем нейроны следующего слоя связаны только с небольшим

числом нейронов предыдущего слоя из окрестности локальной области (как в зрительной коре человека). Веса в каждой точке одной плоскости одинаковы (сверточные слои). За сверточным слоем следует слой, уменьшающий его размерность путем локального усреднения. Затем опять сверточный слой, и так далее. Таким образом, достигается иерархическая организация. Более поздние слои извлекают более общие характеристики, меньше зависящие от искажений изображения. Обучается СНС стандартным методом обратного распространения ошибки. Сравнение МНС и СНС показало существенные преимущества последней как по скорости, так и по надежности классификации. Полезным свойством СНС также является то, что характеристики, формируемые на выходах верхних слоев иерархии, могут быть применимы для классификации по методу ближайшего соседа (например, вычисляя евклидово расстояние), причем СНС может успешно извлекать такие характеристики и для образов, отсутствующих в обучающем наборе. Для СНС характерны высокая скорость обучения и работы. Эта архитектуру является перспективной для дальнейших разработок в области распознавания изображений пространственных объектов. МНС применяются и для обнаружения объектов определенного типа. Помимо этого, любая обученная МНС в какой-то мере может определять принадлежность образов к "своим" классам, ее можно специально обучить надежному детектированию определенных классов. В этом случае выходными классами будут классы, принадлежащие и не принадлежащие к заданному типу образов.

Скрытые Марковские модели

Марковские модели являются мощным средством моделирования различных процессов и распознавания образов [13,24]. По своей природе Марковские модели позволяют учитывать непосредственно пространственно–временные характеристики сигналов, и потому получили широкое применение в распознавании речи, а в последнее время – изображений (в частности, изображений лиц). Каждая модель представляет собой набор N состояний между этими состояними возможны переходы. В каждый момент времени система находится в строго определенном состоянии. В наиболее распространенных

Марковских моделях первого порядка полагается, что следующее состояние зависит только от текущего состояния. При переходе в каждое состояние генерируется наблюдаемый символ, соответствующий физическому сигналу с выхода моделируемой системы. Выход, генерируемый моделью, может быть так же непрерывным. Существуют так же модели, с одинаковым набором символов для всех состояний.

Для того чтобы сократить вычисления, в распознавании речи используются линейные модели. В таких моделях каждое состояние имеет только одно последующее, так же переход возможен обратно в то же состояние. Такие модели учитывают временные характеристики речевого сигнала: определенный порядок следования участков сигнала, их взаимное расположение, возможность локальных растяжений или сжатий. Это позволяет их применять и в распознавании изображений.

Суть двумерных Марковских моделей заключается в том, что, в отличие от одномерных линейных СММ, они позволяют моделировать искажения изображения и взаимное расположение участков не отдельно по горизонтали или вертикали, а в обоих направлениях одновременно. Для уменьшения вычислительной сложности применяются псевдодвумерные СММ. Такая модель состоит из нескольких линейных вертикальных моделей нижнего уровня и одной линейной горизонтальной модели верхнего уровня, на её вход поступают выходы моделей нижнего уровня. Каждое состояние модели верхнего уровня включает в себя последовательность состояний соответствующей модели нижнего уровня. Модели нижнего уровня не связаны между собой. Изначально в модели верхнего уровня были вертикальными. В последующих работах модели верхнего уровня были сделаны горизонтальными, для того чтобы вертикальные модели нижнего уровня могли учесть тот факт, что глаза могут находиться на разной высоте. Таким образом, псевдодвумерная модель позволяет учесть локальные деформации, а также взаимное расположение участков изображений. Но в отличие от оптических потоков и других методов сопоставления деформациями, псевдодвумерная модель учитывает характер деформаций, а то,

какими именно могут быть возможные деформации, псевдодвумерные СММ усваивают в процессе обучения. Другими словами, участок, соответствующий глазу, никогда не будет сопоставлен, например, участку на месте рта.

Для СММ важное значение имеет начальная инициализация модели. В качестве начальной инициализации всех моделей используются все изображения из тренировочного набора. Затем модель каждого класса настраивается на свои изображения.

Недостатком СММ является то, что СММ не обладает различающей способностью, т. е. алгоритм обучения только максимизирует отклик каждой модели на свои классы, но не минимизирует отклик на другие классы и не выделяются ключевые признаки, отличающие один класс от другого. Таким образом, похожие классы могут оказаться слабо различимыми и при увеличении объема базы или использования в более широких условиях СММ может оказаться ненадежными.

Метод гибкого сравнения на графах

Суть метода гибкого сравнения на графах (Elastic graph matching) сводится к эластичному сопоставлению графов, описывающих изображения лиц [25,26]. В этом методе (Elastic Bunch Graph Matching) лицо представляется в виде графа, вершины этого графа расположены на ключевых точках лица таких, как контуры головы, губ, носы, и их крайних точках. Каждая грань помечена расстояниями между ее вершинами. В каждой такой точке вычисляются коэффициенты разложения по функциям Габора для пяти различных частот и восьми ориентаций. Набор таких коэффициентов $J=\{J_j\}$ называется джетом (jet). Джеты характеризуют локальные области изображений и служат для двух целей: во–первых, для нахождения точек соответствия в заданной области на двух различных изображениях; во–вторых – для сравнения двух соответствующих областей различных изображений. Каждый коэффициент $J_j=aj_{exp}(i\phi_j)$ для точек из одной области различных изображений характеризуется амплитудой a_j, медленно меняющаяся с изменением положения точки, и фазой ϕ_j, которая вращается со скоростью, пропорциональной частоте волнового вектора базисной

функции. Поэтому в простейшем случае для поиска на новом изображении точки с аналогичными характеристиками в функции подобия фазу не учитывают , этот случай описывается формулой (1)

$$S_a = \frac{\sum_j a_j a_j'}{\sqrt{\sum_j a_j^2 \sum_j a_j^{r2}}} \tag{1}$$

Функция подобия с одним джетом в фиксированной позиции и другим с переменной позицией является достаточно гладкой, для того чтобы получить быструю и надежную сходимость при поиске с применением простейших методов таких, как диффузия или градиентный спуск. Более совершенные функции подобия привлекают информацию о фазе. Для различных ракурсов соответствующие ключевые точки отмечаются вручную на тренировочном наборе. Кроме того, чтобы для одного и того же лица представить различные вариации его изображения в одном и том же графе, для каждой точки используются несколько джетов, соответствующих различным локальным характеристикам данной точки, например, открытому и закрытому глазу. Процесс распознавания неизвестного лица состоит в сравнении графа изображения лица GI со всеми остальными графами из набора B при помощи функции подобия, описываемой с помощью формулы (2):

$$S_b(G', B) = \frac{1}{N}\sum_n \max_m S_{\text{ф}}\left(J_n^I, J_n^{B_m}\right) - \frac{\lambda}{E}\sum_e \frac{(\Delta x_e^I - \Delta x_e^B)^2}{(\Delta x_e^B)^2} \tag{2}$$

Левая сумма характеризует подобие джетов, вычисленное с применением фазочувствительной функции, правая – топографическое соответствие, пропорциональное квадрату разности расстояний между соответствующими вершинами сравниваемых изображений, N – количество вершин, E – количество граней, λ – коэффициент относительной важности топографической информации.

В представленном выше виде метод способен достаточно надежно распознавать при изменениях ракурса до 20°; при больших углах точность распознавания резко уменьшается, функция подобия оказывается более

чувствительной к ракурсу, чем к межклассовым различиям. Дальнейшее развитие метода заключается в извлечении коэффициентов важности на основе анализа обучающей выборки. Для каждого джета симплекс–методом вычисляется коэффициент важности, затем использующийся впоследствие в функции подобия. Коэффициенты важности вычисляются из условия максимизации функции подобия для одного и того же лица и минимизации – для различных лиц. Существуют также более ранние разновидности этого метода, не использующие изначально определенные ключевые точки и структуры графа. Одни из них используют для сравнения решетки джетов, наложенные на изображение. В неизвестном изображении отыскиваются точки соответствия, затем по найденным точкам строится искаженная решетка и измеряется мера ее искажения для определения наиболее похожего изображения. В других методах точки извлечения джетов изначально образуют решетку, а затем наименее пригодные для распознавания точки отсеиваются в процессе обучения.

Недостатки: высокая вычислительная сложность процедуры распознавания, низкая технологичность при запоминании новых эталонов.

1. 1. 4. 2 Холистические методы

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

- Линейный дискриминантный анализ;
- Метод главных компонент.

Метод линейного дискриминантного анализа

Линейный дискриминантный анализ (линейный дискриминант Фишера) [13] использует такую проекцию пространства изображений на пространство признаков, минимизирующее внутриклассовое и максимизирующее межклассовое расстояние в пространстве признаков. В этих методах предполагается, что классы линейно разделимы.

Матрица W для проецирования пространства изображения на пространство признаков выбирается из следующего условия (3):

$$W_{opt} = \arg\max_{W} \frac{W^T S_B W}{W^T S_W W} \tag{3}$$

где S_B – матрица межклассовой дисперсии, S_W – матрица внутриклассовой дисперсии.

Может существовать до c−1 векторов составляющих базис пространства признаков, где c – общее число классов. С помощью этих векторов пространство изображений переводится в пространство признаков.

Поскольку работа непосредственно с матрицей $S_W \in R^{n \times n}$ затруднительна из–за ее размерности, используется предварительное уменьшение размерности с помощью метода главных компонент, затем вычисления производятся в пространстве меньшей размерности (4):

$$W_{fld} = \arg\max_{W} \frac{W^T W_{pca}^T S_B W_{pca} W}{W^T W_{pca}^T S_w W_{pca} W} \tag{4}$$

где W_{pca} – матрица для проецирования в пространство меньшей размерности (пространство главных компонент).

Обычно тренировочный набор содержит изображения лиц при нескольких базовых условиях освещенности, на основе этих изображений при помощи линейных комбинаций можно получить любые другие условия освещенности. Этот метод дает высокую точность распознавания (около 96%) для широкого диапазона условий освещенности, различных выражений лица и наличия или отсутствия очков. Однако остаются невыясненными вопросы, применим ли этот метод для поиска в больших базах данных, может ли метод работать, когда в

тренировочной выборке для некоторых лиц имеется изображение только в одних условиях освещенности. Для задачи детектирования лица с помощью LDA– классы лиц и "не лиц" разбивают на подклассы. Вышеописанный метод основывается на предположении о линейной разделимости классов в пространстве изображений. В общем случае такое предположение несправедливо. Инструмент для построения сложных разделяющих поверхностей предлагают нейросетевые методы

Метод главных компонент

Основной идеей применения метода главных компонент [3,4] для распознавания лиц является введение понятия «собственных лиц» (eigenfaces), представляющих изображения лиц в виде набора главных компонент изображений. Процесс распознавания лиц заключается в сравнении главных компонент неизвестного изображения с главными компонентами всех известных изображений. Основные преимущества применения МГК проявляются при хранении и поиске изображений в больших базах изображений, а также при реконструкции изображений. Недостаток МГК: высокие требования к условиям съемки изображений.

Метод главных компонент (Principal Component Analysis) применяется для сжатия информации без существенных потерь информативности. Он состоит в линейном ортогональном преобразовании входного вектора x размерности N в выходной вектор y размерности M, N>M. При этом компоненты вектора y являются некоррелированными и, следовательно, общая дисперсия после преобразования остается неизменной. Матрица X состоит из всех примеров изображений обучающего набора. Решив уравнение, описываемое формулой (5):

$$\Lambda = \Phi^T \sum \Phi, \tag{5}$$

получаем матрицу собственных векторов Φ, где Σ – ковариационная матрица для x, а Λ – диагональная матрица собственных чисел. Выбрав из Φ подматрицу Φ_M , соответствующую M наибольшим собственным числам, получим, что

преобразование из формулы (6):

$$y = \Psi_M^T \breve{X}, \tag{6}$$

где $\breve{X} = X - \bar{X}$ нормализованный вектор с нулевым математическим ожиданием, характеризует большую часть общей дисперсии и отражает наиболее существенные изменения x. Выбор первых M главных компонент разбивает векторное пространство на главное (собственное) пространство $F = \{\Phi_i\}_{i=1}^{M}$, содержащее главные компоненты, и его ортогональное дополнение $F = \{\Phi_i\}_{i=M+1}^{M}$. В качестве индикаторов принадлежности в методе главных компонент используют:

DIFS – distance in feature space, расстояние от образа анализируемого изображения в собственном пространстве, до эталонного образа;

DFFS – distance from feature space, расстояние от представления анализируемого изображения в пространстве наблюдения до проекции эталона в собственном пространстве.

Применение для задачи распознавания человека по изображению лица имеет следующий вид. Входные векторы представляют собой отцентрированные и приведенные к единому масштабу изображения лиц. Собственные векторы, вычисленные для всего набора изображений лиц, называются собственными лицами (eigenfaces). Метод главных компонент в применении к изображениям лиц также называют методом собственных лиц. С помощью вычисленных ранее матриц входное изображение разлагается на набор линейных коэффициентов, называемых главными компонентами. Сумма N первых главных компонент, умноженных на соответствующие собственные векторы, является аппроксимацией изображения порядка N.

Для каждого изображения лица вычисляются его главные компоненты. Обычно берется от 5 до 200 главных компонент. Остальные компоненты кодируют мелкие различия между лицами и шум. Процесс распознавания заключается в сравнении главных компонент неизвестного изображения с компонентами всех остальных изображений. Для этого обычно применяют

какую–либо метрику (простейший случай – Евклидово расстояние). Дополнительное повышение надежности достигается за счет дополнительного применения анализа главных компонент к отдельным участкам лица таким, как глаза, нос, рот.

Кроме того метод главных компонент применяется для обнаружения лица на изображении. Для лиц значения компонент в собственном пространстве имеют большие значения, а в дополнении собственного пространства – близки к нулю. По этому факту можно обнаружить, является ли входное изображение лицом. Для этого проверяется величина ошибки реконструкции; чем больше ошибка, тем больше вероятность, что это не лицо. При наличии в наборе изображений лиц вариаций таких, как раса, пол, эмоции, освещение, будут появляться компоненты, их величина в основном определяется этими факторами. Поэтому по значениям соответствующих главных компонент можно определить, например, расу или пол человека. Основные недостатки PCA таковы. Метод главных компонент требует для своего применения идеализированных условий таких, как единые параметры освещенности, нейтральное выражение лица, отсутствие помех вроде очков и бород. При несоблюдении этих условий главные компоненты не будут отражать межклассовые вариации. Например, при различных условиях освещенности метод собственных лиц практически неприменим, поскольку первые главные компоненты преимущественно отражают изменения освещения, и сравнение выдает изображения, имеющие похожий уровень освещенности.

Вычисление набора собственных векторов отличается высокой трудоемкостью. Один из способов – это свертка изображений по строкам и столбцам; в такой форме представление изображения имеет на порядок меньший размер, вычисления и распознавание происходит быстрее, но восстановить исходное изображение уже невозможно.

1. 1. 4. 3 Кластерные подходы

Суть кластерных подходов для сравнения лиц заключается в разделении изображения лица на некоторое количество участков, к каждому из этих

участков можно применить холистический подход и сравнивать их попарно, то есть независимо от других участков. В таком случае можно надеяться на то, что если некоторые пары участков сравниваемых изображений не будет совпадать между собой в силу вышеупомянутых мешающих факторов вроде теней и перекрытий, оставшиеся пары участков обеспечат нужную степень взаимной корреляции, которой будет достаточно для правильного итогового решения совпадения/несовпадения лиц.

Метод локальных бинарных шаблонов

Существует метод описания 3D модели с помощью локальных бинарных шаблонов. Этот метод описывает локальный текстурный паттерн с помощью бинарного кода. Он построен на пороговой классификации некоторой окрестности точки некоторого радиуса (обычно рассматривают 8 соседних пикселей). То есть если значения яркости (в черно–белого изображении) соседнего пикселя больше, чем в рассматриваемом пикселе, то он кодируется единицей, иначе нулём. Так каждому пикселю текстуры соответствует набор из 8 бит (количество соседей). Потом для каждой области стоится гистограмма подобных наборов и таким образом описывается лицо. В случае трехмерного образца этот алгоритм немного изменяется, чтобы описывать внутреннюю геометрическую информацию, снижая эффект от изменения выражения лица эффективно. На стадии обучения спектральная регрессия применяется для обучения главных компонент для каждого трехмерного –изображения. Изначально этот метод применялся для сравнения трехмерных образцов, однако существует применения локальных бинарных шаблонов и для сравнения объектов имеющих различное количество измерений распознавания.

Способы построения выборки

Для оценки эффективности распознавания и идентификации лиц на изображении необходимо общий набор изображений разделить на два подмножества.

Одно из этих подмножеств станет обучающим набором, то есть теми изображениями, принадлежность лиц, присутствующих на котором заранее однозначно определена и соответствует тому или иному человеку. Второе подмножество (оставшиеся изображения из общего набора) будет использовано в качестве тестовой выборки. Принадлежность лиц, присутствующих на них будет определяться исследуемым алгоритмом. Соотношение верных и ошибочных идентификаций в итоге станет основным критерием оценки эффективности метода. Чем больше окажется значение этого параметра, тем эффективнее данный метод справляется с поставленной задачей идентификации человека по изображению лица.

1. 2. 1 Перекрестная проверка

Кросс–валидация, иногда называющаяся перекрестной проверкой, – это техника валидации модели для проверки того, насколько успешно применяемый в модели статистический анализ способен работать на независимом наборе данных.

Обычно кросс–валидация используется в ситуациях, где целью является предсказание, требуется оценить, насколько предсказывающая модель способна работать на практике.

Перекрестная проверка – это способ оценить способность модели работать на гипотетическом тестовом наборе, когда такой набор в явном виде получить невозможно.

В этом случае исходный набор данных разбивается на K одинаковых по размеру блока. Из K блоков один оставляется для тестирования модели, а остающиеся K–1 блока используются как обучающий набор. Процесс повторяется K раз, каждый из блоков используется один раз как тестовый набор.

По завершении, получается K результатов, по одному на каждый блок, они усредняются или комбинируются каким–либо другим способом, и дают одну оценку.

Основное преимущество такого способа перед случайным выбором заключается в том, что все наблюдения используются и для тренировки, и для

тестирования модели, и каждое наблюдение используется для тестирования в точности один раз.

На рисунке 2 показана схема формирования обучающего и тестового наборов для перекрестной проверки.

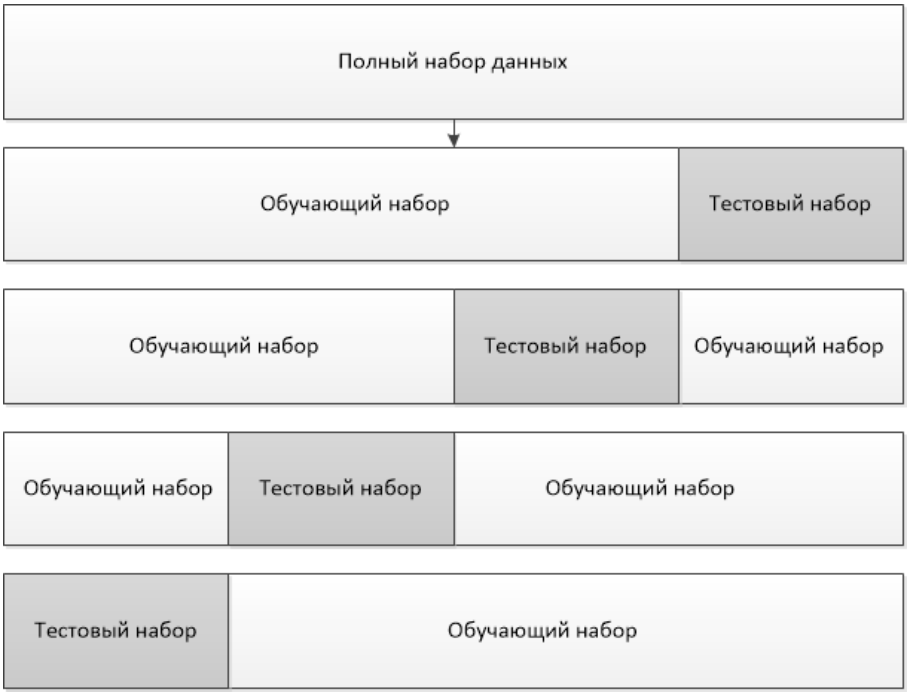

Рисунок 2 – Схема формирования обучающего и тестового наборов для перекрестной проверки

1. 2. 1. 1 Валидация последовательным случайным выбором.

Этот метод случайным образом разбивает набор данных на обучающий и тестовый наборы. Для каждого такого разбиения, модель подгоняется под тренировочные данные, а точность предсказания оценивается на тестовом наборе. Результаты затем усредняются по всем разбиениям. Преимущество такого метода перед кросс–валидацией на K блоках в том, что пропорции тренировочного и тестового наборов не зависят от числа повторений (блоков). Недостаток метода в том, что некоторые наблюдения могут ни разу не попасть в тестовый набор, тогда как другие могут попасть в него более, чем один раз. Другими словами, тестовые наборы могут перекрываться. Кроме того, поскольку разбиения проводятся случайно, результаты будут отличаться в случае повторного анализа.

В послойном варианте этого метода, случайные выборки генерируются таким способом, что средний ответ модели равен по тренировочному и тестовому наборам. Это особенно полезно, когда ответ модели бинарен, с неравными пропорциями ответов по данным.

1. 2. 1. 2 Поэлементная кросс–валидация

Здесь отдельное наблюдение используется в качестве тестового набора данных, а остальные наблюдения из исходного набора – в качестве тренировочного. Цикл повторяется, пока каждое наблюдение не будет использовано один раз в качестве тестового. Это то же самое, что и K–блочная кросс–валидация, где K равно числу наблюдений в исходном наборе данных.

Разделение общего набора на обучающую и тестовую выборку может основываться на различных подходах к данному вопросу.

Для формирования выборки могут применяться вероятностные и детерминированные методы.

1. 2. 2 Вероятностные способы

Вероятностные методы включают в свой состав: простой случайный отбор, систематический отбор, кластерный отбор и стратифицированный отбор.

1. 2. 2. 1 Простой случайный отбор

Простой случайный (собственно–случайный) отбор предполагает, что вероятность быть избранным в выборку известна и является одинаковой для всех единиц совокупности. Вероятность быть включенным в выборку определяется отношением объема выборки к размеру совокупности.

Простой случайный отбор может осуществляться с помощью следующих методов: формирование выборки вслепую или с помощью таблицы случайных чисел.

При использовании метода формирования выборки вслепую единицы совокупности в из данной общего набора случайным образом выбирается некоторое количество изображений, определяемое объемом выборки. Для определения идентификатора случайного изображения может использоваться любой программный генератор случайных (псевдослучайных) чисел.

1. 2. 2. 2 Систематический отбор

Начальная часть метода систематического (механического) отбора соответствует начальной части метода простого случайного отбора: необходимо получить полный список единиц генеральной совокупности.

Однако далее вместо присвоения идентификаторов используется показатель «интервал скачка», рассчитанный как отношение размера совокупности к объему выборки.

Особенно широко метод систематического отбора используется, когда для различных видов совокупностей имеются различные справочники, списки и т.д.

1. 2. 2. 3 Кластерный отбор

Другим методом вероятностного отбора является кластерный отбор, основанный на делении совокупности на подгруппы. Базовая концепция данного метода очень похожа на базовую концепцию метода систематического отбора, однако реализация этой концепции осуществляется по–другому. Генеральная совокупность разбивается на четко определяемые части (кластеры).. Далее один из кластеров выбирается случайным образом, определяется совокупность для этой части, в ней проводится соответствующее исследование, а выводы обобщаются на совокупность всего кластера.

В основе всех описанных методов лежит предположение, что любая совокупность характеризуется симметричным распределением ее ключевых характеристик. Говоря другими словами, каждая выборка достаточно полно характеризует всю совокупность, различные крайности в выборке уравновешивают друг друга. Такая ситуация на практике встречается крайне редко.

В случае несимметричного распределения совокупности последняя разделяется на различные подгруппы (страты), по некоторым независимым параметрам, и выборки формируются из этих подгрупп. Такой метод носит название стратифицированного отбора.

Далее для каждой страты с помощью случайного отбора формируется выборка.

1. 2. 3 Детерминированные способы

При применении детерминированных методов отбора формирование выборки осуществляется без использования понятий теории вероятностей, вследствие чего невозможно рассчитать вероятность включения в выборку единицы совокупности.

Охарактеризуем следующие детерминированных методы отбора: отбор на основе принципа удобства, отбор на основе суждений, формирование выборки в процессе обследования и формирование выборки на основе квот.

1. 2. 3. 1 Отбор на основе принципа удобства

Смысл метода отбора на основе принципа удобства заключается в том, что формирование выборки осуществляется самым удобным с позиций исследователя образом, например, с позиций минимальных затрат времени и усилий, с позиции доступности образцов.

1. 2. 3. 2 Отбор на основе суждения

Формирование выборки на основе суждения основано на использовании мнения квалифицированных экспертов относительно состава выборки.

1. 2. 3. 3 Отбор на основе квот

Формирование выборки на основе квот (квотный отбор) предполагает предварительное, исходя из целей исследования, определение численности групп образцов, соответствующих определенным критериям.

Решение об объеме выборки является компромиссом между теоретическими предположениями о точности результатов обследования и возможностями их практической реализации, прежде всего имеются ввиду затраты на проведение исследования.

Глава 2. Теоретический базис: Исследуемые методы распознавания лиц

2. 1 Метод главных компонент

Основной идеей применения метода главных компонент [1,3] (Principal Component Analysis, PCA) для распознавания лиц является введение понятия «собственных лиц» (eigenfaces), представляющих изображения лиц в виде набора главных компонент изображений. Процесс распознавания лиц заключается в сравнении главных компонент неизвестного изображения с главными компонентами всех известных изображений. Основные преимущества применения PCA проявляются при хранении и поиске изображений в больших базах изображений, а также при реконструкции изображений. Недостаток PCA: высокие требования к условиям съемки изображений.

Метод главных компонент применяется для сжатия информации без существенных потерь информативности. Он состоит в линейном ортогональном преобразовании входного вектора x размерности N в выходной вектор y размерности M, N>M. При этом компоненты вектора y являются некоррелированными и, следовательно, общая дисперсия после преобразования остается неизменной. Матрица X состоит из всех примеров изображений обучающего набора. Решив уравнение $\Lambda = \Phi^T \sum \Phi$, получаем матрицу собственных векторов Φ, где Σ – ковариационная матрица для x, а Λ – диагональная матрица собственных чисел. Выбрав из Φ подматрицу Φ_M , соответствующую M наибольшим собственным числам, получим, что преобразование $y = \Psi_M^T \check{X}$, где $\check{X} = X - \bar{X}$ нормализованный вектор с нулевым математическим ожиданием, характеризует большую часть общей дисперсии и отражает наиболее существенные изменения x. Выбор первых M главных компонент разбивает векторное пространство на главное (собственное) пространство $F = \{\Phi_i\}_{i=1}^{M}$, содержащее главные компоненты, и его ортогональное дополнение $F = \{\Phi_i\}_{i=M+1}^{M}$. В качестве индикаторов принадлежности в методе главных компонент используют:

DIFS – distance in feature space, расстояние от образа анализируемого изображения в собственном пространстве, до эталонного образа;

DFFS – distance from feature space, расстояние от представления анализируемого изображения в пространстве наблюдения до проекции эталона в собственном пространстве.

Применение для задачи распознавания человека по изображению лица имеет следующий вид. Входные векторы представляют собой отцентрированные и приведенные к единому масштабу изображения лиц. Собственные векторы, вычисленные для всего набора изображений лиц, называются собственными лицами (eigenfaces). Метод главных компонент в применении к изображениям лиц также называют методом собственных лиц. С помощью вычисленных ранее матриц входное изображение разлагается на набор линейных коэффициентов, называемых главными компонентами. Сумма N первых главных компонент, умноженных на соответствующие собственные векторы, является аппроксимацией изображения порядка N.

Для каждого изображения лица вычисляются его главные компоненты. Обычно берется от 5 до 200 главных компонент. Остальные компоненты кодируют мелкие различия между лицами и шум. Процесс распознавания заключается в сравнении главных компонент неизвестного изображения с компонентами всех остальных изображений. Для этого обычно применяют какую–либо метрику (простейший случай – Евклидово расстояние). Дополнительное повышение надежности достигается за счет дополнительного применения анализа главных компонент к отдельным участкам лица таким, как глаза, нос, рот.

Также метод главных компонент применяется для обнаружения лица на изображении. Для лиц значения компонент в собственном пространстве имеют большие значения, а в дополнении собственного пространства – близки к нулю. По этому факту можно обнаружить, является ли входное изображение лицом. Для этого проверяется величина ошибки реконструкции; чем больше ошибка, тем больше вероятность, что это не лицо. При наличии в наборе изображений

лиц вариаций таких, как раса, пол, эмоции, освещение, будут появляться компоненты, их величина в основном определяется этими факторами. Поэтому по значениям соответствующих главных компонент можно определить, например, расу или пол человека. Основные недостатки PCA таковы. Метод главных компонент требует для своего применения идеализированных условий таких, как единые параметры освещенности, нейтральное выражение лица, отсутствие помех вроде очков и бород. При несоблюдении этих условий главные компоненты не будут отражать межклассовые вариации. Например, при различных условиях освещенности метод собственных лиц практически неприменим, поскольку первые главные компоненты преимущественно отражают изменения освещения, и сравнение выдает изображения, имеющие похожий уровень освещенности. На рисунке 3 показаны реконструции первых 20 собственных лиц.

Рисунок 3 – Пример изображений собственных векторов (собственные лица)

Вычисление набора собственных векторов отличается высокой трудоемкостью. Один из способов – это свертка изображений по строкам и столбцам; в такой форме представление изображения имеет на порядок меньший размер, вычисления и распознавание происходит быстрее, но восстановить исходное изображение уже невозможно.

Проблема представления изображения заключается в его высокой размерности. Двумерное p*q изображение в оттенках серого отображается в векторное пространство m = pq, таким образом, изображение 100x100 точек уже лежит в 10000–мерном пространстве. Вопрос заключается в том, все ли измерения для нас одинаково полезны? Мы можем лишь делать выводы о том, имеется ли какой–либо разброс в данных, таким образом, ищутся это компоненты, содержащие большую часть информации. Анализ главных компонент был предложен независимо Карлом Пирсоном в 1901 году и Гарольдом Хотелингом, для того чтобы превратить набор вероятностно коррелированных переменных в меньший набор некоррелированных переменных. Идея состоит в том, что многоразмерный набор данных зачастую описывается переменными коррелирующими между собой. И поэтому куда меньшее количество более значимых переменных содержит основную часть информации. Метод анализа главных компонент позволяет выбрать основное в огромном многообразии данных.

Алгоритм работы метода главных компонент.

Пусть X={x_1,x_2,…,x_n} случайный вектор , в котором

1. Вычислить среднее μ по формуле (7):

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{7}$$

2. Вычислить ковариационную матрицу S, рассчитывающуюся по формуле (8):

$$S = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)(x_i - \mu)^{\tau_i} \tag{8}$$

3. Вычислить собственные значения λ_i и собственные вектора ν_i из S, определенного по формуле (9)

$$S_{\nu_i} = \lambda_i \nu_{i,} \tag{9}$$

где i=1,2,…,n

4. Упорядочить собственные вектора по убыванию их собственных значений. к главных компонент это собственные вектора, соответствующие самым большим к–собственных значениям. К–главных компонент рассматриваемого вектора заданы равенством (10):

$$y = W^{\tau}(x - \mu), \tag{10}$$

Восстановленный с помощью анализа главных компонент базис представлен равенством (11):

$$x = W_y + \mu \tag{11}$$

Метод главных компонент выполняет распознавание лиц с помощью:

- Проецирования всех обучающих примеров в подпространство главных компонент;
- Проецирования обрабатываемого изображения в подпространство главных компонент;
- Нахождение ближайшего соседа для обрабатываемого изображения среди обучающих примеров.

Остается еще одна нерешенная проблема. Предположим, дано 400 изображений, размерностью 100 на 100 пикселей. Методом Главных Компонент получаем ковариационную матрицу S = XX^{τ}, где размер получается равным 10000 x 400. Решением будет матрица 10000 x 10000, примерно 0,8 гигабайт, это практически невыполнимо. Поэтому, используется следующий вариант. Матрица M x N, при M > N, может иметь только N – 1 ненулевых собственных значений, поэтому можно взять декомпозицию S = $X^{\tau}X$, размерностью N x N описанную формулой (12):

$$X^{\tau}X_{\nu_i} = \lambda_i \nu_i, \tag{12}$$

и получить исходные собственные вектора S =XX^{τ}, умноженые слева на матрицу данных: $\mathrm{X}X^{\tau}\left(X_{v_i}\right) = \lambda_i\left(X_{v_i}\right)$, полученные собственные вектора ортогональны, чтобы получить ортонормальные вектора они должны быть нормализованы.

2. 2 Метод локальных бинарных шаблонов

Локальный бинарный шаблон – это определенный вид признака, используемый для классификации в компьютерном зрении, и представляющий собой простой оператор. Локальные бинарные шаблоны впервые были предложены в 1996 году для анализа текстуры полутоновых изображений. При этом дальнейшие исследования показали, что ЛБШ инвариантны к небольшим изменениям в условиях освещения и небольшим поворотам изображения.

ЛБШ представляет собой описание окрестности пикселя изображения в двоичном представлении. Базовый оператор ЛБШ, применяемый к пикселю изображения, использует восемь пикселей окрестности, принимая значение интенсивности центрального пикселя в качестве порога (рисунок 4). Пиксели со значением интенсивности большим или равным значению интенсивности центрального пикселя принимают значения равные «1», остальные принимают значения равные «0». Таким образом, результатом применения базового оператора ЛБШ к пикселю изображения является восьмиразрядный бинарный код, описывающий окрестность этого пикселя.

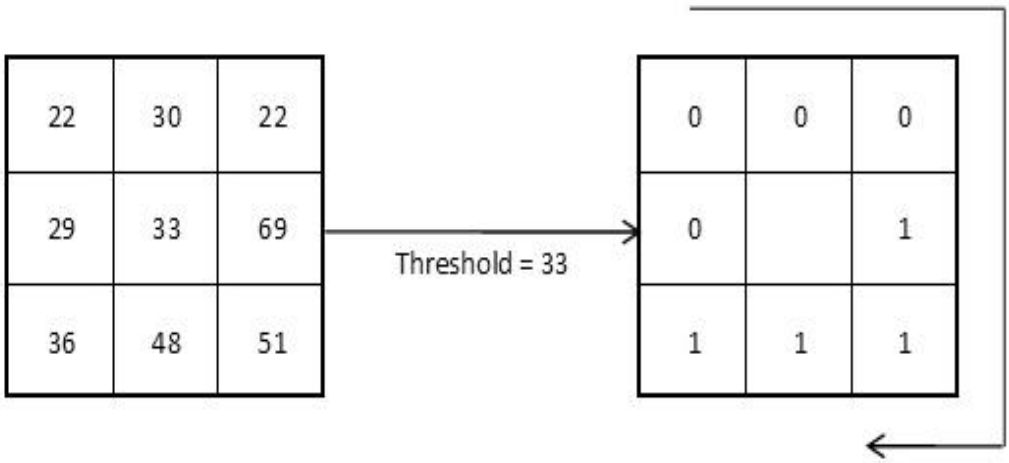

Рисунок 4 – Схема базового оператора ЛБШ

Использование круговой окрестности и билинейной интерполяции значений интенсивностей пикселей позволяет построить локальный бинарный шаблон с произвольным количеством точек P и радиусом R .

Некоторые бинарные коды несут в себе больше информации, чем остальные. Так, локальный бинарный шаблон называется равномерным, если он содержит не более трех серий «0» и «1» (например, 00000000, 001110000 и 11100001). Во–первых, равномерные ЛБШ определяют только важные локальные особенности изображения, такие как концы линий, грани, углы и пятна (рисунок 5). Во–вторых, они обеспечивают существенную экономию памяти ($P(P–1)+2$ различных шаблонов вместо $2P$).

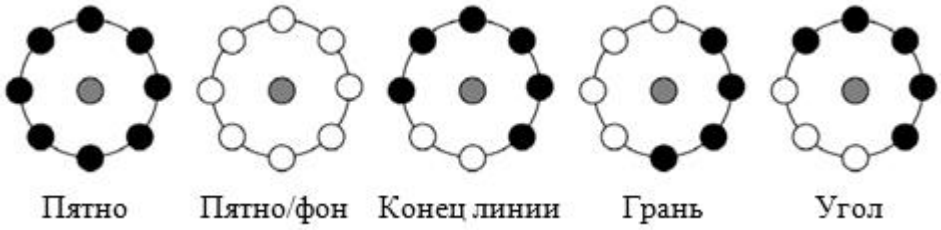

Рисунок 5 – Операторы, соответствующие некоторым графическим примитивам

Применяя оператор ЛБШ к каждому пикселю изображения, мы можем построить гистограмму, где каждому равномерному коду ЛБШ соответствует отдельный столбец. Также имеется еще один дополнительный столбец, содержащий информацию обо всех неравномерных шаблонах.

Алгоритм работы метода локальных бинарных шаблонов

Формальное описание ЛБШ оператора выглядит, как показано в формуле (13):

$$\mathrm{LBP}(x_c, y_c) = \sum_{p=0}^{p-1} 2^p \, S\left(i_p - i_c\right) \tag{13}$$

где $\left(i_p - i_c\right)$ интенсивность центрального пикселя; а i_n интенсивность соседнего пикселя. S– знаковая функция определенная равенством (14) :

$$s(x) = \begin{cases} 1 \text{ if x } \geq 0 \\ 0 \quad \text{else} \end{cases} \tag{14}$$

Это описание позволило обозначить 5 различных видов детали на изображении. Вскоре после публикации было замечено, что при фиксированных соседях

распознавание деталей в масштабе не удается. По этому стали использовать переменное число соседей. Идея в том, чтобы привязывать произвольное количество соседей по кругу с переменным радиусом, что позволило улавливать следующие соседства.

Для данной точки положение соседей может быть по следующей формулам (15) и (16):

$$x_p = x_c + R\cos\left(\frac{2\pi p}{p}\right) \tag{15}$$

$$y_p = y_c - R\,sin(\frac{2\pi p}{p}) \tag{16}$$

где R радиус круга и p количество точек образца.

«Оператор» – это расширение изначального ЛБШ кода, Поэтому его иногда называют расширенный ЛБШ (также называется круговой ЛБШ). Если координаты точки не соответствуют координатам изображения, точка интерполируется. В компьютерной науке существует множество «умных» интерполяционных схем, в OpenCV применяется билинейная интерполяция, определяющаяся формулой (17):

$$\text{f(x,y)} \approx [1 - x\ x]\begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1 \end{bmatrix}\begin{bmatrix} 1-y \\ y \end{bmatrix} \tag{17}$$

Остается добавить пространственную информацию в модель распознавания лица. В представление Ахонена предлагается разделить ЛБШ изображение на m участков и извлекать гистограмму для каждого. Векторы пространственных черт получаются путем объединения локальных гистограмм (не сливая их). Эти гистограммы называются гистограммы локальных бинарных шаблонов.

Изображения лиц могут рассматриваться как набор всевозможных локальных особенностей, они хорошо описываются с помощью локальных бинарных шаблонов. Однако гистограмма, построенная для всего изображения в целом, кодирует лишь наличие тех или иных локальных особенностей, но при этом не содержит никакой информации об их расположении на изображении. Для учета такого рода информации изображение разбивается на подобласти, в каждой из

которых вычисляется своя гистограмма ЛБШ (рисунок 6). Путем конкатенации этих гистограмм может быть получена общая гистограмма, учитывающая как локальные, так и глобальные особенности изображения.

Рисунок 6 – Разделение изображения на подобласти, для вычисления отдельной гистограммы ЛБШ на различных участках

Глава 3. Экспериментальное исследование на основе метода главных компонент и метода локальных бинарных шаблонов.

3. 1 Исходные данные: базы изображений лиц

Исходный набор изображений был сформирован из двух свободно–распространяемых баз изображений лиц ORL и FERET.

3. 1. 1 База ORL

База изображений лиц, подготовленная в научно–исследовательской лаборатории компании Olivetti (Olivetti Research Laboratory или ORL; после приобретения корпорацией AT&T в 1999 г. и до закрытия в 2002 г. лаборатория носила название AT&T Laboratories Cambridge)[35]. В базу входят фронтальные изображения лиц. База содержит 400 фронтальных изображений 40 человек. Все изображения в базе полутоновые, с 256 градациями яркости. Размер каждого изображения — 92×112 пикселей. Файлы представлены в формате PGM. На рисунке 7 показаны примеры исходных из изображений базы ORL.

Рисунок 7 – Примеры изображений базы ORL

3. 1. 2 База FERET

База изображений лиц FERET это стандартный набор, используемый для оценки систем распознавания лиц. Технология распознавания лиц – FERET – программа под управлением агентство передовых оборонных исследовательских проектов DARPA и Национального института стандартов и технологий США [36]. Данная база изображений собиралась с декабря 1993 по август 1996 года В 2003 году DARPA выпустила базу в высоком разрешении и 24–битной цветности. Размер изображений лиц базы данных FERET равен 512×768

пикселей. Примеры исходных изображений из базы FERET показаны на рисунке 8.

Рисунок 8 – Примеры изображений базы FERET

3. 2 Формирование выборки

Поскольку цель данной работы исследование эффективности методов распознавания и идентификации лиц. Весь набор изображений был обработан с помощью каскадов Хаара, обнаруженные каскадом области изображений, выделенные как изображение лица и стали основой для дальнейшего исследования. Таким образом, для обработки предлагались лишь те части изображений, где машина обнаружила лицо, и эффективность методов обнаружения лица не оказала влияния на результаты исследования

Рисунки 9 и 10 показывают, какая область исходных изображений выделяется в качестве содержащей изображение лица

Рисунок 9 – Результат выделения области лица на изображениях базы ORL

Рисунок 10 – Результат выделения области лица на изображениях базы FERET

После процедуры выделения лиц в итоговых наборах оказалось 515 классов, соответствующих выбранным критериям. 477 из них в базе FERET и 38 в базе ORL. Всего было выделено 4705 изображений лиц (4366 принадлежали классам базы FERET, 339 – ORL). В каждом классе из базы ORL получилось от 5 до 10 (3 – 8 в обучающей выборке) изображений, а в классах из базы FERET от 4 до 36 (2–28 для обучающей выборки).

3. 2. 1 Создание обучающих и тестовых наборов

После того как определены количественные характеристики для всех классов обеих выборок, необходимо решить вопрос об оптимальном выборе того, как разделять наборы изображений каждого отдельного класса на тестовый и обучающий для проведения перекрестной проверки в нескольких вариантах. Такой подход позволяет решить две задачи. Во–первых, сама по себе перекрестная проверка позволяет избежать так называемой ошибки «третьего рода» – связанной с параметрами, присутствующими в исходном наборе данных, учитывающихся классификаторами, такими как освещение и геометрические искажения изображений лица. Во–вторых, использование различных пропорций соотношения тестового и обучающего набора позволяет рассмотреть варианты, где количество изображений предъявляемых для обучения классификатора будет различным для одного и того же класса, но при этом сохранятся преимущества перекрестной проверки и, практически каждое изображение попадет в тестовую выборку, но будет отсутствовать в обучающем наборе.

В конечном итоге, выбраны варианты разбиения выборки каждого класса на 2, 3, 4 и 5 частей. Выбор минимального количества очевиден, так как одна часть необходима для тестового набора, а вторая для обучающего. Максимальное рассматриваемое число – пять, обусловлено, с одной стороны тем, что большее количество частей требовало несравненно больше временных и машинных ресурсов для проведения соответствующей обработки всеми классификаторами со всеми выбранными вариантами параметров. С другой стороны, характеристики исходного набора данных не позволяет включить в обработку более тридцати процентов всех классов, количество изображений внутри

которых было менее 7, что негативно может сказаться на репрезентативности результатов.

3. 2. 1. 1 Создание тестового и обучающего набора для отдельного класса изображений

К каждому классу изображений лиц применены выше описанные виды разбиения общего набора, в тестовый набор, с помощью программного генератора случайных чисел выбраны изображения таким образом, что каждое изображение попадает в тестовый набор не более одного раза для данного варианта разбиения. В обучающий набор соответственно попадают все оставшиеся изображения класса. Схема алгоритма формирования тестового и обучающего набора для каждого класса в исходном наборе показана на рисунке А. 1

3. 2. 1. 2 Создание общей выборки для обучения классификаторов.

Для каждого варианта собирается общая для всех классов обучающая выборка, путем конкатенации всех обучающих наборов изображений каждого класса, попавших в число классов, имеющих достаточное количество изображений для данного варианта разбиения общего количества изображений в классе на выбранное количество частей. Аналогичным образом собирается тестовая выборка для текущего варианта разбиения, то есть в тестовый набор попадают все изображения каждого класса не попавшие в обучающий набор.

На рисунке А.2 показан алгоритм формирования общего тестового и обучающего набора для всех классов.

3. 2. 2 Наборы, сформированные для обучения и проверки классификаторов

После проведения этих процедур были сформированы наборы для обработки с помощью исследуемых методов распознавания, в таблице 1 показано, как распределились по выборкам изображения, в зависимости от того, на какое количество частей делился исходный набор изображений.

Таблица 1 – Объем тестового и обучающего набора при разбиении исходного набора изображений класса на различное количество частей

Доля изображений класса выбираемых для составления тестового набора.	1/2	1/3	1/4	1/5
Количество изображений, составляющих один обучающий набор				
Из базы FERET	2306	3083	3418	3438
Из базы ORL	174	239	270	280
Итого	2480	3322	3688	3718
Количество изображений, составляющих один тестовый набор				
Из базы FERET	2060	1283	900	704
Из базы ORL	165	100	69	59
Итого	2225	1383	969	763

3. 3 Программное обеспечение исследования

3. 3. 1 Программная платформа

Программной платформой для разработки приложения стала платформа . NET Framework — программная платформа, выпущенная компанией Microsoft в 2002 году. Основой платформы является общеязыковая среда исполнения Common Language Runtime (CLR), подходящая для разных языков программирования. Функциональные возможности CLR доступны в любых языках программирования, использующих эту среду.

Хотя . NET является патентованной технологией корпорации Microsoft и официально рассчитана на работу под операционными системами семейства Microsoft Windows, существуют независимые проекты (прежде всего это Mono и Portable. NET), позволяющие запускать программы . NET на некоторых других операционных системах.

3. 3. 2 Среда разработки

Средой разработки приложения стала Microsoft Visual Studio [32].

Линейка продуктов компании Microsoft, включающих интегрированную среду разработки программного обеспечения и ряд других инструментальных средств. Данные продукты позволяют разрабатывать как консольные приложения, так и приложения с графическим интерфейсом, в том числе с поддержкой технологии Windows Forms, а также веб–сайты, веб–приложения, веб–службы как в родном, так и в управляемом кодах для всех платформ, поддерживаемых Windows, Windows Mobile, Windows CE, . NET Framework, Xbox, Windows Phone . NET Compact Framework и Silverlight.

Visual Studio включает в себя редактор исходного кода с поддержкой технологии IntelliSense и возможностью простейшего рефакторинга кода. Встроенный отладчик может работать как отладчик уровня исходного кода, так и как отладчик машинного уровня. Остальные встраиваемые инструменты включают в себя редактор форм для упрощения создания графического интерфейса приложения, веб–редактор, дизайнер классов и дизайнер схемы базы данных. Visual Studio позволяет создавать и подключать сторонние дополнения (плагины) для расширения функциональности практически на каждом уровне, включая добавление поддержки систем контроля версий исходного кода (как, например, Subversion и Visual SourceSafe), добавление новых наборов инструментов (например, для редактирования и визуального проектирования кода на предметно–ориентированных языках программирования) или инструментов для прочих аспектов процесса разработки программного обеспечения (например, клиент Team Explorer для работы с Team Foundation Server).

3. 3. 3 Язык программирования

В качестве языка программирования был выбран C Sharp [33].

Объектно–ориентированный язык программирования. Разработан в 1998—2001 годах группой инженеров под руководством Андерса Хейлсберга в компании Microsoft как язык разработки приложений для платформы Microsoft . NET Framework и впоследствии был стандартизирован как ECMA–334 и ISO/IEC 23270.

C Sharp относится к семье языков с C–подобным синтаксисом, из них его синтаксис наиболее близок к C++ и Java. Язык имеет статическую типизацию, поддерживает полиморфизм, перегрузку операторов (в том числе операторов явного и неявного приведения типа), делегаты, атрибуты, события, свойства, обобщённые типы и методы, итераторы, анонимные функции с поддержкой замыканий, LINQ, исключения, комментарии в формате XML.

Переняв многое от своих предшественников — языков C++, Pascal, Модула, Smalltalk и, в особенности, Java — C#, опираясь на практику их использования, исключает некоторые модели, зарекомендовавшие себя как проблематичные при разработке программных систем, например, C# в отличие от C++ не поддерживает множественное наследование классов (между тем допускается множественное наследование интерфейсов).

3. 3. 4 Расширения для работы с изображениями

Для работы с изображениями использовалась OpenCV (Open Source Computer Vision Library, библиотека компьютерного зрения с открытым исходным кодом) — библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом[34]. Реализована на C/C++, также разрабатывается для Python, Java, Ruby, Matlab, Lua и других языков.

Для того, чтобы воспользоваться возможностями OpenCV в среде . Net понадобилась Emgu CV – кросс–платформенная . Net обертка для Open CV, библиотеки обработки изображений. Она позволяет вызывать функции Open CV из . NET совместимых языков программирования, таких как C#, VB, VC++, IronPython и т. п. Обертка может быть скомпилирована в следующих средах программирования: Visual Studio, Xamarin Studio and Unity, и использоваться под операционными системами Windows, Linux, Mac OS X, iOS, Android и Windows Phone.

3. 4 Результаты, полученные в ходе исследования

3. 4. 1 Результаты, показанные на материале базы ORL

По результатам разбиения наборов изображений базы ORL для каждого класса количество изображений в классе, для различных вариантов разбиения, принимает значения от 3 до 8. Соответственно, количество попыток идентификации изображения, предпринятые каждым классификатором варьируется в диапазоне от 16 до 398. Процент успешных распознаваний для всех получившихся комбинаций показан в таблице 2

Таблица 2 – Результаты, показанные на материале базы ORL

Количество изображений в классе	Количество попыток распознавания	Доля успешных распознаваний, %					
		Алгоритм распознавания					
		Главные компоненты			Локальные бинарные шаблоны		
		80 ГК	160 ГК	160 ГК и световая норм.	Радиус = 1	Радиус = 5	Радиус = 5 и световая нормализация
3	16	69	69	75	75	56	56
4	96	80	80	80	92	74	74
5	284	86	86	84	92	85	86
6	156	87	88	88	95	87	86
7	251	90	91	92	95	94	94
8	398	93	93	92	98	96	96
	1201	89	89	89	95	90	90

Число испытаний при количестве изображений в классе равном 3 заметно ниже прочих вариантов, поэтому на этот результат можно ориентироваться лишь для определения некоей общей тенденции, основное же внимание следует уделить результатам во всех остальных случаях.

В целом, очевидно, что с увеличением количества изображений в классе эффективность распознавания заметно растет, в независимости от классификатора и его параметров.

Кроме того, метод локальных бинарных шаблонов, практически во всех случаях, показывает лучшие результаты, чем метод главных компонент. В среднем, преимущество около 5% (в отдельных случаях более 10%) эффективности.

Световая нормализация не оказывает значимого влияния на эффективность обоих алгоритмов (различия 1–2%).

На рисунках 11 и 12 показаны результаты световой нормализации для изображений лиц для каждой базы. Очевидно, что такая обработка не вносит существенных изменений в изображения базы ORL, в то время как изменения в изображениях FERET заметны невооруженным взглядом. Это и объясняет различия во влиянии данного параметра на эффективность методов распознавания.

Рисунок 11 – Световая нормализация изображений лиц базы ORL

Рисунок 12 – Световая нормализация изображений лиц базы FERET

Для метода главных компонент, очевидно, что увеличение количества главных компонент с 80 до 160 также не дает принципиального изменения в результатах. Это можно объяснить тем, что значимая в данной ситуации информация вполне достаточным образом описывается 80 компонентами, и дальнейшее добавление компонент не имеет смысла, так как информация, попадающая в эти компоненты, не содержит особенностей, позволяющих отличить изображения других классов и отнести обрабатываемое изображение к «своему» классу.

На рисунке 13 видно, что кривые графиков всех результатов метода главных компонент схожи и очень близки друг относительно друга, так что практически сливаются в одну линию.

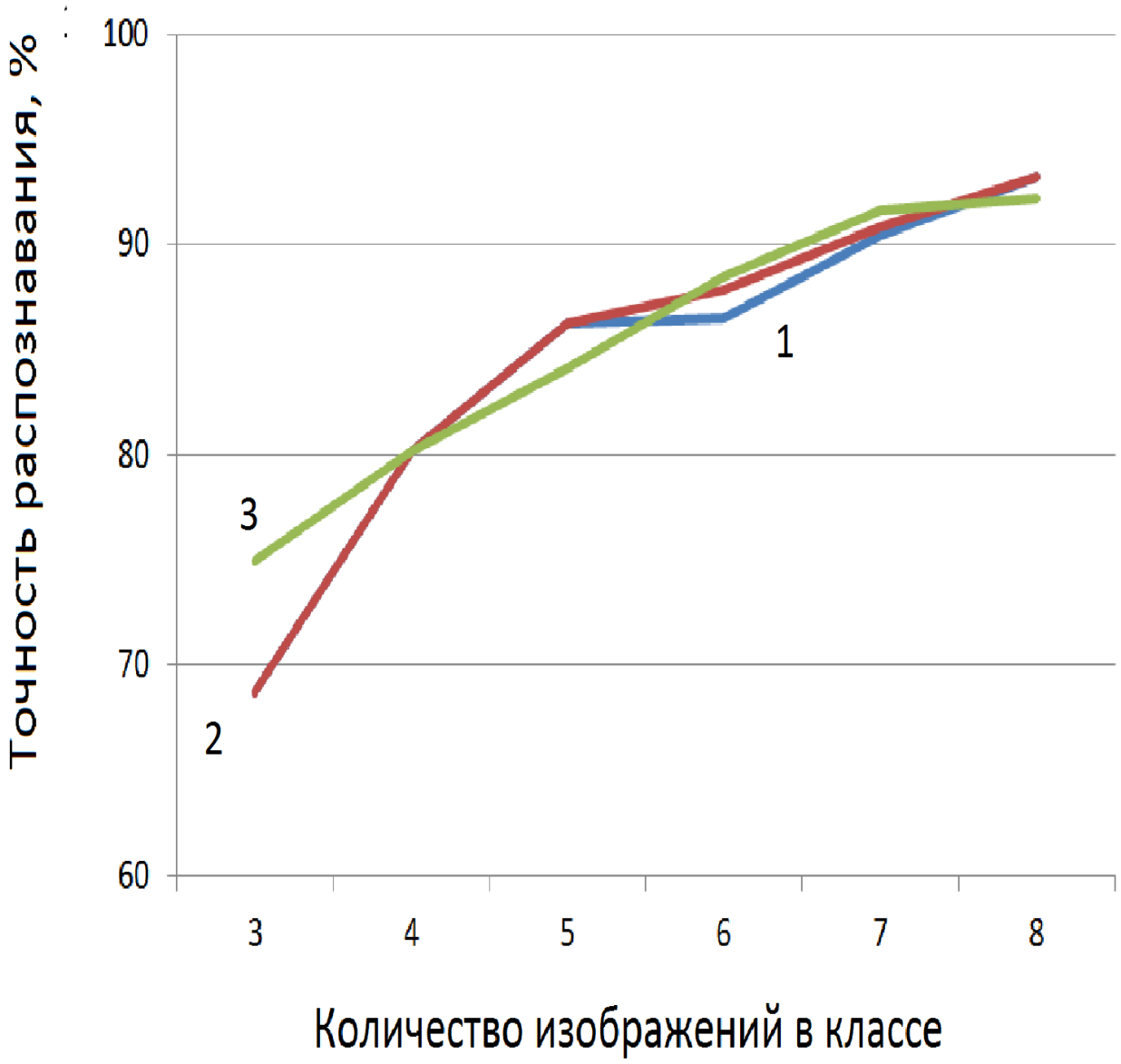

Рисунок 13 – График результатов метода главных компонент с различными параметрами для базы ORL (1 – 80 главных компонент,2 – 160 главных компонент без световой нормализации, 3 – 160 главных компонент со световой нормализацией)

Для метода локальных бинарных шаблонов изменения параметра радиуса – величины каждого отдельного обрабатываемого участка оказало негативное влияние на выборках с малым количеством обучающих примеров, однако с увеличением этого количества влияние уменьшается, и кривые графиков

сближаются (рисунок 14) Вероятнее всего, это объясняется не столько количеством изображений в классе, сколько увеличением количества испытаний, а следовательно и снижением значимости каждого отдельного испытания, связанным с перебором различных сочетаний изображений в тестовом и обучающем наборе, при использовании метода перекрестной проверки.

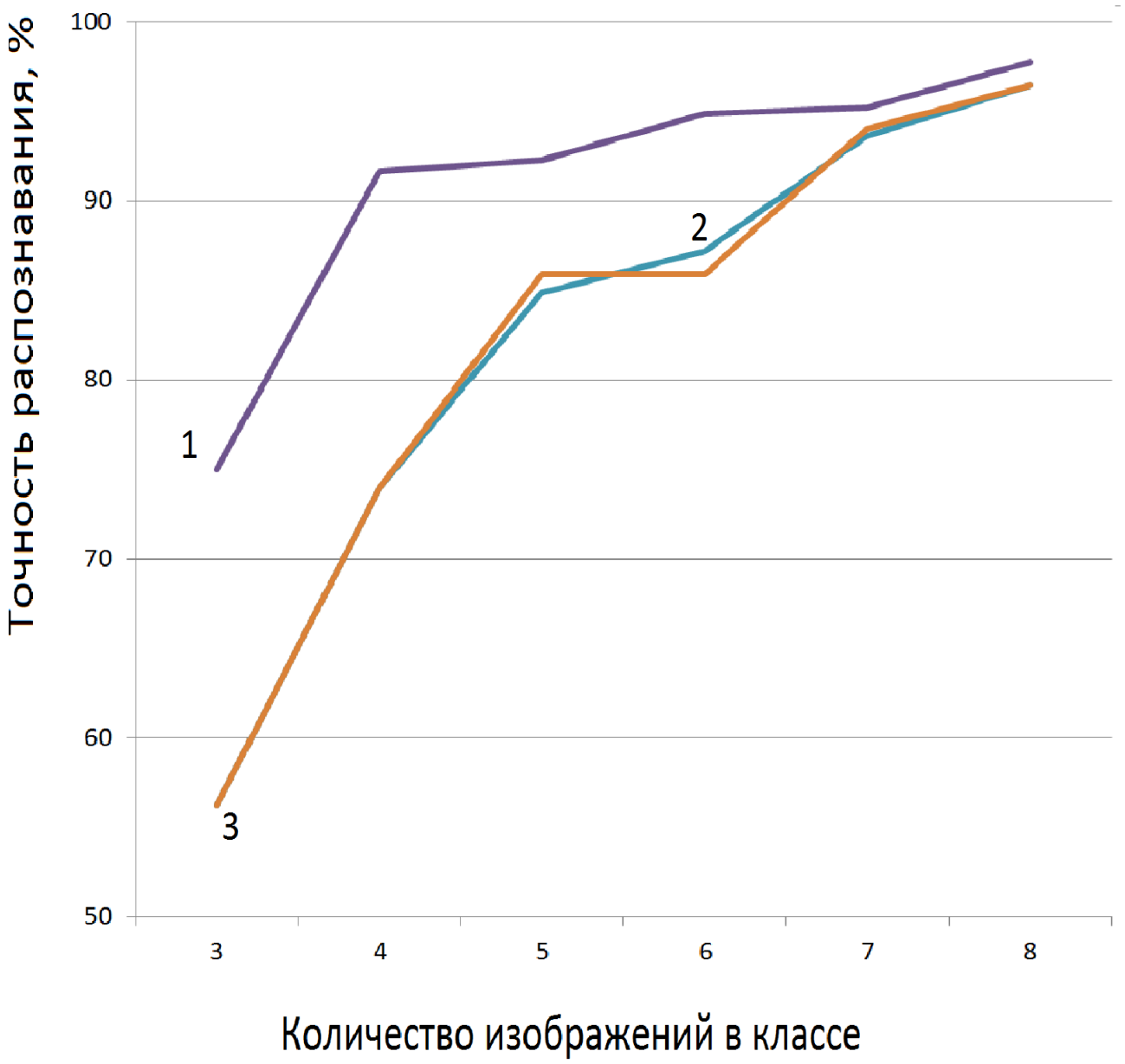

Рисунок 14 – График результатов метода локальных бинарных шаблонов с различными параметрами для базы ORL (1 – Радиус паттерна = 1 ,2 Радиус паттерна = 5 без световой нормализации, 3 – Радиус паттерна = 5 со световой нормализацией)

В целом, средние значения эффективности для обоих алгоритмов на материале базы ORL достаточно похожи, тем не менее, результаты метода локальных бинарных шаблонов в лучшем сочетании параметров все же несколько выше результатов метода главных компонент.

Графики лучших результатов обоих методов на материале базы ORL показаны на рисунке 15

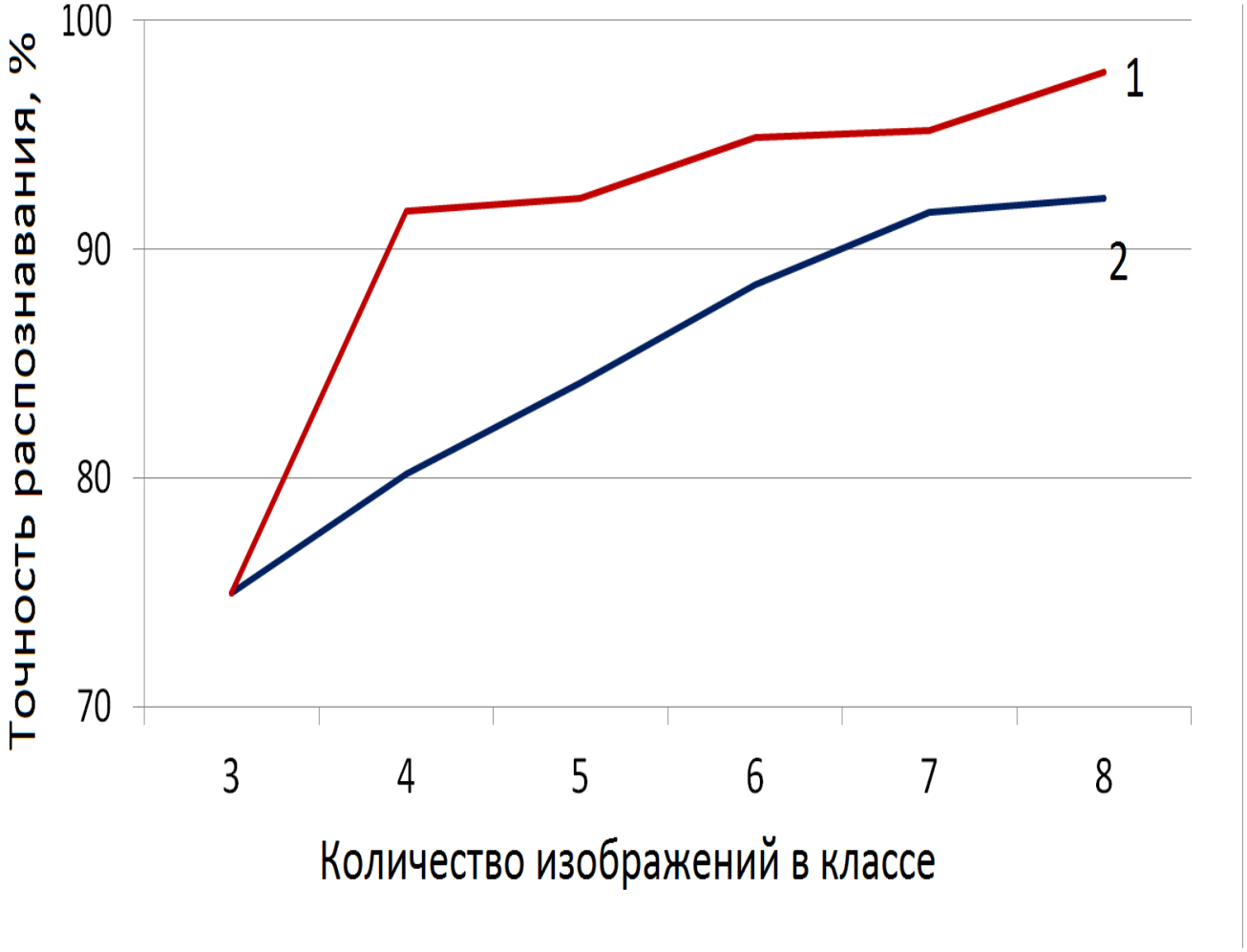

Рисунок 15 – Графики лучших результатов обоих методов для базы ORL (1 – Радиус паттерна = 1 без световой нормализации,2 – 160 главных компонент со световой нормализацией)

3. 4. 2 Результаты, показанные на материале базы FERET

Изображения лиц, представленные в базе FERET по своим особенностям существенно отличаются от изображений базы ORL. Во–первых, ракурс съемки многих изображений куда сильнее отклоняется от запечатления анфас, чем в ORL, как и положение головы человека относительно осей изображения, во–

вторых, условия освещенности имеют намного больше вариаций расположения относительно источника света и его интенсивности.

Немаловажен также тот факт, что в базе FERET вариативность мимики человека на снимке намного больше, чем в ORL.

Кроме того, количество изображений в классе, зачастую, куда больше, чем для ORL, да и общее количество классов на порядок больше. Само собой, такие различия в исходном материале оказывают существенное влияние на результативность обработки изображений классификаторами, усложняя задачу идентификации. Количество классов из базы изображений FERET 477, а количество изображений в классе от 2 до 28 изображений. А количество попыток распознавания изображений для различного количества изображений в классе составляет минимум 32 и максимум 2571.

Наибольшее количество испытаний более 1000 для каждого варианта проводилось на наборах, где количество изображений в классе составляло от 4 до 9 изображений.

В рамках исследования эти особенности полезны для того, чтобы с одной стороны, зафиксировать более объективные результаты работы классификаторов при увеличении количества изображений в классе, а с другой пронаблюдать, насколько измениться их эффективность, при работе в условиях более близких к реальным практическим задачам подобных систем.

Разумеется, все эти факторы оказали влияние на итоговые результаты распознавания, хотя наиболее общие закономерности, показанные на материале базы ORL отчетливо прослеживаются и для базы FERET.

Увеличение количества изображений в классе также положительно сказывается на эффективности распознавания, в независимости от применяемого алгоритма. Использование количества главных компонент больше 80 практически не влияет на результаты метода главных компонент, а световая нормализация не оказывает влияния на эффективность метода локальных бинарных шаблонов. Результаты, показанные на материале базы FERET ,

исследуемыми методам со всеми выбранными вариантами исходных параметров приведены в таблице 3.

Таблица 3 – Результаты, показанные на материале базы FERET

Количество изображений в классе	Кол-во обработанных изображений	Доля успешных распознаваний, %					
		Алгоритм распознавания					
		МГК			ЛБШ		
		80 ГК	160 ГК	160 ГК и свет. норм.	Радиус = 1	Радиус = 5	Радиус = 5 и свет. норм.
2	256	27	29	29	31	32	32
3	824	32	32	36	44	50	50
4	1502	32	32	38	45	56	55
5	1848	44	45	45	54	70	68
6	2006	39	40	41	50	62	62
7	1509	45	47	47	56	74	73
8	2571	45	47	47	55	75	74
9	1342	47	49	53	64	75	76
10	514	46	47	58	63	71	70
11	363	41	41	52	56	65	65
12	548	44	46	57	58	72	73
13	281	43	44	59	59	73	74
14	230	46	47	59	57	73	69
15	230	44	44	54	63	76	72
16	175	53	52	65	72	74	74
17	158	41	41	55	63	71	72
18	213	34	34	47	63	69	70
19	44	68	75	86	77	82	86
20	123	50	50	72	74	83	82
21	74	51	55	76	69	89	84
22	25	60	60	76	64	88	76
23	61	51	51	62	75	85	85
24	63	56	56	78	76	84	83
26	32	47	47	66	69	91	88
27	32	50	50	78	75	84	81
28	65	38	38	72	71	85	80
	15089	42	43	47	55	68	68

В то же время, для метода главных компонент световая нормализация дает ощутимый прирост в эффективности, малозаметный при небольшом количестве изображений в классе, растущий до 10% с ростом количества изображений до 18. И до 20–30% при последующем увеличении количества изображений в классе. Графики результатов исследованных вариантов комбинации параметров классификатора, работающего методом главных компонент, представлены на рисунке 16

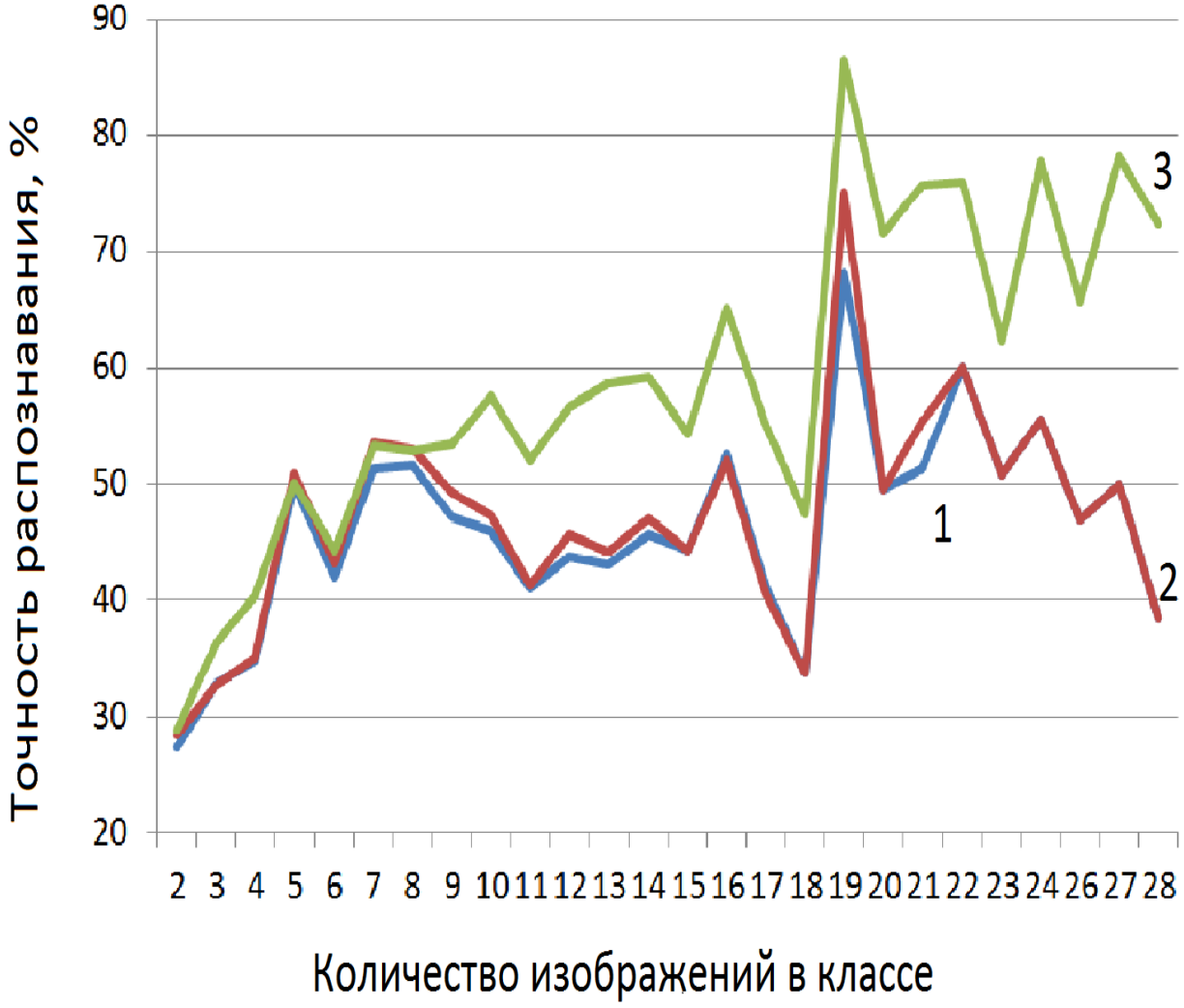

Рисунок 16 – Графики результатов метода главных компонент с различными параметрами для базы FERET(1 – 80 главных компонент,2 – 160 главных компонент без световой нормализации, 3 – 160 главных компонент со световой нормализацией)

Для результатов метода локальных бинарных шаблонов, заметно что увеличение радиуса каждого отдельного паттерна на материале базы FERET

сказывается положительно. Так, увеличение радиуса до 5, дает преимущество в результатах в 5 – 20% для различного количества изображений в классе. Эта картина очевидна при изучении графиков результатов метода локальных бинарных шаблонов, изображенных на рисунке 17

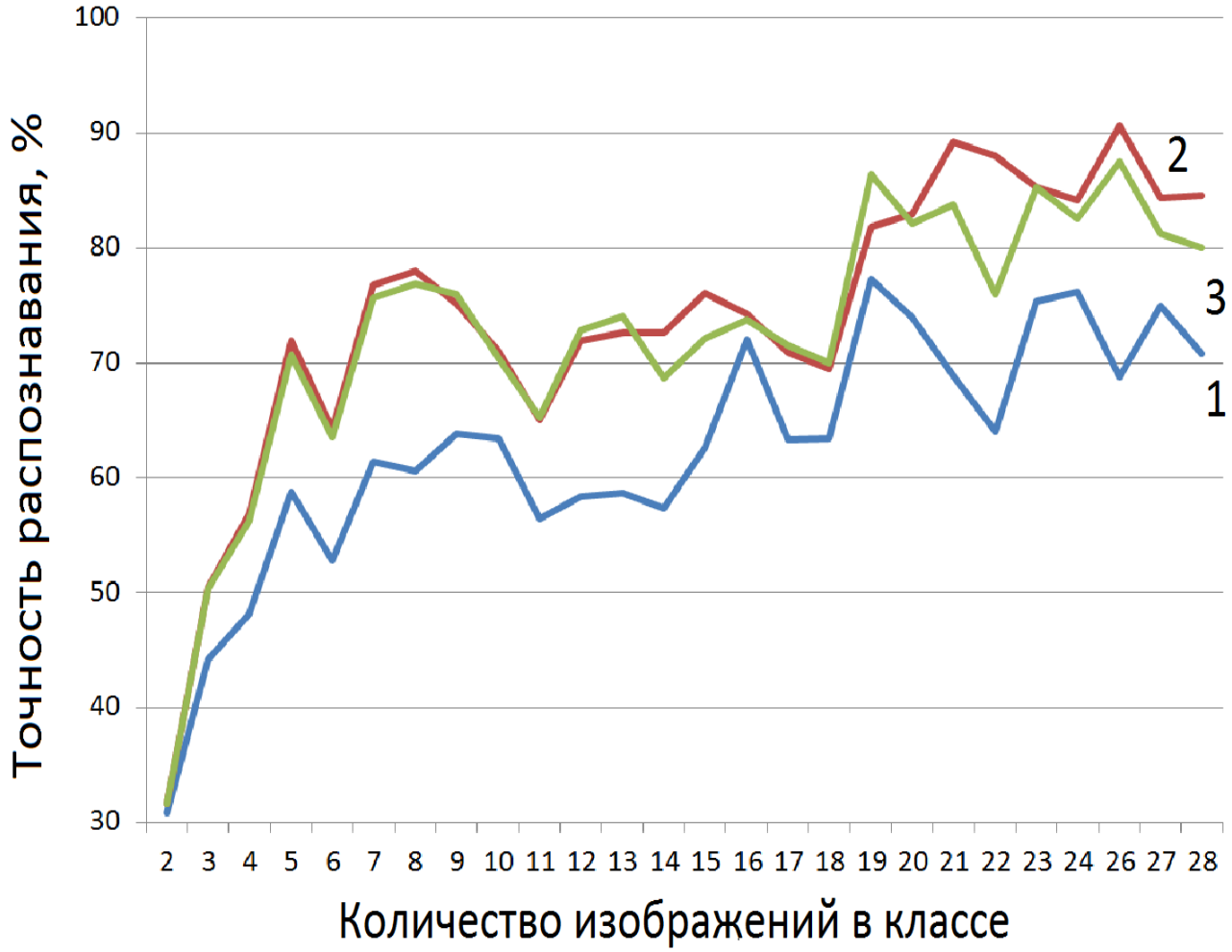

Рисунок 17 – Графики результатов метода локальных бинарных шаблонов с различными параметрами для базы FERET. (1 – Радиус паттерна = 1 ,2 Радиус паттерна = 5 без световой нормализации, 3 – Радиус паттерна = 5 со световой нормализацией)

Таким образом, среди исследованных методов и различными вариантами исходных параметров, на базе изображений FERET лучшие результаты для метода локальных бинарных шаблонов показал вариант с увлеченным радиусом паттерна, а для метода главных компонент определяющим фактором оказалась световая нормализация. Среди лучших результатов обоих методов, алгоритм локальных бинарных шаблонов, имеет преимущество над методом главных

компонент, в среднем, более 10%, в отдельных случаях эта разница варьируется от 5 до 27%. На графиках (рисунок 18) видно, что наибольшее преимущество наблюдается при объеме выборки в 5 – 8 изображений, где количество испытаний, которые прошли методы ощутимо выше. Это обстоятельство позволяет предположить, что данное наблюдение можно расценивать как наиболее объективное и независимое от случайных особенностей отдельных классов изображений.

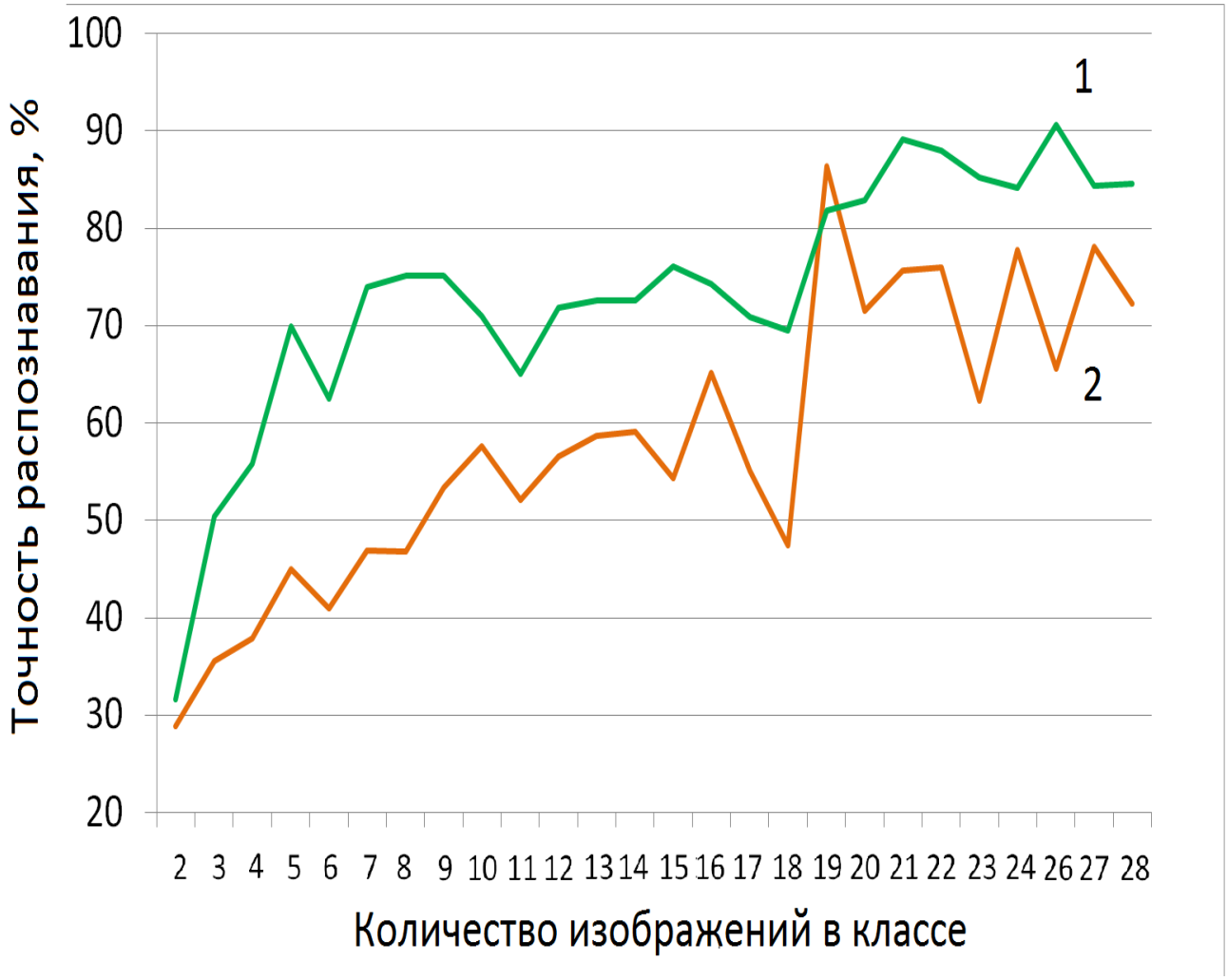

Рисунок 18 – Графики лучших результатов методов для базы FERET. (1 – Радиус паттерна = 5 без световой нормализации, 2 – 160 главных компонент со световой нормализацией)

3. 4. 3 Сравнение результатов методов для баз FERET и ORL

Особенности результатов, показанных классификаторами на материале, различных баз изображений, во многом определяются спецификой изображений, включенных в эти базы. Однако, они позволяют сделать выводы об общих

закономерностях в результативности рассматриваемых методов. Поскольку зависимость эффективности идентификации от количества изображений в классе не вызывает сомнений, правомерно сравнивать лишь результативность, показанную на выборках с одинаковым объемом тестового набора. В данной ситуации, это количество в пределах от 3 до 8 изображений. В таблице 4 приведены полученные результаты.

Таблица 4 – Сравнение результатов методов для баз FERET и ORL

Кол-во изображений в классе	Метод главных компонент						Метод локальных бинарных шаблонов					
	80 ГК		160 ГК		160 ГК и световая нормализ.		Радиус = 1		Радиус = 5		Радиус = 5 и световая нормализ.	
	FERET	ORL	FERET	ORL	FERET	ORL	FERET	ORL	FERET	ORL	FERET	ORL
3	32	69	32	69	36	75	44	75	50	56	50	56
4	32	80	32	80	38	80	45	92	56	74	55	74
5	44	86	45	86	45	84	54	92	70	85	68	86
6	39	87	40	87	41	88	50	95	62	87	62	86
7	45	90	47	90	47	92	56	95	74	94	73	94
8	45	93	47	93	47	92	55	98	75	96	74	96

Рисунок 19 наглядно отображает соотношение лучших результатов метода главных компонент и метода локальных бинарных шаблонов на материалах баз ORL и FERET.

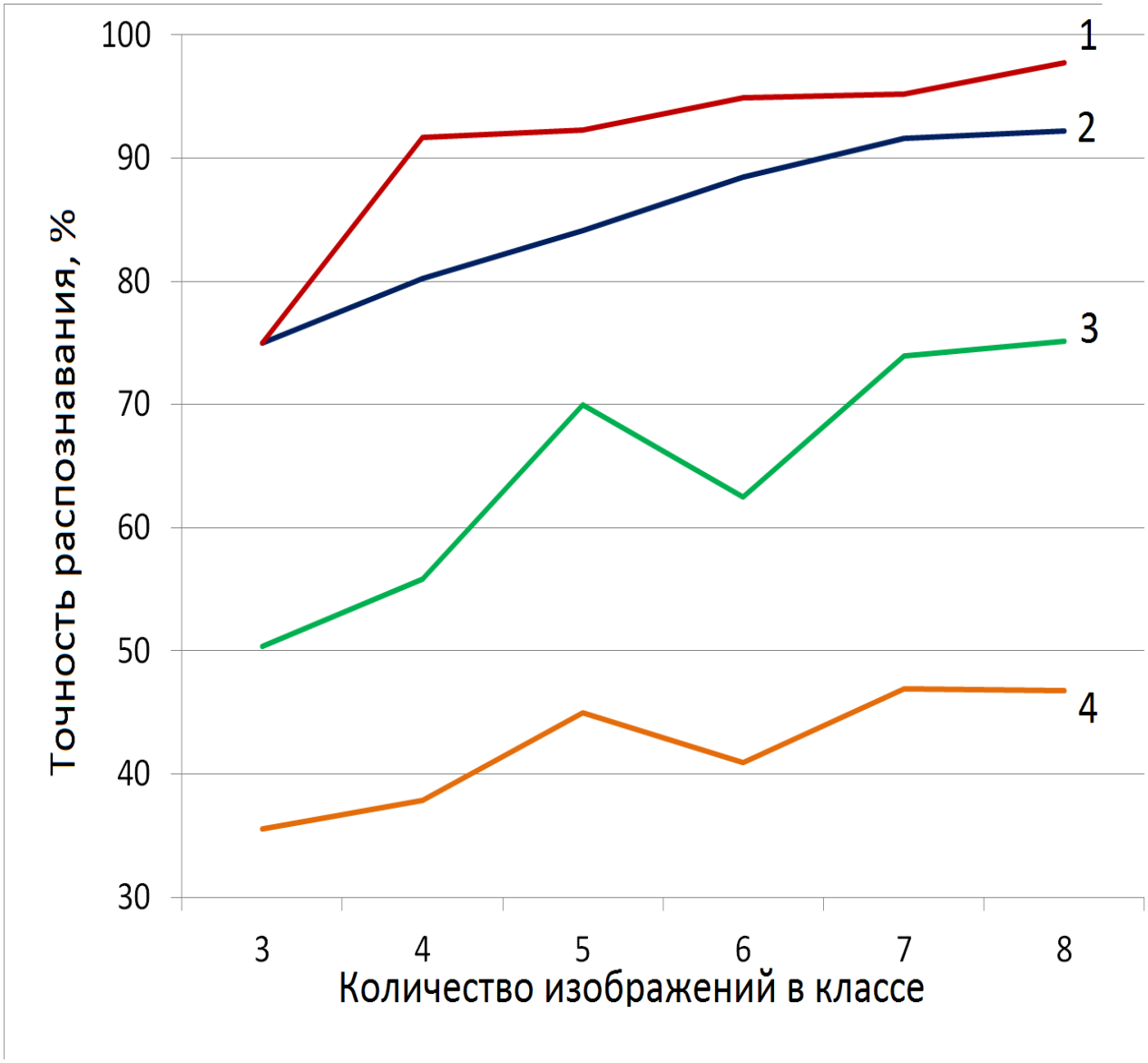

Рисунок 19 – Сравнение лучших результатов методов для баз FERET и ORL (1–метод локальных бинарных шаблонов на материале ORL; 2–метод главных компонент на материале ORL; 3–метод локальных бинарных шаблонов на материале FERET; 4–метод главных компонент на материале FERET)

.

ЗАКЛЮЧЕНИЕ

В ходе работы были исследованы различные подходы к классификации методов распознавания лиц, проанализированы соответствующие методов. Рассматриваются существующие подходы к построению выборочной совокупности.

По результатам изучения этих вопросов были выбраны методы для исследования и сравнения эффективности идентификации человека по изображению лица: метод главных компонент и метод локальных бинарных шаблонов.

Исходным материалом для оценки и сравнения эффективности методов стали базы данных ORL и FERET. Для объективного сравнения результатов идентификации были подготовлены несколько вариантов выборок изображений с разбиением имеющихся наборов изображений в классах на различное количество частей. Путем перекрёстной проверки результаты для аналогичных наборов данных. Была проведена обработка полученные наборы были с помощью исследуемых методов с применением различных параметров.

Сравнение результатов данной обработки позволило сделать следующие выводы:

1. В целом, метод локальных бинарных шаблонов обладает большей эффективностью, чем метод главных компонент как на материале базы FERET, так и на материале ORL.
2. С увеличением количества изображений в классе эффективность идентификации растет на материале обоих баз для обоих методов с любым сочетанием выбранных параметров;
3. На материале базы ORL с увеличением количества изображений в классе различия между эффективностью методов уменьшаются.
4. На материале базы FERET с увеличением количества изображений в классе различия между эффективностью методов незначительно увеличиваются.

5. На материале базы FERET преимущество в эффективности метода локальных бинарных шаблонов над методом главных компонент больше, чем на материале базы ORL.

6. Влияние световой нормализации на результаты метода локальных бинарных шаблонов минимально для обеих баз.

7. Метод главных компонент несколько эффективнее при световой нормализации на материале базы FERET, влияние этого параметра менее заметно, чем на материале базы ORL.

8. На материале базы ORL увеличение радиуса снижает эффективность метода локальных бинарных шаблонов на FERET ситуация обратная.

9. Самый низкий показатель эффективности идентификации изображений лиц у метода главных компонент без световой нормализации, при малом количестве изображений в классе.

10. Лучшие результаты эффективности идентификации изображений лиц показывает метод локальных бинарных шаблонов с минимальным радиусом без световой нормализации на материале базы ORL при большом количестве изображений в классе.

На основании сделанных выводов, можно заключить, что метод локальных бинарных шаблонов, практически во всех исследованных вариантах сочетания входных параметров и исходных данных, представленных для обработки, превосходит по эффективности распознавания метод главных компонент. Это преимущество становится очевидным, когда исходный набор данных имеет достаточно большой объем, тем самым снижается вероятность ошибки в результатах исследования, навязанной спецификой обрабатываемого материала и фактором случайности. При использовании в качестве исходного материала изображений изначально имеющих меньшую «степень нормализации» такое преимущество куда более ощутимо. По этой причине, в реальных условиях, можно рекомендовать использование метода ЛБШ.

В процессе исследования стало очевидно, что обучение классификаторов при большом объеме исходных данных весьма требовательно к аппаратным ресурсам и достаточно продолжительно по времени. Частично данную проблему удалось решить путем распараллеливания процесса обработки данных и запуска её в несколько потоков. Однако, на отдельной рабочей станции ограничения производительности все же оказались весьма ощутимыми. Вследствие этого для решения масштабных задач рекомендуется использовать распределённые вычисления непосредственно для обработки данных и производительную систему хранения данных для записи и хранения обученных классификаторов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Kirby, M. Application of the KL procedure for the characterization of human faces / M. Kirby, L. Sirovich // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1990. – Vol. 12(1). – P. 103-108

2. Viola P. , Jones M. Rapid object detection using a boosted cascade of simple features // IEEE CVPR01. – 2001. – P. 511–518.

3. Moghaddam B. and Pentland A. Probabilistic Visual Learning for Object Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 1997, Vol. 19, P. 696-710.

4. D. B. Graham and N. M. Allinson “Face recognition using virtual parametric eigenspace signatures,” Image Processing and its Applications, P. 106-110, 1997.

5. Brunelli R. , Poggio T. Face recognition: features versus templates // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993. – Vol. 15. – No 10. – P. 235-241.

6. Kolsch M. , Turk M. Robust hand detection // Proc. of the 6th IEEE Intern. Conf. on Automatic Face and Gesture Recognition. – Seoul (Korea), May 2004. – Seoul, 2004. – P. 614–619.

7. Hallinan P. L. , Gordon G. G. , Yuille A. L. , Giblin P. , Mumford D. Two- and Three-Dimensional Patterns of the Face. Natick:A. K. Peters Ltd. 1999. – 260 P.

8. Golfarelli M, Maio D. and Maltoni D. On the Error-Reject Trade-Off in Biometric Verification Systems. IEEE Transactions on Pattern Analysis and Machine Intelligence 1997, Vol. 19, P. 786-796.

9. Ng C. W. , Ranganath S. Gesture recognition via pose classification // Proc. 15th Intern. Conf. on Pattern Recognition. – Barcelona (Spain), 2000. – V. 3. – P. 699–704.

10. Oka K. , Sato Y. , Koike H. Realtime fingertip tracking and gesture recognition // Proc. IEEE Computer Graphics and Applications. – 2002. – V. 22. – №6. – P. 64–71.

11. Huang C. , Jeng S. A modelbased hand gesture recognition sy stem// Machine Vision and Application. – 2001. – V. 12. – №5.

12. Foltyniewicz R. Efficient High Order Neural Network for Rotation, Translation and Distance Invariant Recognition of Gray Scale Images. Lecture Notes in Computer Science - Computer Analysis of Images and Patterns, 1995, P. 424-431
13. Belhumeur P. N. , Hespanha J. P. and Kriegman D. J. Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection. IEEE Transactions on Pattern Analysis and Machine Intelligence 1997, Vol. 19, P. 711-720.
14. Samaria F. Face Recognition Using Hidden Markov Models // PhD thesis, Engineering Department, Cambridge University, 1994
15. Chen F. , Fu C. , Huang C. Hand gesture recognition using a realtime tracking method and Hidden Markov Models // Image and Vi sion Computing. – Oxford: Elsevier, 2003. – V. 21. – №8. – P. 745–758.
16. L. Wiskott et al. , «Face Recognition by Elastic Bunch Graph Matching», Trans. IEEE Pattern Analysis and Machine Intelligence, July 1997, pp. 775-779
17. Wiskott L. , Fellous J. -M. , Krueger N and Malsburg C. Face Recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence 1997, Vol. 19, P. 775-779
18. Lawrence S. , Giles C. L. , Tsoi A. C. , Back A. D. Face Recognition: A Convolutional Neural Network Approach // IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition, 1997. P. 1-24
19. Le Hoang Thai, Vo Nhat Truong// Face Alignment Using Active Shape Model And Support Vector Machine // International Journal of Biometrics and Bioinformatics, 2015. P. 224 – 234
20. Введение в оценку биометрических систем / П. Дж. Филлипс, Э. Мартин, С. Л. Уилсон, М. Пржибоски // Открытые Системы, №03, 2000. http://www. osp. ru/os/2000/03/ (дата об ращения: 17. 06. 2013)
21. Самаль Д. И. , Старовойтов В. В. - Подходы и методы распознавания людей по фотопортретам. - Минск, ИТК НАНБ, 1998. - 54с.

22. Пентланд А. С. , Чаудхари Т. Распознавание лиц для интеллектуальных сред // Открытые Системы, №03, 2000. http://www. osp. ru/os/2000/03/ (дата об ращения: 14. 09. 2015)
23. Галушкин А. И. , Томашевич Д. С. , Томашевич Н. С. Методы реализации инвариантности к аффинным преобразованиям двумерных изображений // Приложение к журналу «Информационные технологии», 2001, №1, с. 1-19
24. Прэтт У. Цифровая обработка изображений. – М:Мир, 1982. – Том 2. – С, 792
25. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика, 1992 - С. 184
26. Мокеев, В. В. О задаче нахождения собственных значений и векторов больших матричных систем / В. В. Мокеев // Журнал вычислительной математики и математической физики. – 1992. – № 32 (10). – С. 1652-1657.
27. Мокеев, В. В. О решении проблемы выборки малого размера при использовании линейного дискриминантного анализа в задачах распознавания лиц / В. В. Мокеев, С. В. Томилов // Бизнес информатика. – 2013. – № 1. – С. 37-43
28. Мокеев, А. В. О точности и быстродействии метода синтеза главных компонент / А. В. Мокеев // Бизнесинформатика. – 2010. – № 3(13). – С. 65-68.
29. Мокеев, В. В. О повышение эффективности вычислений главных компонент в задачах анализа изображений / В. В. Мокеев // Цифровая обработка сигналов. – 2011. – № 4. – С. 29-36.
30. Головко В. А. Нейроинтеллект: Теория и применения. Книга 2. Самоорганизация, отказоустойчивость и применение нейронных сетей - Брест:БПИ, 1999, - 228с.
31. Чернодуб А. Генетические алгоритмы в лицах/ URL: http://habrahabr.ru/post/221137/ (дата обращения 13. 08. 2014)
32. Library documentation/Microsoft /Developer Network// 2013URL: https://msdn.microsoft.com/library/67ef8sbd.aspx (дата обращения: 01. 03. 2014)

33. Library documentation/Microsoft /Developer Network// 2013. URL: https://msdn.microsoft.com/library/dd831853.aspx (дата обращения: 01. 03. 2014)
34. OpenCV Documentation/ Face Recognition with OpenCV /OpenCV API Reference //URL: http://docs.opencv.org/2.4/modules/contrib/doc/facerec/ (дата обращения 25. 12. 2013)
35. The ORL Database of Faces, /AT&T Laboratories Cambridge 2002. //URL: http://www. cl. cam. ac. uk/research/dtg/attarchive/facedatabase.html (дата об ращения: 18. 04. 2015)
36. The Color FERET Database, / The National Institute of Standards and Technology (NIST) 2009. // URL: http://www. nist.gov/itl/iad/ig/colorferet.cfm (дата обращения: 18. 04. 2015).

ПРИЛОЖЕНИЕ А

Алгоритмы формирования тестового и обучающего наборов.

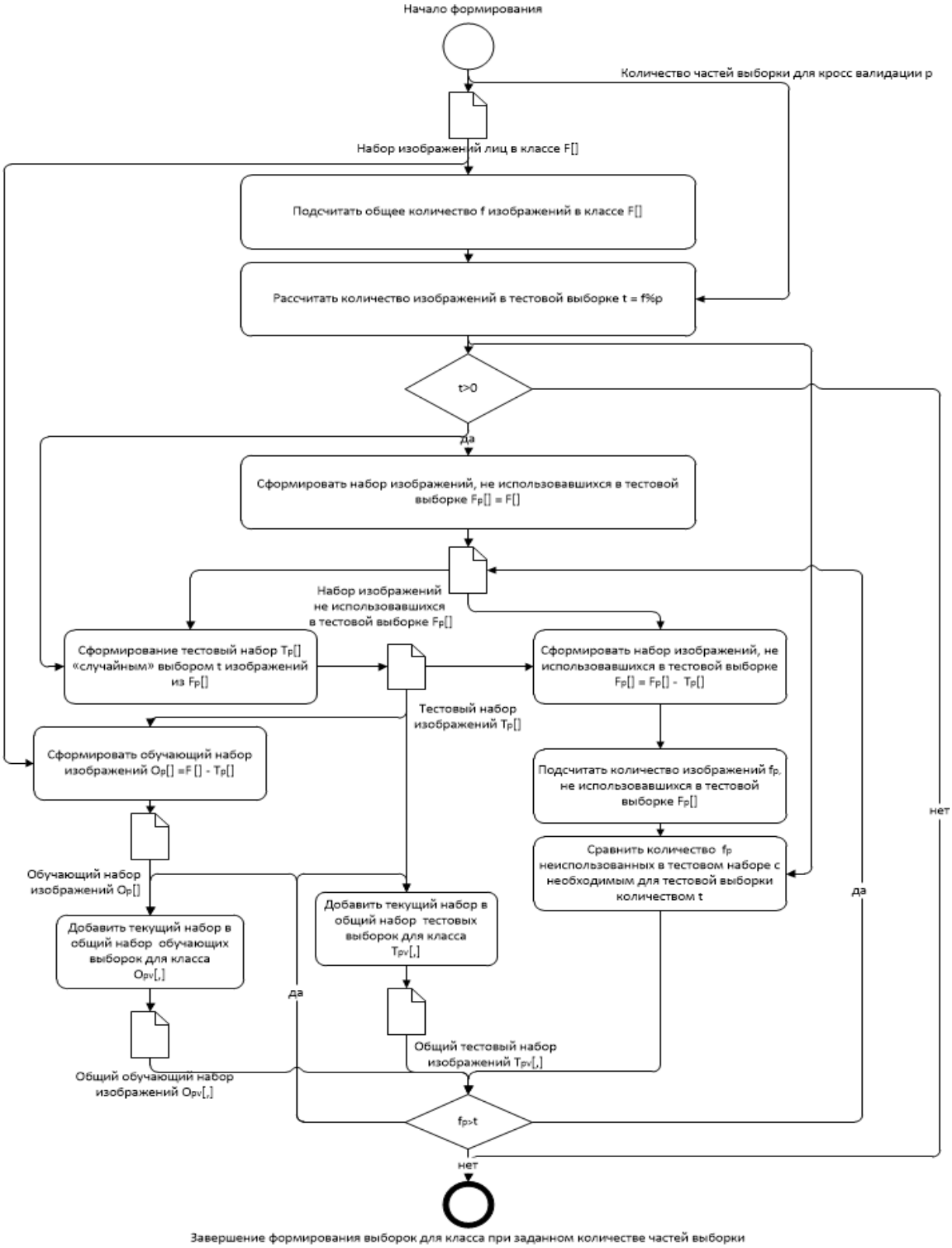

Рисунок А 1 – Алгоритм формирования тестового и обучающего набора для каждого класса в исходном наборе

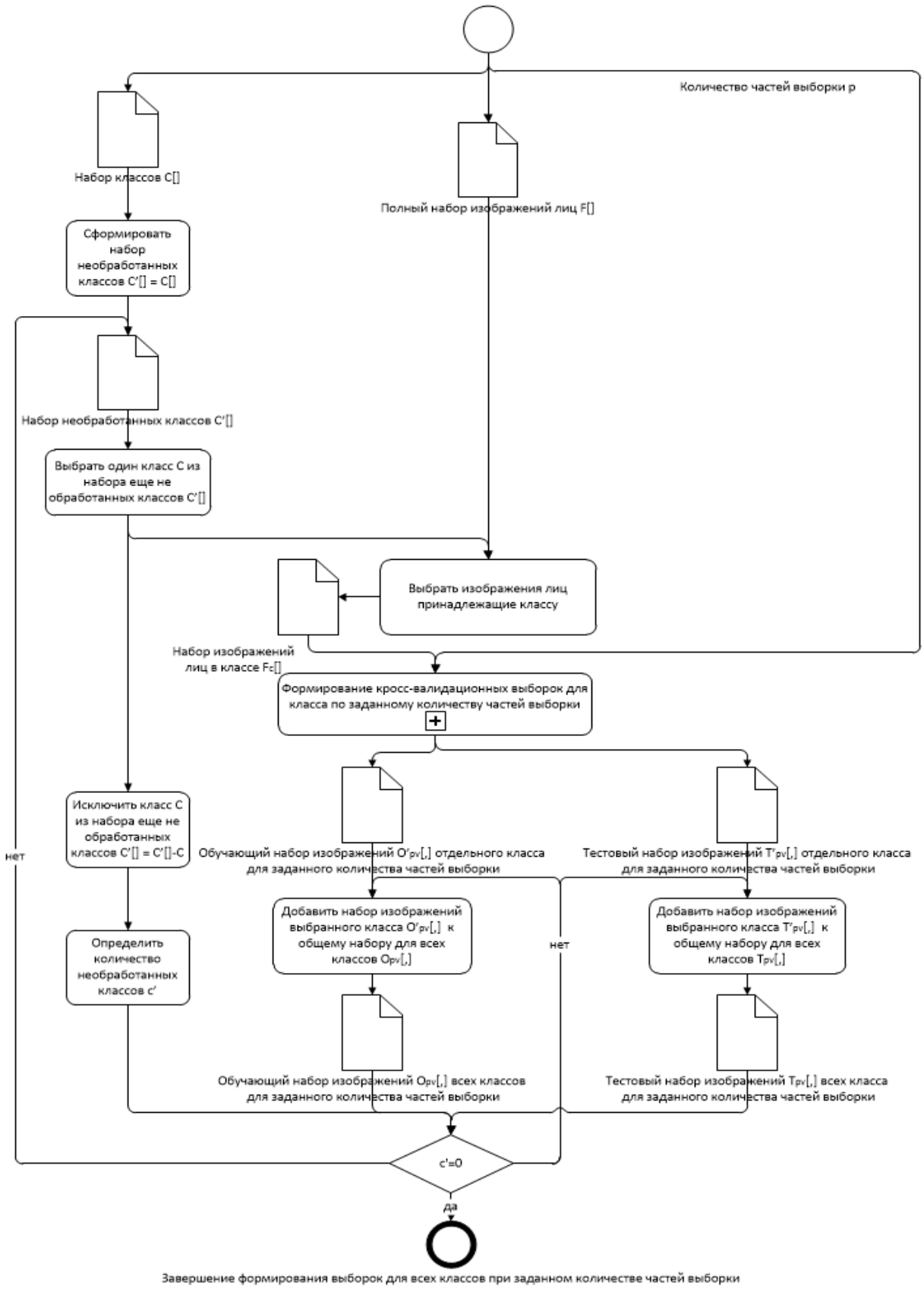

Рисунок Б 1 – Алгоритм формирования общего тестового и обучающего набора для всех классов

ПРИЛОЖЕНИЕ Б

Листинг многопотоковой обработки выборки изображений различными классификаторами с различным набором параметров

```
private static void runRecognizer(Object pc)
    {
        List<string> recognizerList = createRecognizerList();
        string equalize = "False";
        int partsCount = Convert. ToInt32(pc);
        for (int recList = 0; recList < recognizerList. Count(); recList++)
        {
            equalize = recognizerList[recList]. Split('\t')[3]. ToString();
            int recognizerId = Convert. ToInt32(recognizerList[recList]. Split('\t')[0]);
            FaceRecognizer oneRecognizer =
createRecognizer(recognizerList[recList]. Split('\t')[1], recognizerList[recList].
Split('\t')[2]);
            for (int variant = 0; variant < partsCount; variant++)
            {
                string[] finalSample =
selectFinalSampleForPartsAndVariant(partsCount, variant, dbConnection);
                int finalSampleId = Convert. ToInt32(finalSample[0]);
                List<int> learningPersonList = new List<int>();
                List<Image<Gray, byte>> learningImageList = new List<Image<Gray,
byte>>();
                for (int l = 0; l < finalSample[2]. Split(';'). Count(); l++)
                {
                    learningPersonList. Add(Convert. ToInt32(finalSample[2].
Split(';')[l]));
                    Image<Gray, byte> oneFace = new Image<Gray,
byte>(finalSample[4]. Split(';')[l]);
                    if (equalize == "True")
                    {
                        oneFace. _EqualizeHist();
                    }
                    learningImageList. Add(oneFace);
                }
                Console. WriteLine("parts " + partsCount + " variant " + variant + ", " +
oneRecognizer. GetType(). Name + " training started " + DateTime. Now);
                oneRecognizer. Train(learningImageList. ToArray(), learningPersonList.
ToArray());
                oneRecognizer. Save(Path. Combine(Path. Combine(@"D:\faceBase",
"recognizer"), oneRecognizer. GetType(). Name + "_" + partsCount + "_" + variant +
"_alt. xml"));
                Console. WriteLine("parts " + partsCount + " variant " + variant + ", " +
oneRecognizer. GetType(). Name + " training completed " + DateTime. Now);
```

```
                List<int> testingPersonList = new List<int>();
                List<string> testingFacePathList = new List<string>();
                List<Image<Gray, byte>> testingImageList = new List<Image<Gray,
byte>>();
                List<int> predictedPersonList = new List<int>();
                List<int> predictionDistanceList = new List<int>();
                for (int t = 0; t < finalSample[1]. Split(';'). Count(); t++)
                {
                    testingPersonList. Add(Convert. ToInt32(finalSample[1]. Split(';')[t]));
                    Image<Gray, byte> oneFace = new Image<Gray,
byte>(finalSample[3]. Split(';')[t]);
                    if (equalize == "True")
                    {
                        oneFace. _EqualizeHist();
                    }
                    testingImageList. Add(oneFace);
                    testingFacePathList. Add(finalSample[3]. Split(';')[t]);
                    int prediction = oneRecognizer. Predict(testingImageList[t]). Label;
                    int dis = 999999999;
                    dis = Convert. ToInt32(oneRecognizer. Predict(testingImageList[t]).
Distance);
                    predictionDistanceList. Add(dis);
                    predictedPersonList. Add(prediction);
                }
            }
            Console. WriteLine("parts " + partsCount + ", " + oneRecognizer.
GetType(). Name + " overall completed " + DateTime. Now);
        }
        threadsFinished++;
    }
```

ПРИЛОЖЕНИЕ В

Листинг вспомогательные функций формирования выборок для каждого класса

Создание полного списка изображений для отдельного класса

```
static string[] createPersonFullSampleList()
    {
        List<string> personFaceIdList = new List<string>();
        string[] pArray = selectPersonsArray();
        string[] pfArray = selectPersonFaceIdsArray();

        foreach (string onePerson in pArray)
        {
            string oneFullSample = onePerson. Split(';')[0] + ";";
            foreach (string onePersonsFace in pfArray)
            {
                if (onePerson. Split(';')[0] == onePersonsFace. Split(';')[0])
                {
                    oneFullSample += onePersonsFace. Split(';')[1] + ";";
                }
            }
            oneFullSample = oneFullSample. Substring(0, oneFullSample. Length – 1);
            personFaceIdList. Add(oneFullSample);
        }
        return personFaceIdList. ToArray();
    }
```

Создание вариантов выборки для одного класса по количеству частей разбиения исходного набора

```
    static List<string> createOnePersonsSamplesVariants(string
onePersonFullSample, int partsCount)
    {
        int facesQuantity = onePersonFullSample. Split(';'). Count() – 1;
        int testingFacesQuantity = (facesQuantity – (facesQuantity % partsCount)) /
partsCount;
        List<int> variantfaceIndicesList = new List<int>();
        List<int> defaultFaceIndicesList = new List<int>();
        List<string> allVariantsSamples = new List<string>();
        if (partsCount <= facesQuantity)
        {
```

```
for (int i = 0; i < facesQuantity; i++)
            {
                defaultFaceIndicesList. Add(Convert. ToInt32(onePersonFullSample.
Split(';')[i + 1]));
            }
            for (int v = 0; v < partsCount; v++)
            {
                string testingSampleString = "";
                string learningSampleString = "";
                string oneVariantSampleString;
                variantfaceIndicesList. Clear();
                for (int i = 0; i < facesQuantity; i++)
                {
                    variantfaceIndicesList. Add(Convert. ToInt32(onePersonFullSample.
Split(';')[i + 1]));
                }
                for (int t = 0; t < testingFacesQuantity; t++)
                {
                    Random randomFaceSelection = new Random();
                    int selectedIndex = randomFaceSelection. Next(0,
defaultFaceIndicesList. Count());
                    testingSampleString += (defaultFaceIndicesList[selectedIndex]).
ToString() + ";";
                    variantfaceIndicesList.
Remove(defaultFaceIndicesList[selectedIndex]);
                    defaultFaceIndicesList. RemoveAt(selectedIndex);
                }
                testingSampleString = testingSampleString. Substring(0,
testingSampleString. Count() – 1);
                for (int l = 0; l < variantfaceIndicesList. Count(); l++)
                {
                    learningSampleString += (variantfaceIndicesList[l]). ToString() + ";";
                }
                learningSampleString = learningSampleString. Substring(0,
learningSampleString. Count() – 1);
                oneVariantSampleString = partsCount. ToString() + "\t" +
testingSampleString + "\t" + learningSampleString + "\t" + onePersonFullSample.
Split(';')[0];
                allVariantsSamples. Add(oneVariantSampleString);
            }
        }
        return allVariantsSamples;
    }
    static FaceRecognizer createRecognizer(string title, string parameters)      {
```

```
FaceRecognizer currentRecognizer = null;
  switch (title)
  {
     case "EigenFaceRecognizer":
        int eComponents = Convert. ToInt32(parameters. Split(';')[0]);
        double eThreshold = (parameters. Split(';')[1] == "+i") ? (Double.
PositiveInfinity) : (Convert. ToDouble(parameters. Split(';')[1]));
        currentRecognizer = new EigenFaceRecognizer(eComponents,
eThreshold);
        break;
     case "LBPHFaceRecognizer":
        int radius = Convert. ToInt32(parameters. Split(';')[0]);
        int neigbors = Convert. ToInt32(parameters. Split(';')[1]);
        int gridX = Convert. ToInt32(parameters. Split(';')[2]);
        int gridY = Convert. ToInt32(parameters. Split(';')[3]);
        double lThreshold = Convert. ToDouble(parameters. Split(';')[4]);
        currentRecognizer = new LBPHFaceRecognizer(radius, neigbors, gridX,
gridY, lThreshold);
        break;
  }
  return currentRecognizer;
```