

Министерство науки и высшего образования РФ  
Федеральное государственное автономное  
образовательное учреждение  
высшего образования  
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»  
Институт космических и информационных технологий  
Кафедра систем искусственного интеллекта

УТВЕРЖДАЮ

Заведующий кафедрой

\_\_\_\_\_ Г. М. Цибульский

подпись

« \_\_\_\_\_ » \_\_\_\_\_ 2019 г.

## БАКАЛАВРСКАЯ РАБОТА

09.03.02 — «Информационные системы и технологии»

Разработка системы распознавания лиц по визуальным данным

Руководитель \_\_\_\_\_ канд. техн. наук, доцент каф. СИИ А. В. Пятаева  
подпись, дата

Выпускник \_\_\_\_\_ М.В. Верхотурова  
подпись, дата

Красноярск 2019

Продолжение титульного листа бакалаврской работы по теме «Разработка системы распознавания лиц по визуальным данным».

Нормоконтролер

\_\_\_\_\_

подпись, дата

А. В. Пятаева

Министерство науки и высшего образования РФ  
Федеральное государственное автономное  
образовательное учреждение  
высшего образования  
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»  
Институт космических и информационных технологий  
Кафедра систем искусственного интеллекта

УТВЕРЖДАЮ  
Заведующий кафедрой  
\_\_\_\_\_ Г. М. Цибульский  
подпись  
« \_\_\_\_ » \_\_\_\_\_ 2019 г.

**ЗАДАНИЕ  
НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ  
в форме бакалаврской работы**

Студентке Верхотуровой Марии Владимировне

Группа КИ15-12Б, направление 09.03.02 «Информационные системы и технологии», профиль подготовки 09.03.02.05 «Информационные системы и технологии в административном управлении».

Тема выпускной квалификационной работы «Разработка системы распознавания лиц по визуальным данным».

Утверждена приказом по университету № 6220/с от 08.05.2019

Руководитель ВКР А.В. Пятаева кандидат технических наук, доцент кафедры систем искусственного интеллекта ИКИТ СФУ.

Исходные данные для ВКР: разработана система распознавания и обнаружения лиц по визуальным данным.

Перечень разделов ВКР:

- введение;
- методы распознавания лиц;
- эмпирический метод;
- методы характерных инвариантных признаков;
- активные модели внешнего вида;
- метод главных компонент;
- метод гибкого сравнения на графах;
- скрытые Марковские модели;
- выводы по главе 1;
- разработка системы распознавания лиц по визуальным данным;
- метод Виолы–Джонса для обнаружения лиц;
- локальные бинарные шаблоны для распознавания лиц;
- выводы по главе 2;
- экспериментальные исследования;
- обнаружение лиц;
- распознавание лиц;
- выводы по главе 3;

- заключение;
- список использованной литературы;
- список публикаций по теме выпускной квалификационной работы;
- приложения А (плакаты презентации).

Перечень графического материала: презентация «Разработка системы распознавания лиц по визуальным данным»

Руководитель ВКР

\_\_\_\_\_

подпись

Задание принял к исполнению

\_\_\_\_\_

подпись

А. В. Пятаева

М.В. Верхотурова

« \_\_\_\_\_ » \_\_\_\_\_ 2019 г.

## График

Выполнения выпускной квалификационной работы студентом направления 09.03.02 «Информационные системы и технологии», профиля подготовки 09.03.02.05 «Информационные системы и технологии в административном управлении».

График выполнения выпускной квалификационной работы приведен в таблице 1.

Таблица 1 – График выполнения этапов ВКР

Наименование этапа	Срок выполнения этапа	Результат выполнения этапов	Примечание руководителя (отметка о выполнении этапа)
Ознакомление с целью и задачами работы	04.02-9.02	Краткое эссе по теме ВКР	Выполнено
Сбор литературных источников	10.02-18.02	Список использованных источников	Выполнено
Анализ собранных источников литературы	19.02-25.02	Реферат о проблемно предметной области	Выполнено
Уточнение и обоснование актуальности цели и задач ВКР	26.02-29.02	Окончательная формулировка цели и задач ВКР	Выполнено
Решение первой задачи ВКР	18.03-31.03	Доклад и презентация по решению первой задачи	Выполнено
Решение второй задачи ВКР	01.04-14.04	Доклад и презентация по решению второй задачи	Выполнено
Решение третьей задачи ВКР	15.04-31.05	Доклад и презентация по третьей задаче ВКР	Выполнено
Подготовка доклада и презентации по теме ВКР	01.06-05.06	Доклад с презентацией по теме ВКР	Выполнено

Окончание таблицы 1

Компоновка отчета по результатам решения задач ВКР	06.06-13.06	Отчет по результатам решения задач ВКР	Выполнено
Первичный нормоконтроль (Н/К)	14.06	Пояснительная записка, презентация ВКР	Выполнено
Предварительная защита результатов ВКР	18.06	Доклад с презентацией по теме ВКР	Выполнено
Вторичный нормоконтроль (Н/К)	24.06	Пояснительная записка, презентация ВКР	Выполнено
Итоговый нормоконтроль (Н/К)	27.06	Пояснительная записка, презентация ВКР	Выполнено
Защита ВКР	03.07	Пояснительная записка, доклад и презентация по результатам бакалаврской работы	

Руководитель ВКР

Студентка гр.КИ15-12Б

\_\_\_\_\_

подпись

\_\_\_\_\_

подпись

А. В. Пятаева

М.В. Верхотурова

« \_\_\_\_ » \_\_\_\_\_ 2019 г.

## СОДЕРЖАНИЕ

Введение.....	9
1 Методы распознавания лиц.....	11
1.1 Эмпирический метод .....	11
1.2 Методы характерных инвариантных признаков.....	12
1.3 Активные модели внешнего вида.....	13
1.4 Метод главных компонент .....	15
1.5 Метод гибкого сравнения на графах .....	16
1.6 Скрытые Марковские Модели.....	18
1.7 Выводы по 1 главе.....	20
2 Разработка системы распознавания лиц по визуальным данным.....	21
2.1 Метод Виолы-Джонса для обнаружения лиц.....	21
2.2 Локальные бинарные шаблоны для распознавания лиц .....	27
2.3 Выводы по 2 главе.....	32
3 Экспериментальные исследования.....	33
3.1 Обнаружение лиц .....	34
3.2 Распознавание лиц .....	39
3.3 Выводы по 3 главе.....	45
Заключение .....	47
Список использованной литературы.....	48
Список публикаций по теме выпускной квалификационной работы .....	53
Приложение А .....	54



## ВВЕДЕНИЕ

В настоящее время технологии распознавания лиц завоевали большую популярность, благодаря богатым возможностям и эффективности использования. Применение технологии распознавания лиц позволяет решить множество практических задач, таких как защита персональных данных и личной информации, обеспечение человеко-машинного взаимодействия, реализация доступа на территории объектов с повышенным уровнем безопасности, идентификация личности для обеспечения прокторинга в системах электронного обучения и многие другие [1].

Подходы к решению задачи распознавания лиц также значительно различаются. На начальном этапе развития подходов к распознаванию лиц использовались геометрические признаки для выделения характерных особенностей лица [1, 2]. В настоящее время для решения этой задачи используют технологии глубокого обучения [3, 4], эволюционные алгоритмы [5], метод роя частиц [6] и другие подходы. На эффективность распознавания могут оказывать влияние различные факторы, такие как низкий уровень освещенности и зашумленность изображений [7, 8], положение лица [9], пол и возраст и этническая принадлежность человека на изображении [10-12] и пр.

На сегодняшний день алгоритм обнаружения Виолы-Джонса и метод Локальных бинарных шаблонов является самым востребованным ввиду своей высокой скорости работы и высокой точности срабатывания.

Целью данной работы является разработка алгоритма и последующая его программная реализация, позволяющая решить задачу обнаружения и распознавания лиц по визуальным данным.

Поставленная цель достигается решением следующих задач:

1. обзор современных методов распознавания лиц;
2. определение недостатков существующих методов;
3. разработка алгоритма и его программная реализация обнаружения и распознавания лиц;

4. тестирование разработанной программной системы.

## **1 Методы распознавания лиц**

Распознавание лиц является активно разрабатываемой областью научных исследований как минимум пять десятилетий [1].

Подходы к решению задачи распознавания лиц эволюционировали от сопоставления геометрических признаков [2, 3] и учета текстурных особенностей [4] до применения технологий глубокого обучения [5, 6]. При этом выполняется распознавание лиц не только людей, но и обезьян [7], свиней для решения задачи идентификации животных на ферме [8] и даже статуй Бодхисаттвы [9].

К решению задачи распознавания лиц выделяют следующие подходы:

- эмпирический метод;
- методы характерных инвариантных признаков;
- активные модели внешнего вида;
- метод главных компонент;
- скрытые Марковские модели;
- метод гибкого сравнения на графах.

Рассмотрим особенности перечисленных подходов.

### **1.1 Эмпирический метод**

Эмпирический подход (knowledge based top-down methods) предполагает создание алгоритма, реализующего набор правил, которым должен отвечать фрагмент фотографии, для того чтобы считаться человеческим лицом. Этот набор правил – попытка формализовать эмпирические знания о том, как выглядит лицо на изображении, то есть формализовать то, чем руководствуется человек при принятии решения: видит он лицо или не лицо. Основные правила эмпирического подхода[18]:

- центральная часть лица имеет однородную яркость и цвет;

- у центральной и верхней частей лица разная яркость;
- лицо содержит в себе нос, рот и два симметрично расположенных глаза, резко отличающиеся по яркости относительно остальной части лица.

Достоинствами данного метода являются, достаточно просто реализуются, не требуют мощных вычислительных ресурсов, позволяют построить некоторую модель изображения лица и свести задачу к выполнению некоторого количества относительно простых проверок, возможность гибкой настройки под конкретную задачу путем модификации интуитивно понятных параметров, являются основой для адаптации эмпирических методов во многие современные методики обнаружения и распознавания лиц. К недостаткам метода относится, высокий процент ложных обнаружений или пропусков (в зависимости от жёсткости выбранных решающих правил), ограниченность возможных реализаций процессов, происходящих в мозгу во время решения задачи распознавания изображений, так как тот набор эмпирических знаний о человеческом лице, которые доступны исследователям на "сознательном уровне", далеко не исчерпывает инструментарий, используемый мозгом "подсознательно".

## **1.2 Методы характерных инвариантных признаков**

Методы характерных инвариантных признаков (Feature invariant approaches) являются методом для обнаружения лиц. В этих методах характерных инвариантных признаков не требуются в явном виде формализовать процессы, которые происходят в человеческом мозге. Сторонники этого подхода пытаются выявить свойства и закономерности изображения лица неявно, то есть найти инвариантные особенности лица, независимо от положения или угла наклона. Основные этапы алгоритма[4]:

- детектирование на изображении явных признаков лица: глаз, носа, рта;
- обнаружение: границы лица, форма, яркость, текстура, цвет;

- объединение всех найденных инвариантных признаков и их верификация.

Метод обнаружения лиц в сложных сценах [3] предполагает поиск правильных геометрических расположений черт лица. Для этого применяется гауссовский производный фильтр с множеством различных масштабов и ориентаций. После этого производится поиск соответствия выявленных черт и их взаимного расположения случайным перебором. Суть метода группировки признаков [4] в применении второй производной гауссовского фильтра для поиска интересующих областей изображения. Далее группируются края вокруг каждой такой области при помощи порогового фильтра. А затем используется оценка при помощи байесовской сети для комбинирования найденных признаков, таким образом происходит выборка черт лица. Данный метод в качестве достоинств имеют возможность распознавать лицо в различных положениях. Но даже при небольшом загромождении лица другими объектами, возникновении шумов или засветке процент достоверного распознавания сильно падает. Большая изменчивость объекта распознавания, зависимость вида лица на изображении от условий съемки и освещения позволяют без колебаний отнести обнаружение лица на изображении к задачам высокой сложности.

### **1.3 Активные модели внешнего вида**

Active Appearance Models (AAM) и Active Shape Models (ASM). — это статистические модели изображений, которые с помощью разных деформаций могут быть преобразованы для анализа реальных изображений. Этот тип моделей в двумерном варианте предложил Тим Кутс и Крис Тейлор в 1998 году.[] Изначально ААМ применялись для того, чтобы оценить параметры изображений лиц, такие как параметры формы и параметры внешнего вида. Перед использованием модель обучают на множестве заранее размеченных изображений. Такие изображения размечают вручную. Каждая метка имеет свой

номер и определяет характерную точку, которую алгоритм находит во время адаптации к новому изображению [20]. Процедура обучения активных моделей внешнего вида начинается с нормализации форм на размеченных изображениях для того, чтобы компенсировать различия в наклоне, смещении и масштабе. Активные модели внешнего вида состоят из набора параметров, одни из которых представляют форму лица, а другие задают его текстуру. При решении задачи обнаружения лица на фотографии выполняется поиск параметров, таких как форма, текстура и расположение, которые представляют синтезируемое изображение, наиболее близкое к наблюдаемому. По степени близости активных моделей внешнего вида исследуемому изображению принимается решение о наличии либо отсутствии лица на нем [9]. Модель лица, используемая в этих методах, показана на рисунке 1.

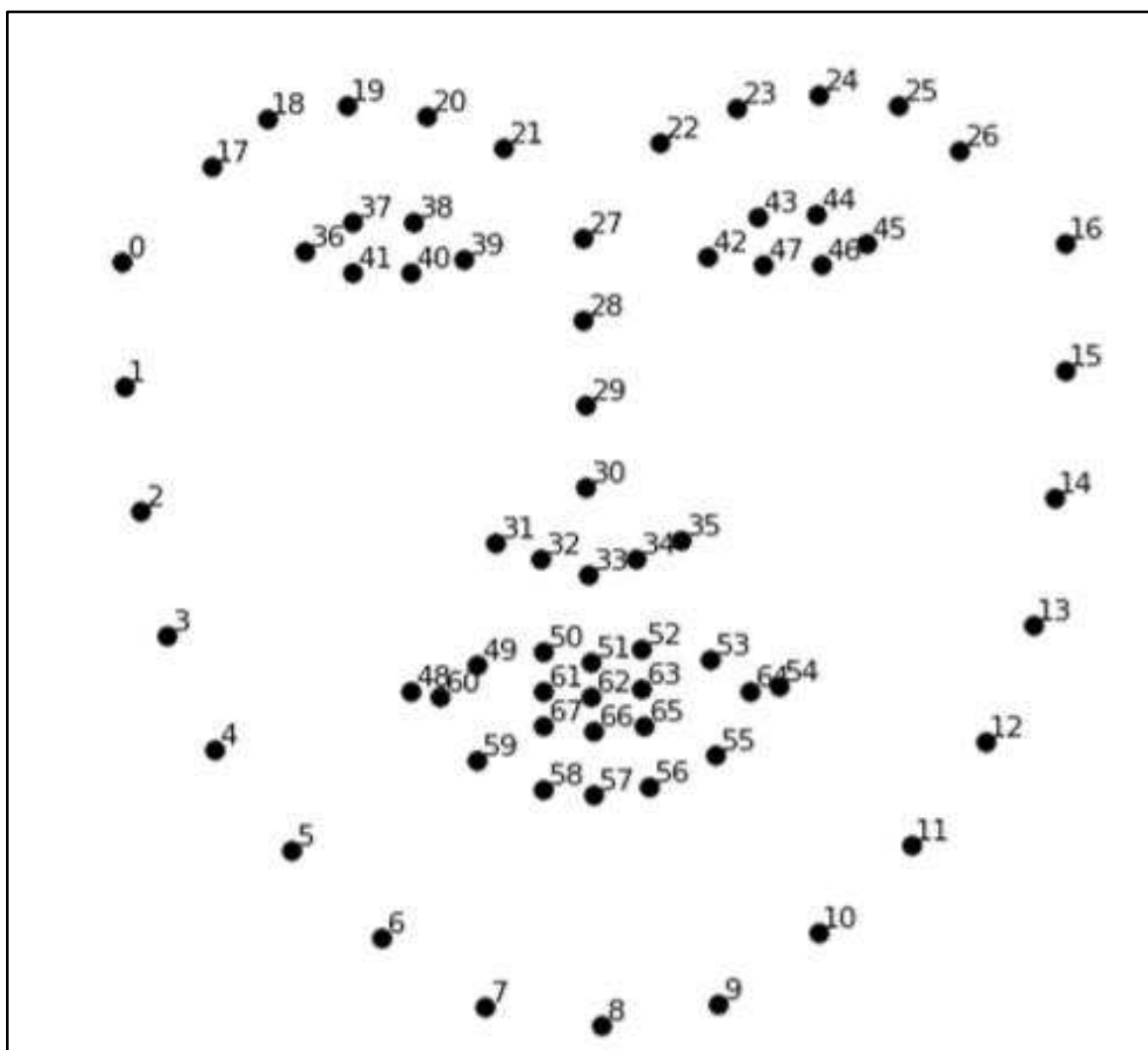


Рисунок 1 – Модель лица

Приведенный выше алгоритм достаточно эффективен, однако он имеет достаточно серьезный недостаток, который ограничивает его применение в приложениях реального времени: он медленно сходится и требует большого объема вычислений.

#### **1.4 Метод главных компонент**

Метод главных компонент (Principal component method) первоначально использовался для снижения пространства признаков без существенной потери информации. В задаче распознавания лиц его применяют главным образом для представления изображения лица вектором малой размерности (главных компонент), который сравнивается затем с эталонными векторами, заложенными в базу данных. Основной целью метода главных компонент является значительное уменьшение размерности пространства признаков таким образом, чтобы оно как можно лучше описывало «типичные» образы, принадлежащие множеству лиц. Используя этот метод можно выявить различные вариации в обучающей выборке изображений лиц и описать эту изменчивость в базисе нескольких ортогональных векторов, которые называются собственными (eigenface). Полученный один раз на обучающей выборке изображений лиц, набор собственных векторов используется для кодирования всех остальных изображений лиц, которые представляются взвешенной комбинацией этих собственных векторов. Используя ограниченное количество собственных векторов можно получить сжатую аппроксимацию входному изображению лица, которую затем можно хранить в базе данных в виде вектора коэффициентов, служащего одновременно ключом поиска в базе данных лиц. На первом этапе весь обучающий набор лиц преобразуется в одну общую матрицу данных, где каждая строка представляет собой один экземпляр изображения лица. Все лица обучающего набора приводятся к одному размеру, гистограммы нормируются.

Метод главных компонент хорошо зарекомендовал себя в практических приложениях. Однако, в тех случаях, когда на изображении лица присутствуют значительные изменения в освещенности или выражении лица, эффективность метода значительно падает.

### 1.5 Метод гибкого сравнения на графах

Суть метода сводится к эластичному сопоставлению графов, описывающих изображения лиц. Лица представлены в виде графов со взвешенными вершинами и ребрами. На этапе распознавания один из графов – эталонный – остается неизменным, в то время как другой деформируется с целью наилучшей соответствия первому. В подобных системах распознавания графы могут представлять собой как прямоугольную решетку, так и структуру, образованную характерными (антропометрическими) точками лица[15].

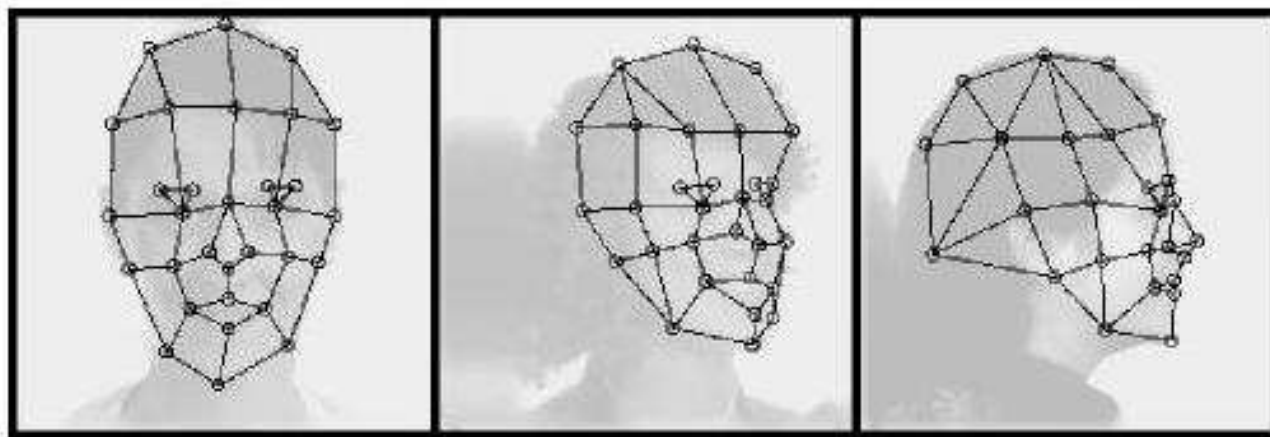


Рисунок 2 – Характерные точки лица

В вершинах графа вычисляются значения признаков, для этого чаще всего используют комплексные значения фильтров Габора или их упорядоченных наборов – Габоровских вейвлет (строи Габора), которые вычисляются в некоторой локальной области вершины графа путем свертки значений яркости пикселей с фильтрами Габора[15].



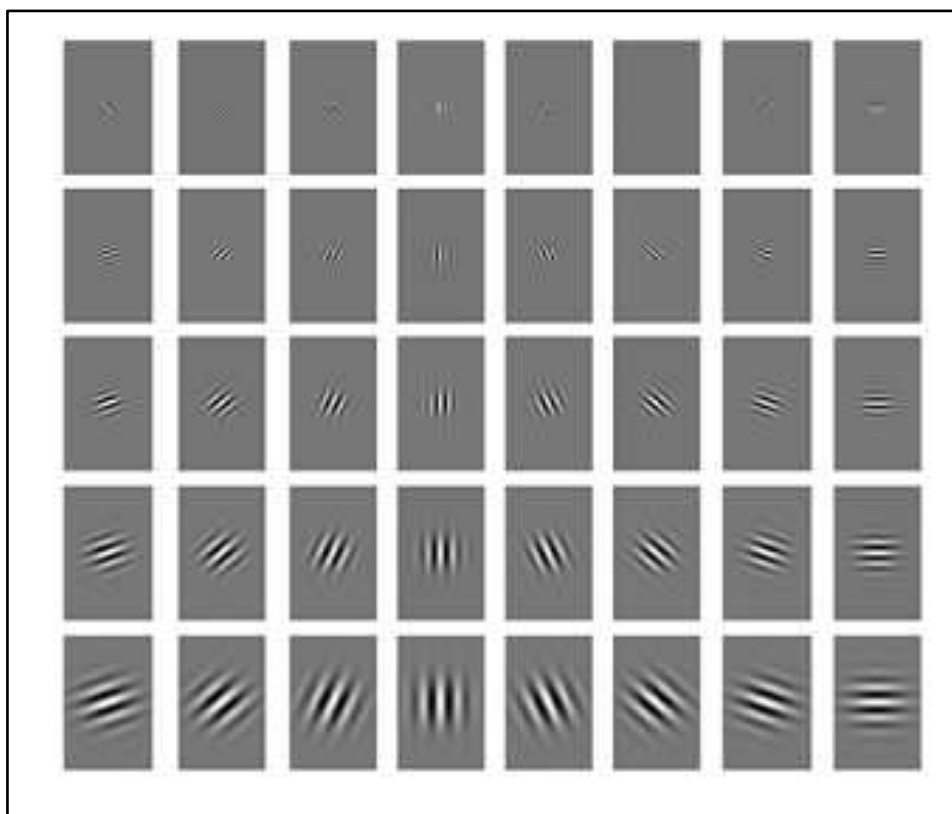


Рисунок 3 – набор фильтров Габора

Ребра графа взвешиваются расстояниями между смежными вершинами. Различие (расстояние, дискриминационная характеристика) между двумя графами вычисляется при помощи некоторой ценовой функции деформации, учитывающей как различие между значениями признаков, вычисленными в вершинах, так и степень деформации ребер графа.

Деформация графа происходит путем смещения каждой из его вершин на некоторое расстояние в определённых направлениях относительно ее исходного местоположения и выбора такой ее позиции, при которой разница между значениями признаков (откликов фильтров Габора) в вершине деформируемого графа и соответствующей ей вершине эталонного графа будет минимальной. Данная операция выполняется поочередно для всех вершин графа до тех пор, пока не будет достигнуто наименьшее суммарное различие между признаками деформируемого и эталонного графов. Значение ценовой функции деформации при таком положении деформируемого графа и будет являться мерой различия между входным изображением лица и эталонным графом.

Пример ценовой функции деформации графа  $G^I$ , построенного для исходного изображения лица, с некоторым графом  $B$ , представлен в формуле:

$$S_B(G^I, B) = \frac{1}{N} \sum_n \max S_f(J_n^I, J_n^B) - \frac{\lambda}{E} \sum_e \frac{(\Delta x_e^I - \Delta x_e^B)^2}{(\Delta x_e^B)^2}$$

Результат распознавания системы – эталон с наилучшим значением ценовой функции деформации.

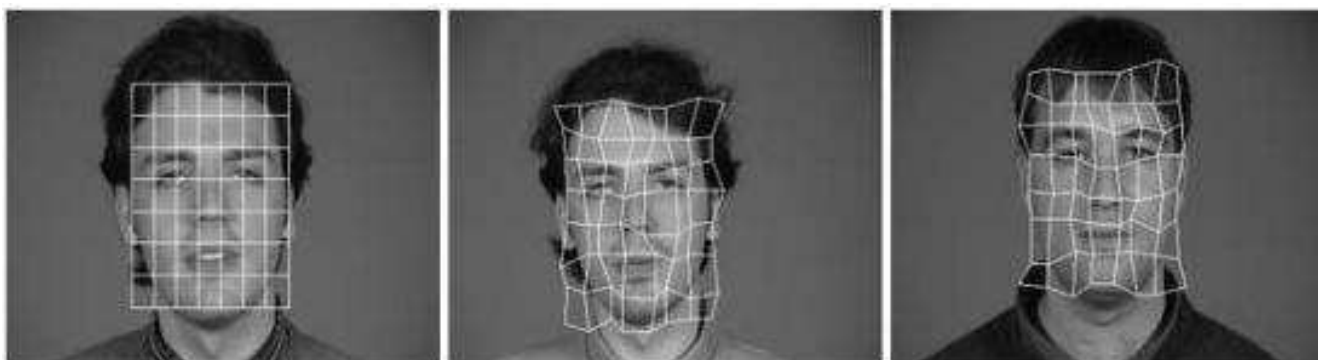


Рисунок 4 – Пример деформации графа

Данным метод эффективен для распознавания даже при наличии различных эмоциональных выражений и изменении ракурса лица до 15 градусов. Недостатками данного метода является высокая вычислительная сложность процедуры, низкая технологичность при запоминании новых эталонов, линейная зависимость времени работы от размера базы данных лиц.

## 1.6 Скрытые Марковские Модели

Скрытые Марковские Модели (СММ) (Hidden Markov Models, HMM) [12] могут использоваться как один из способов создания математической модели, описывающей свойства некоторого регистрируемого сигнала, при этом СММ являются стохастическими моделями, т. е. они стараются охарактеризовать исключительно статистические свойства сигнала, не имея информации о его

особенностях. В основе стохастических моделей лежит допущение о том, что сигнал можно описать некоторым случайным параметрическим процессом, при этом параметры данного процесса можно достаточно точно оценить, используя некоторый вполне определённый способ. Настроенную СММ можно представить в качестве источника некоторого случайного сигнала с определёнными характеристиками; притом возможно рассчитать вероятность генерации тестового сигнала с использованием данной модели.

Применительно к задаче распознавания, если представить вектор признаков объекта в виде конкретного сигнала (набора последовательных измерений), то возможно моделировать класс объектов с использованием СММ. При этом вероятность принадлежности тестового объекта к классу, заданному СММ, оценивается вероятностью генерации сигнала, который соответствует его вектору признаков.

Чтобы использовать СММ для решения задач обнаружения лиц, нужно выбрать способ, преобразования изображений лиц в сигнал (набор последовательных наблюдений). Естественные особенности структуры лица позволяют выделить несколько участков, например рот, лоб, подбородок и глаза, при выделении горизонтальных областей. Лицо человека может быть переведено в сигнал, передающий данные области в строго определённом порядке (чаще всего сверху вниз и слева направо). Так, изображение лица становится представленным в виде строгой последовательности наблюдений векторов, при этом каждый из векторов является горизонтальной полосой пикселей лица, последовательно передающихся случайному процессу, который моделируется СММ во время распознавания и тренировки [12]. Недостатки метода Скрытые Марковские Модели: необходимо подбирать параметры модели для каждой базы данных; не обладает хорошей различающей способностью, то есть алгоритм обучения, только максимизирует отклик каждой модели на свои классы или отклик каждого изображения на свою модель, но не минимизирует отклик на другие классы или на другие модели, и не выделяются ключевые признаки,

отличающие один класс от другого. Таким образом, похожие классы могут оказаться слабо различимыми. СММ невысокая надежность распознавания.

### **1.7 Выводы по 1 главе**

Первая глава посвящена исследованию методов обнаружения и распознаванию лиц, таким как:

- эмпирический метод;
- методы характерных инвариантных признаков;
- активные модели внешнего вида;
- метод главных компонент;
- скрытые Марковские модели;
- метод гибкого сравнения на графах.

Для каждого метода обнаружения и распознавания лиц, произведен анализ, так же приводится оценка их достоинств и недостатков.

## 2 Разработка системы распознавания лиц по визуальным данным

Первым этапом распознавания лиц является обнаружение лица на видеопоследовательности или изображении, то есть классификация «один против всех». В таком случае объекты изображения делятся на два класса «лицо», «не лицо».

### 2.1 Метод Виолы-Джонса для обнаружения лиц

На первом этапе выполняется обнаружение лиц по видеоданным использован метод Виолы – Джонса. Метод Виолы-Джонса разработан и представлен Полом Виолой и Майклом Джонсом [10]. Особенностью этого метода является обработка видеопотока в режиме реального времени. Алгоритм Виолы-Джонса позволяет классифицировать различные объекты, но основной областью его применения в настоящее время является распознавание лиц [11, 12]. В основе работы метода Виолы-Джонса лежит выделение признаков, подобных признакам Хаара и использование каскадной модели классификации. Особенностью метода Виолы – является работа с интегральным способом представления изображения представленная на рисунке 3.

Интегральное представление изображения – это матрица, которая совпадает по размерам с исходным изображением. Каждый ее элемент хранит сумму интенсивностей всех пикселей, находящихся левее и выше данного элемента. Элементы матрицы рассчитываются по следующей формуле:

$$L(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j), \quad (1)$$

где  $I(i, j)$  – яркость текущего пиксела исходного изображения,  $(i, j)$  – координаты текущего пиксела.

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

Расчет матрицы возможен по формуле:

$$L(x,y) = I(x,y) - L(x-1,y-1) + L(x,y-1) + L(x-1,y) \quad (2)$$

По такой интегральной матрице можно быстро вычислить сумму пикселей произвольного прямоугольника и произвольной площади

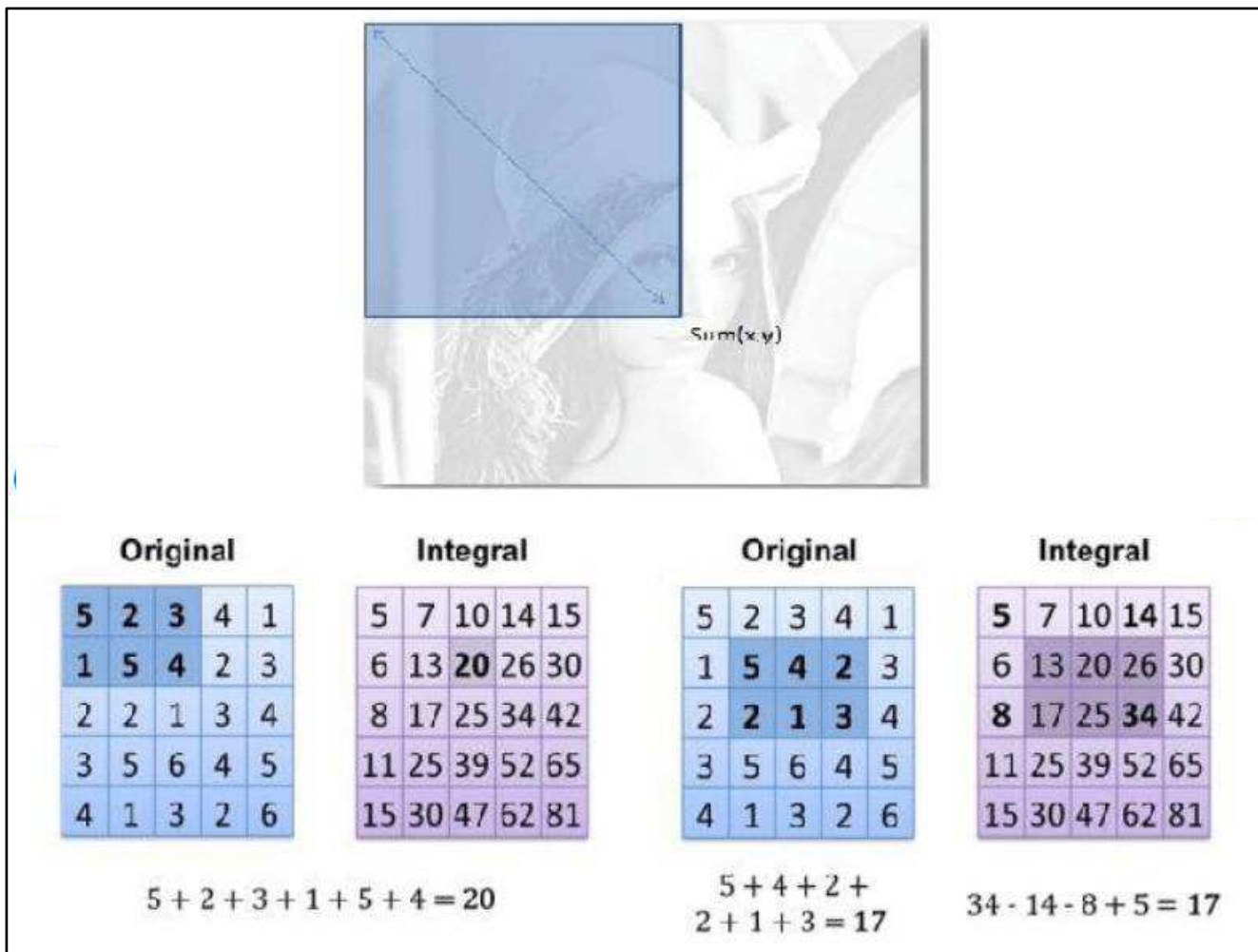


Рисунок 5 – Пример работы с интегральным способом представления изображения

Вычисление признаков Хаара выполняется следующим образом. По полутоновому изображению перемещается сканирующее окно, состоящее из смежных прямоугольников – примитивов Хаара (рис. 6). Далее суммируются интенсивности пикселей в этих областях, вычисляется разность между суммами.

Эта разность и определяет значение признака в заданной точке изображения. Известно, что для лица человека на изображении области в районе глаз темнее чем в районе щек. Выделяя такие характерные особенности изменения интенсивности пикселей для лица, становится возможным отделить его от других объектов изображения. Таким образом, путем перемещения сканирующего окна по всему изображению вычисляются признаки Хаара, дающие точечное значение перепада яркости пикселей области. Затем, по порогу выполняется отнесение исследуемой области к классам «лицо», «нет лица».

Вектор признаков, полученный при обходе сканирующим окном изображения лица с использованием интегрального представления изображения на основе примитивов Хаара вычисляется, согласно формуле:

$$F=X - Y, \tag{3}$$

где  $X$  – сумма значений яркостей точек закрываемых светлой частью признака, а  $Y$  – сумма значений яркостей точек закрываемых темной частью признака.

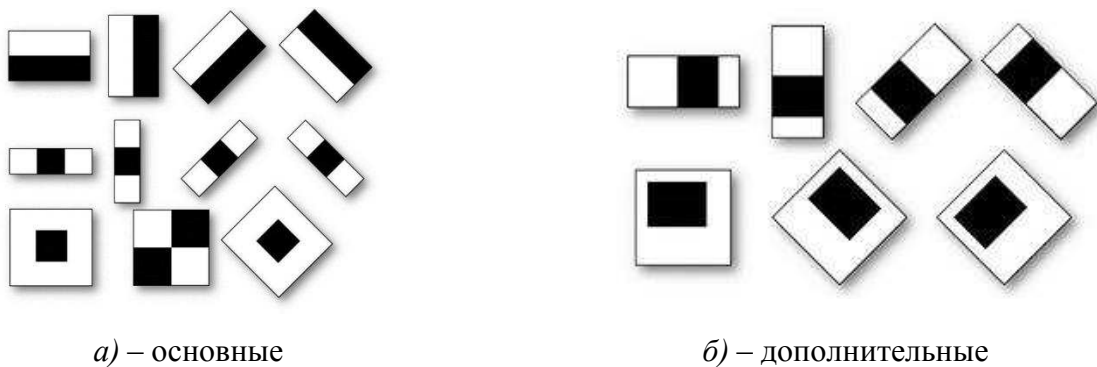


Рисунок 6 – Примитивы Хаара

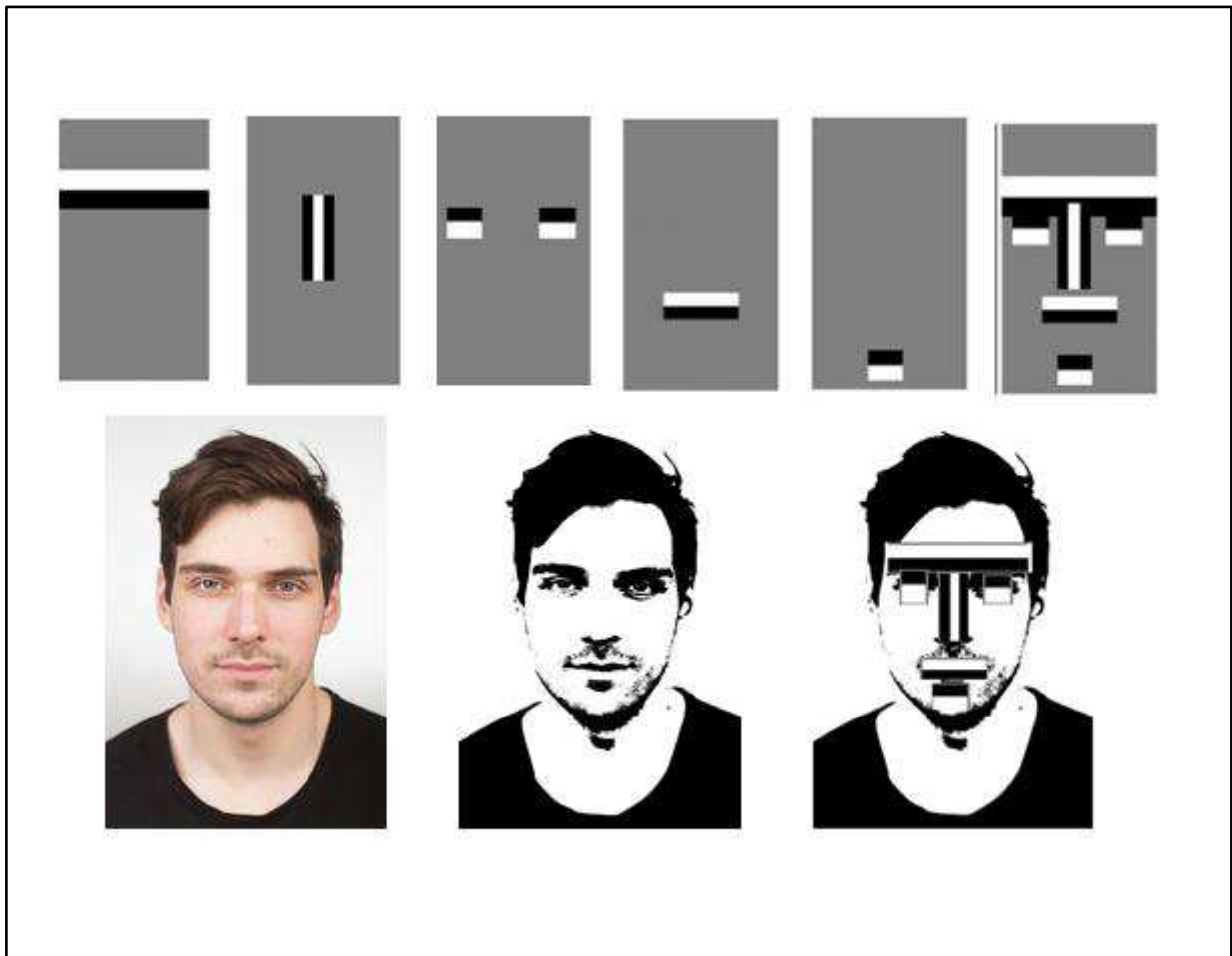


Рисунок 7 – Маска лица с использованием примитивов Хаара

Признаки Хаара в алгоритме Виолы-Джонса организуются в каскадный классификатор. Идея каскадной классификации заключается в постепенном увеличении количества применяемых примитивов, что увеличивает скорость работы алгоритма, отсеивая регионы, в которых лицо не найдено на ранних этапах (рис. 8).





Рисунок 8 – Каскадная классификация в методе Виолы-Джонса.

На основе полученных данных формируются векторы признаков. Далее представлены шаги разработанного алгоритма Виолы-Джонса для решения задачи обнаружения лиц.

Шаг 1. Получение видеопоследовательности (возможно использование данных с веб-камеры).

Шаг 2. Преобразование видеопоследовательности в серию видеоизображений с последующим выбором кадра для работы.

Шаг 3. Преобразование цветного изображения в полутоновое.

Шаг 4. Поиск лица путем выделения признаков Хаара методом вычисления характерных перепадов яркостей пикселей.

Шаг 5. На обнаруженном лице выполняется поиск глаз с использованием примитивов Хаара.

Шаг 6. Обозначение лица и глаз геометрической фигурой для вычисления показателей точности работы алгоритма.

Результатом работы алгоритма классификации Виолы-Джонса является набор признаков для каждой области, состоящий из 200 значений перепадов

яркостей, позволяющий отделить изображения, содержащие лицо, от изображений без него.

Обнаружение лиц может выполняться при повороте головы. Поэтому в работе использовались дополнительные примитивы Хаара, которые помогают обнаруживать лица при большем угле поворота до 55 градусов(рис).

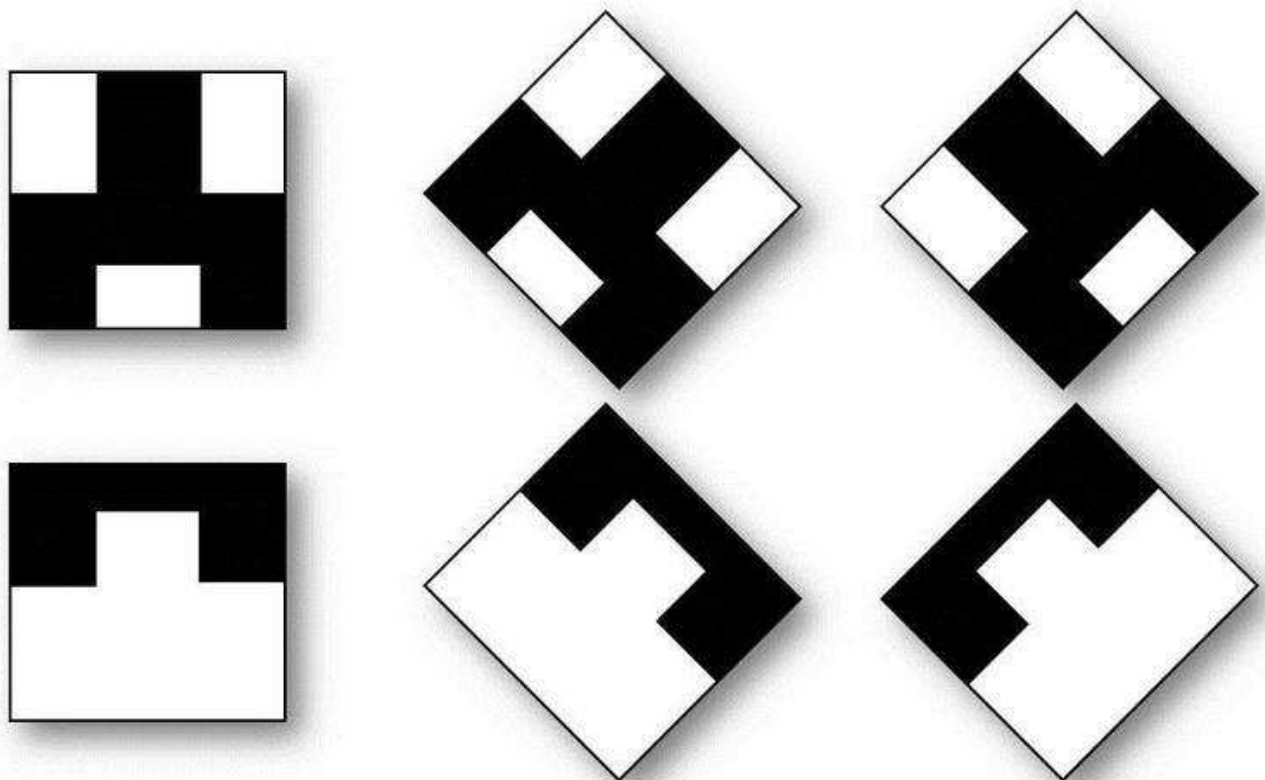


Рисунок 9 – дополнительные примитивы Хаара

Таким образом, настоящей работе предложен алгоритм выделения лиц по видеоизображениям с помощью метода Виолы-Джонса. При работе алгоритма вычисляются признаки Хаара, интегральное представление изображения, классификатор, используется каскадная классификация.

Для обнаружения лиц при угле поворота головы на изображении более 30 градусов использованы дополнительные маски Хаара для лица в профиль. Особенностью применения метода Виолы – Джонса для решения задачи обнаружения является возможность масштабирования и поворота самого изображения лица без влияния на точность обнаружения.

## 2.2 Локальные бинарные шаблоны для распознавания лиц

Для распознавания лиц используются локальные бинарные шаблоны (ЛБШ). Впервые ЛБШ были предложены Т.Оjala в 1994 году [20]. Базовый оператор ЛБШ представляет собой описание окрестности пиксела радиуса  $R$ . Количество пикселей окрестности  $P$  зависит от  $R$  и определяется формулой.

$$P = (2R + 1)^2 - 1, \quad (4)$$

где  $P$  – количество пикселей окрестности,  $R$  – радиус окрестности.

Оператор ЛБШ, который применяется к окрестности пиксела изображения, использует  $P$  пикселей окрестности, принимая яркость центрального пиксела в качестве порогового значения[23].

Пикселы, которые имеют значения яркости больше, чем значение яркости центрального пиксела (или равные ему), принимают значения «1»; те пикселы, которые имеют меньшие значения относительно центрального пиксела, принимают значения «0». В результате получается бинарный код, который описывает окрестность заданного пикселя[23].

При расчете шаблона изображение разбивается на непересекающиеся блоки. Формально обобщенный оператор ЛБШ можно представить в виде:

$$LBP_R(P) = \sum_{n=0}^{P-1} S(I_n - I_c) \cdot 2^n, \quad (5)$$
$$S(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

где  $I_i$  и  $I_c$  – значения яркости текущего и центрального пикселей, представляющие собой значения яркости  $Y$  из цветовой схемы YUV.

Прямое преобразование YUV:

$$Y = 0,299R + 0,587G + 0,114B \quad (6)$$

$$U = -0,16873R - 0,331264G + 0,5B + 128 \quad (7)$$

$$V = 0,5R - 0,418688G - 0,081312B + 128 \quad (8)$$

Обратное преобразование:

$$R = Y + 1,402(U - 128)$$

$$G = Y - 0,7141(U - 128) - 0,34414(V - 128)$$

$$B = Y + 1,772(V - 128),$$

где  $Y$  – яркостная компонента,  $U$ ,  $V$  – цветоразностные компоненты.

Формирование бинарного кода LBP выполняется следующим образом: если для текущего пиксела значение яркости  $Y$  меньше центрального, то текущему разряду LBP присваивается значение «1», в противном случае значение разряда бинарного кода локальных бинарных шаблонов равно «0». Таким образом, получается  $P$ -разрядный бинарный код, который описывает окрестность пиксела.

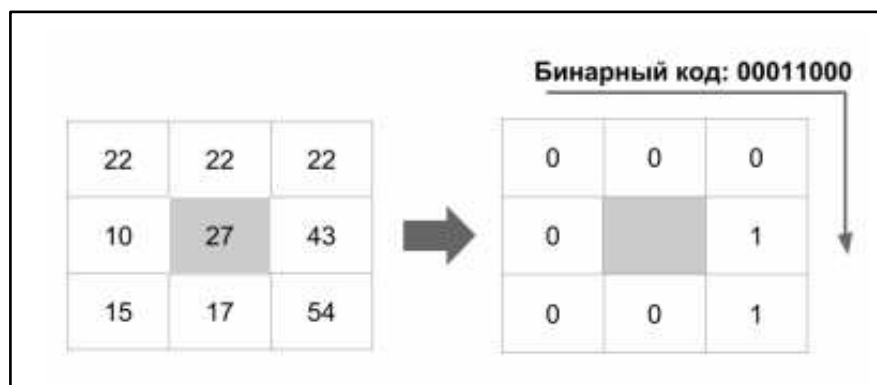


Рисунок 10 – Пример вычисления оператора локального бинарного шаблона

Для описания изображения в пространстве признаков ЛБШ используют гистограмму, столбцы которой соответствуют десятичным значениям бинарных кодов шаблонов[17].

Изображение разделяется на регионы, в последствии, объединив которые, мы получим полное описание изображения.

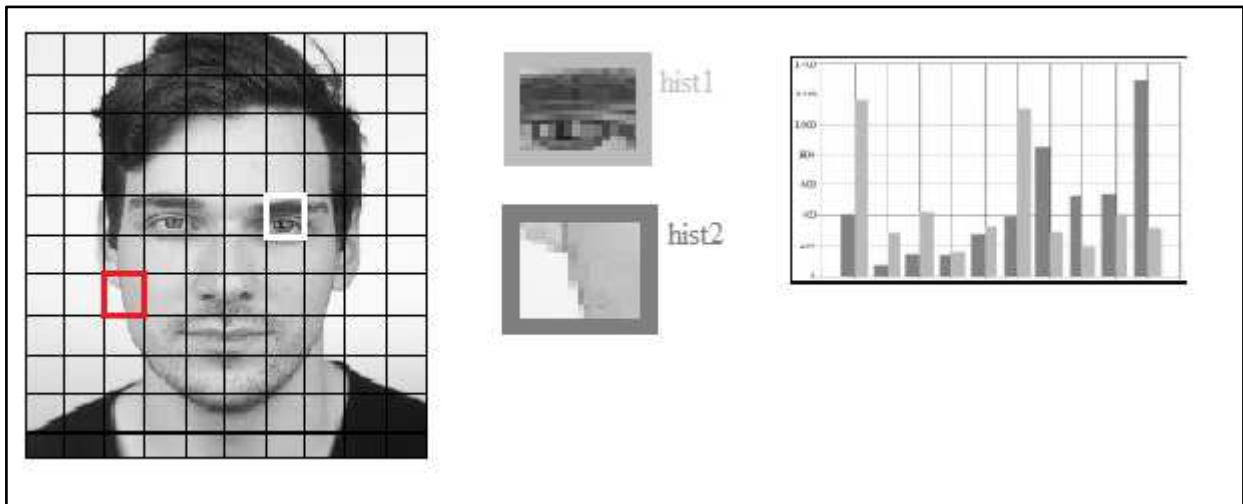


Рисунок 11 – Фрагмент гистограммы

Способ построения гистограммы, заключается в следующем, каждый бинарный код преобразуется в десятичное число, подсчитывается количество одинаковых чисел, определяя положение и высоту столбцов гистограммы. Таким образом, гистограмма каждого региона будет содержать 256 значений для описания текстурных особенностей каждого региона[21]. Итоговая гистограмма для исследуемого изображения лица получается путем конкатенации гистограмм отдельных блоков (рис 12).

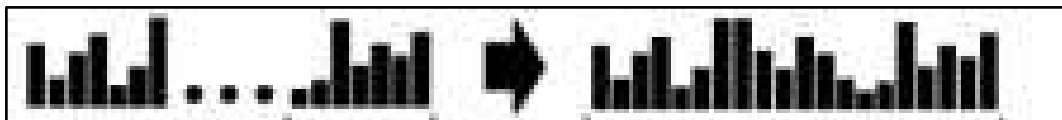


Рисунок 12 – конкатенация гистограмм

Блок-схема алгоритма применения оператора локальных бинарных шаблонов для задачи распознавания лица представлена на рисунке 13.

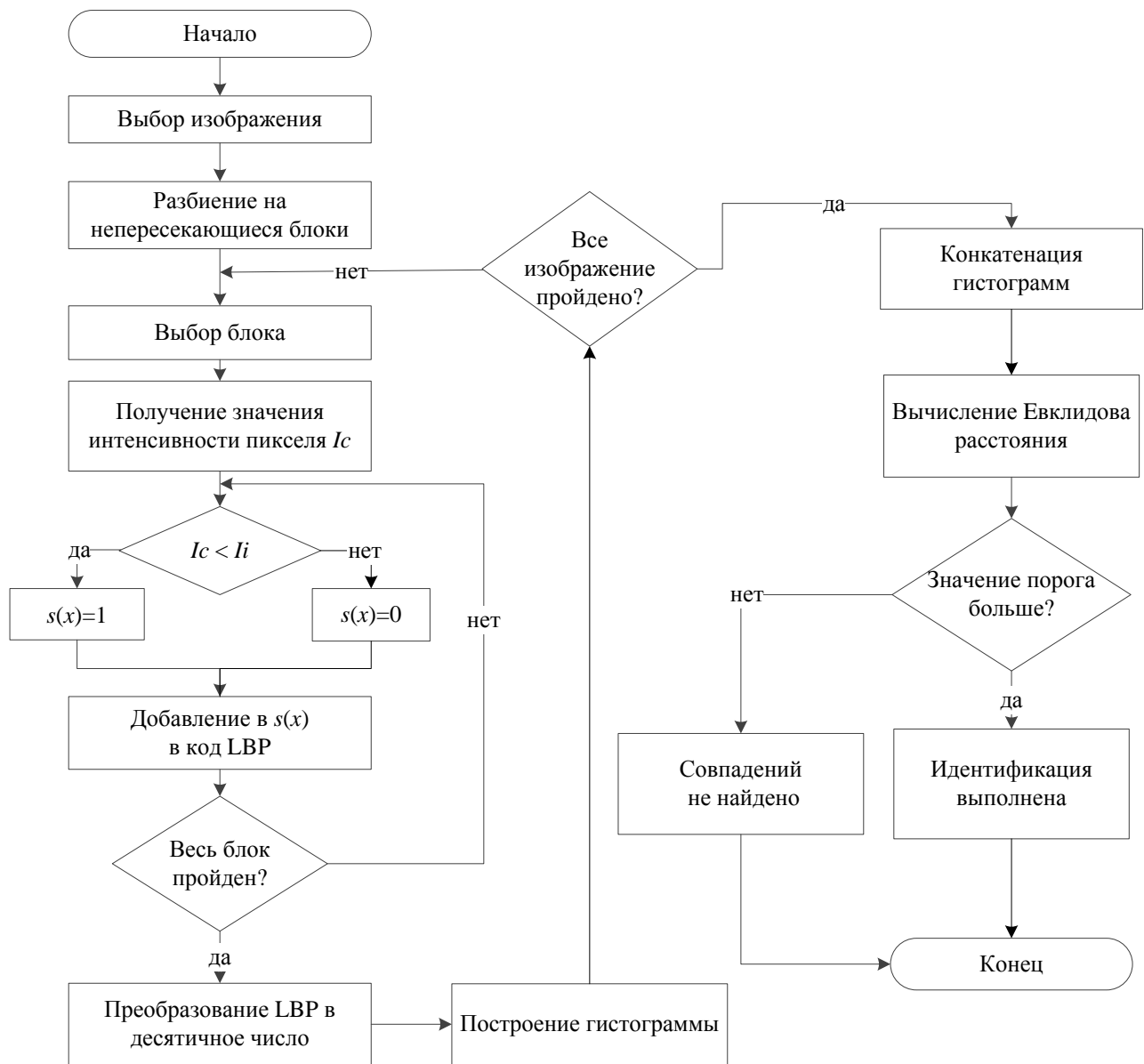


Рисунок 13 – блок схема алгоритма

В качестве меры различия гистограмм, вычисляется евклидово расстояние на основе данной формулы:

$$D = \sqrt{\sum_{i=1}^n (hist1_i - hist2_i)^2}, \quad (9)$$

где,  $hist1_i$  – столбец с номером  $i$  гистограммы исследуемого изображения лица, где  $hist2_i$  – столбец с номером  $i$  гистограммы изображения из базы эталонов лиц,  $n$  - количество столбцов гистограмм.

Таким образом, объединенная гистограмма фрагментов лица сравнивается по порогу с каждой из эталонных гистограмм, на основе этого сравнения выполняется идентификация пользователя.

Шаг1. Разбиение изображения на непересекающиеся блоки.

Шаг2. В каждом блоке производится расчет окрестности пиксела.

Шаг3. Вычисляется бинарная строка.

Шаг4. Преобразование бинарной строки в десятичное число.

Шаг5. Построение гистограммы каждого блока изображения, где подсчитывается количество одинаковых чисел, определяя положение и высоту столбцов гистограммы.

Шаг6. Конкатенация гистограмм каждого блока.

Шаг8. В качестве меры различия гистограмм рассчитываем евклидово расстояние.

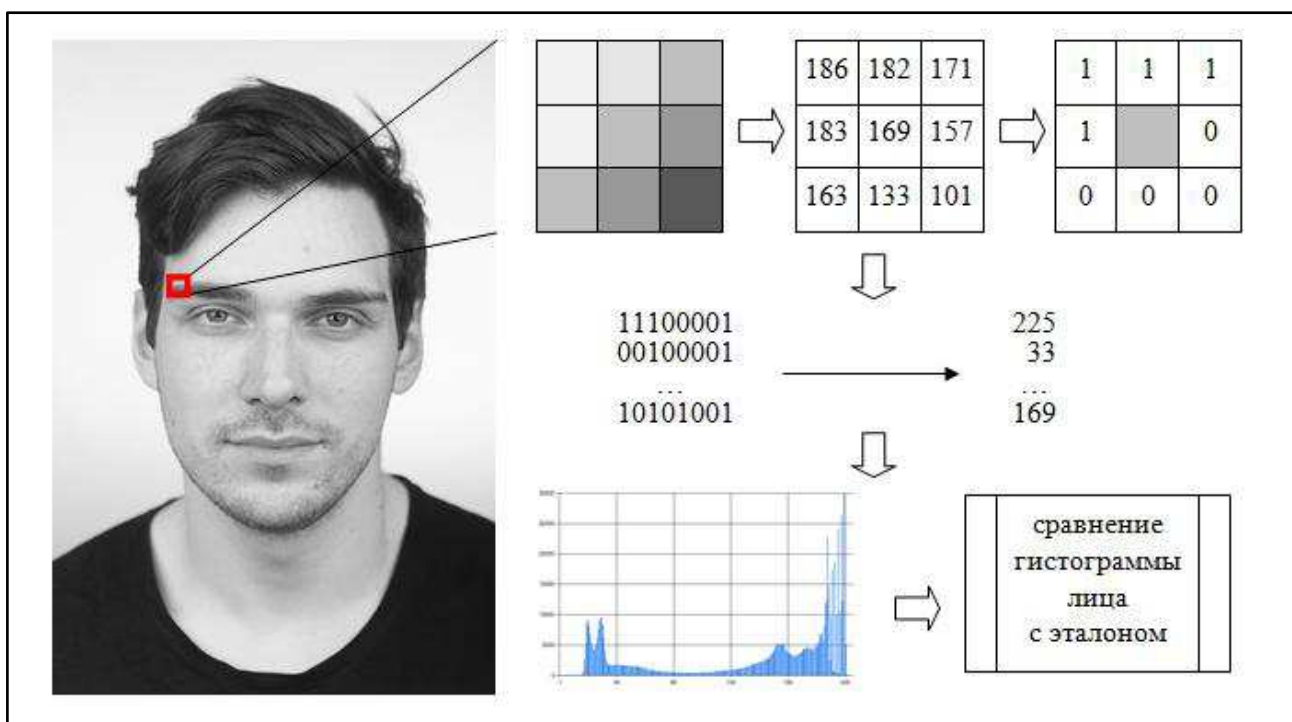


Рисунок 14 –применение ЛБШ

## 2.3 Выводы по 2 главе

В данной главе были представлены решения задач обнаружения и распознавания лиц. Для решения задачи обнаружения лиц был выбран метод «Виолы-Джонса». Для решения задачи идентификации лиц был выбран метод Локальных Бинарных Шаблонов. Произведен подробный разбор методов.

Таким образом, задача распознавания лиц включает в себя два этапа.

На первом этапе происходит обнаружение лиц путем применения метода Виолы – Джонса, основу которого составляют применение заранее обученных признаков Хаара, сканирующего окна и использование каскадной классификации. Так же, для обнаружения лиц в ситуациях, когда угол головы достигает поворота более 30 градусов используются дополнительные признаки Хаара, обученные на выборке лиц со схожим углом поворота.

На втором этапе происходит распознавание лиц с использованием метода ЛБШ, для распознавания вычисляются бинарные строки, описывающие локальные особенности изображения. После чего производится построение гистограмм, с целью упорядочить локальные особенности, и тем самым получить описание изображения. После чего вычисление евклидова расстояния между двумя гистограммами, полученными идентичным образом. Лучшей особенностью метода ЛБШ является достаточно точное распознавание при минимальных требованиях к подготовке и модифицированию изображения, по сравнению с другими аналогичными методами.



### 3 Экспериментальные исследования

Для проведения экспериментальных исследований использованы следующие наборы данных:

- Aberdeen включающая 4916 примеров лиц и 8500 примеров с отсутствием лиц, взятых из хранилища данных Labeled Faces in the Wild Home [13] и [14].

- «YouTube Faces (YTF) data set», база данных видео, предназначена для изучения проблемы распознавания лиц. Набор данных содержит 3425 видео 1595 разных людей. Все видеопоследовательности были загружены с YouTube. Самая короткая продолжительность видеопоследовательности составляет 48 кадров, самая длинная видеопоследовательность составляет 6070 кадров. В базе данных так же присутствуют видеопоследовательности, с различными уровнями освещенности, разным количеством объектов распознавания, что позволяет эффективно симулировать различные условия использования системы.  
<http://www.cs.tau.ac.il/~wolf/ytfaces/index.html#download>.

- «McGillFaces» база данных видео содержит 18000 видеокадров с разрешением 640x480 из 60 видеопоследовательностей, каждая из которых записана на разные темы (31 женщина и 29 мужчина). Каждое видео было собрано в разных условиях (в помещении или на улице), что привело к произвольным условиям освещения и фоновой помехе. Кроме того, субъекты были совершенно свободны в своих движениях, что приводило к произвольным масштабам лица, произвольным выражениям лица, позе головы.  
<https://sites.google.com/site/meltemdemirkus/mcgill-unconstrained-face-video-database>.

- «NRC-ИТ Лицевая Видео База Данных» содержит 22 видеоклипа, каждый из которых показывает людей проходящих под камерой с различным поворотом головы и разным количеством объектов. Разрешение захвата видео

поддерживается на уровне 160 x 120.

<http://www.videorecognition.com/db/video/faces/cvglab/>

### 3.1 Обнаружение лиц

Для проведения экспериментальных исследований обнаружения лиц выбраны наборы данных, которые описанные выше. Обучающая выборка составила 80 %, тестовая выборка – 20 % от общей выборки. Примеры изображений приведены на рисунке 15. Кроме того, авторами проведены экспериментальные исследования с использованием данных, полученных с web-камеры.

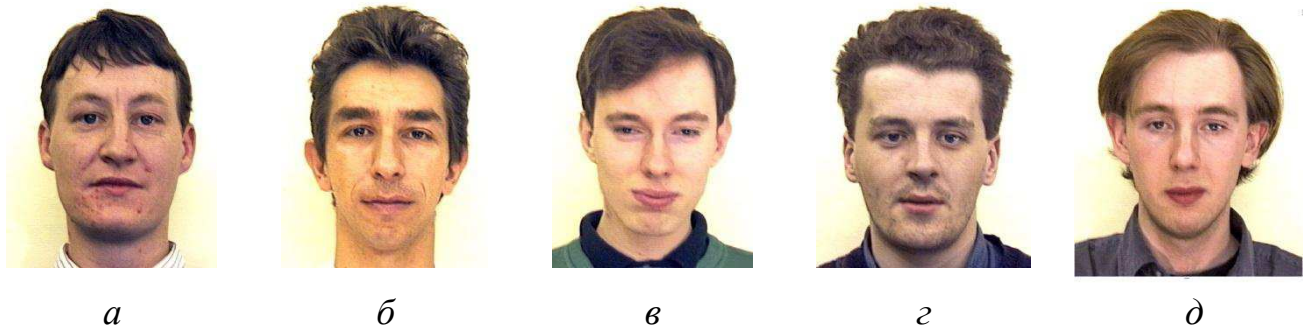
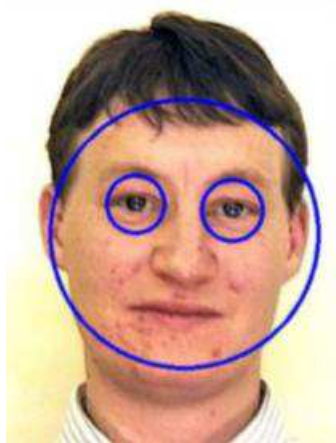


Рисунок 15 – Используемые изображения людей :

*а* – dhands.jpg; *б* – jim.jpg; *в* – iroy.jpg; *г* – paul.jpg; *д* – dpearson.jpg.

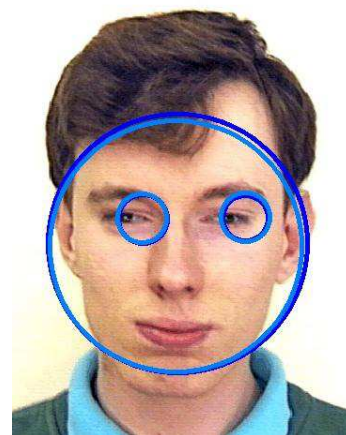
Примеры из обучающей выборки содержали лица людей, с различным наклоном и поворотом головы, в различных условиях освещения. Минимальный размер изображения обучающей выборки составил 24×24 пиксела, максимальный - 624×544. Результаты работы алгоритма Виолы-Джонса показаны на рисунке 16 и таблице 1.



*a*




*б*



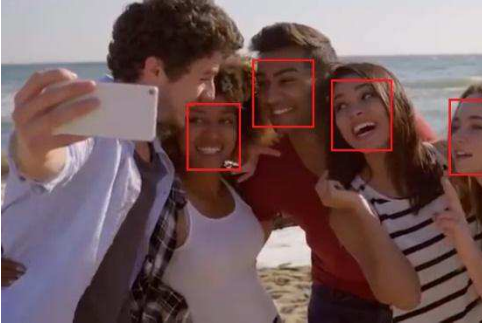
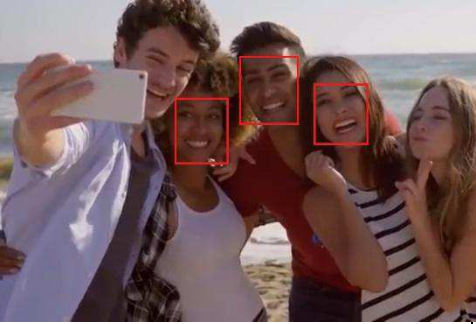
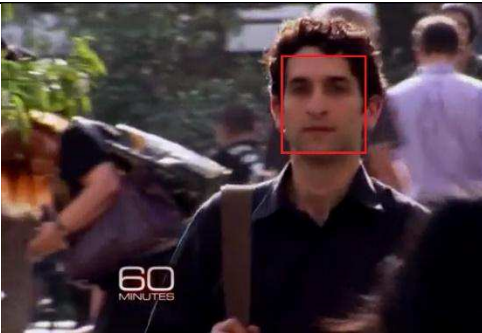
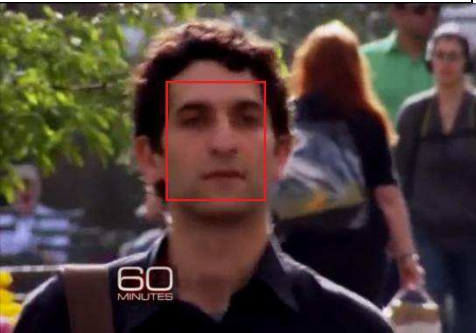
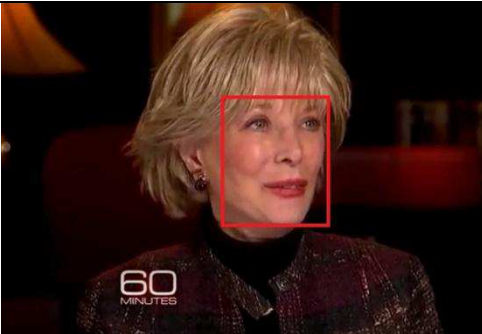
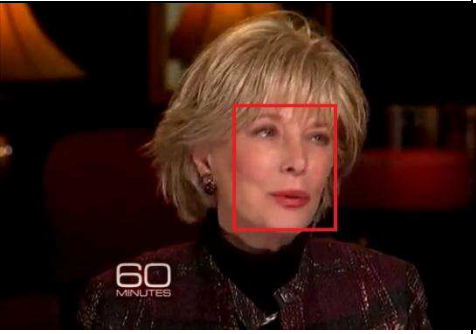
*в*

Рисунок 16 –Результат обнаружения лица: *a* – dhands.jpg; *б* – irene.jpg; *в* – iroy.jpg.






Таблица 1 – примеры работы обнаружения лиц

Описание тестовой видеопоследовательности	Обнаруженные лица на видеопотоке
YouTubeFaces\PIE_S1_C2.mp4, Количество кадров: 375, Разрешение: 640×480, Кол-во различных лиц: 1 мужское лицо.	



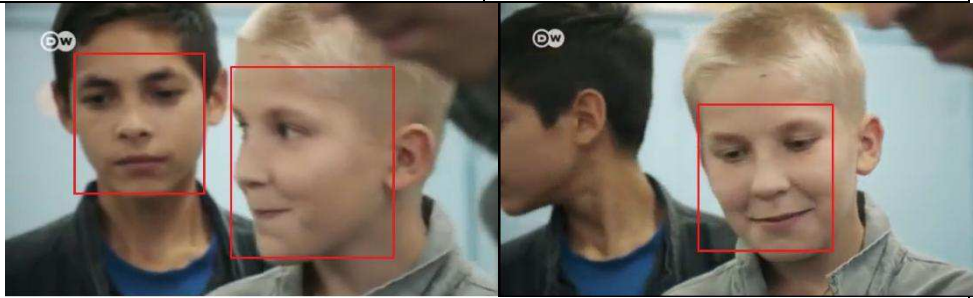
Продолжение таблицы 1

<p>YouTubeFaces\P1E_S1_ C5.mp4, Количество кадров: 300 Разрешение: 640×480, Кол-во различных лиц: 3 женских лица и 2 мужских лица.</p>		
<p>YouTubeFaces\P1E_S1_ C6.mp4, Количество кадров: 350 Разрешение: 640×480, Кол-во различных лиц: 1 мужское лицо.</p>		
<p>YouTubeFaces\P1E_S2_ M2.mp4, Количество кадров: 450 Разрешение: 1280×720, Кол-во различных лиц: 1 женское лицо.</p>		

Продолжение таблицы 1

<p>YouTubeFaces\P1E_S2_ M4.mp4, Количество кадров:550 Разрешение: 640×480, Кол-во различных лиц:1 женское и 1 мужское лицо.</p>		
<p>YouTubeFaces\P1E_S2_ M5.mp4, Количество кадров: 625 Разрешение: 640×480, Кол-во различных лиц: 1 женское лицо и 2 мужских лица.</p>		
<p>YouTubeFaces\P1E_S2_ M9.mp4, Количество кадров: 700 Разрешение: 640×480, Кол-во различных лиц:2 женских лица и 3 мужских лица.</p>		

## Окончание таблицы 1

<p>YouTubeFaces\P1E_S2_M12.mp4,          Количество кадров:775          Разрешение: 640×480,          Кол-во различных лиц:          10 женских лиц.</p>	
<p>YouTubeFaces\P1E_S2_D1.mp4,          Количество кадров:425          Разрешение: 640×480,          Кол-во различных лиц:1          женское и 1 мужское          лицо.</p>	
<p>YouTubeFaces\P1E_S2_D5.mp4,          Количество кадров: 500          Разрешение: 1280×720,          Кол-во различных лиц: 2          мужских лица.</p>	

Для проведения эксперимента необходимо было определить эффективность работы системы, подсчитать количество успешных срабатываний, и частоту возникновения ошибок. Для этого использовались следующие показатели:

TD (%) – точность обнаружения, FRR (%) –ложноположительная ошибка, FAR (%) –ложноотрицательная ошибка, для расчета данных показателей использовались следующие формулы:

$$TR = \frac{N_i}{N_t} \quad (10)$$

$$FAR = \frac{N_i - N_t}{N} \quad (11)$$

$$FRR = 1 - TR, \quad (12)$$

где  $N_t$  – количество кадров с лицом,  $N_i$  – количество кадров с лицом которое распознано,  $N$  – количество кадров в видеопоследовательности.

Результаты вычисления показателей представлены в таблице 2.

Таблица 2 – расчет показателей

Видеоролик	Обнаружение лиц		
	TD,%	FRR,%	FAR,%
YouTubeFaces\P1E_S1_C2	100	0	0
YouTubeFaces\P1E_S1_C5	98,2	1,20	17,8
YouTubeFaces\P1E_S1_C6	100	0	0
YouTubeFaces\P1E_S2_M2	100	0	0
YouTubeFaces\P1E_S2_M4	100	0	0
YouTubeFaces\P1E_S2_M9	100	0	0
YouTubeFaces\P1E_S2_M12	93,9	6,10	5,80
YouTubeFaces\P1E_S2_M15	100	0	0
YouTubeFaces\P1E_S2_D1	100	0	0
YouTubeFaces\P1E_S2_D5	100	0	0

Для видеопотоков, не содержащих лиц, предложенный алгоритм показал точность обнаружения, равную 100% при полном отсутствии ложных срабатываний. Для видеопотока, содержащих лицо, точность обнаружения составила 98%.

### 3.2 Распознавание лиц





Целью эксперимента было протестировать работу алгоритма распознавания лиц в различных условиях, с целью определить процент возникающих в процессе

работы ошибок алгоритма, а так же определить факторы, при которых достигаются наилучшие показатели, а так же факторы, влияющие на ухудшение работы.

Лицом будем считать, когда угол поворота головы достигает не более 55 градусов.







В таблице 3 представлены результаты проведенного эксперимента.

Таблица 3 – примеры видеопотоков



Описание тестовой видеопоследовательности	Образец кадра	Описание тестовой видеопоследовательности	Образец кадра
<p>YouTubeFaces\P1E_S1_C1.mp4,                      Количество кадров: 1125,                      Разрешение: 1280×720,                      Кол-во различных лиц: 2 женских лица.</p>		<p>YouTubeFaces\P1E_S1_C3.mp4,                      Количество кадров: 1250                      Разрешение: 1280×720,                      Кол-во различных лиц: 1 мужское лицо</p>	
<p>YouTubeFaces\P1E_S1_C4.mp4,                      Количество кадров: 1500                      Разрешение: 1280×720,                      Кол-во различных лиц: 1 женское.</p>		<p>YouTubeFaces\P1E_S2_M1.mp4,                      Количество кадров: 875                      Разрешение: 640×480,                      Кол-во различных лиц: 1 женское и 1 мужское лицо.</p>	







Продолжение таблицы 3

<p>YouTubeFaces\P1E_S2_M3.mp4,          Количество кадров: 750          Разрешение: 1280×720,          Кол-во различных лиц: 2 мужских лица.</p>		<p>YouTubeFaces\P1E_S2_M6.mp4,          Количество кадров: 1025          Разрешение: 480,          Кол-во различных лиц: 3 женских лица.</p>	
<p>YouTubeFaces\P1E_S2_M13.mp4,          Количество кадров: 850          Разрешение: 640×480,          Кол-во различных лиц: 5 женских лиц и 1 мужское лицо.</p>		<p>YouTubeFaces\P1E_S2_M7.mp4,          Количество кадров: 700          Разрешение: 640×480,          Кол-во различных лиц: 2 женских лица и 2 мужских лица.</p>	
<p>YouTubeFaces\P1E_S2_D3.mp4,          Количество кадров: 575          Разрешение: 480×320,          Кол-во различных лиц: 1 женское и 2 мужских лица.</p>		<p>YouTubeFaces\P1E_S2_D5.mp4,          Количество кадров: 900          Разрешение: 1280×720,          Кол-во различных лиц: 2 мужских лица.</p>	

Продолжение таблицы 3

<p>YouTubeFaces\P1E_S1_K1.mp4,          Количество кадров: 300          Разрешение: 480×320,          Кол-во различных лиц: 1          детское лицо.</p>		<p>YouTubeFaces\P1E_S1_K2.mp4,          Количество кадров: 400          Разрешение: 480×320,          Кол-во различных лиц: 1          детское лицо.</p>	
<p>McGillFacesDatabase\mmdm2\video\si2028.avi          Количество кадров: 575          Разрешение: 640x480,          Кол-во различных лиц: 2          мужских лица.</p>		<p>db\faces\crglad\000046.avi          Количество кадров: 300          Разрешение: 160x120,          Кол-во различных лиц: 1          мужское лицо.</p>	
<p>McGillFacesDatabase\mmdm2\video\si1425.avi          Количество кадров: 375          Разрешение: 640x480,          Кол-во различных лиц: 1          мужское лицо.</p>		<p>db\faces\crglad\000050.avi          Количество кадров: 300          Разрешение: 160x120,          Кол-во различных лиц: 1          мужское лицо.</p>	

### Окончание таблицы 3

<p>db\faces\crglad\00002.avi .avi</p> <p>Количество кадров: 300 Разрешение: 160x120, Кол-во различных лиц: 1 женское лицо.</p>		<p>McGillFacesDatabase\mmdm2\video\sx102.avi</p> <p>Количество кадров: 425 Разрешение: 640x480, Кол-во различных лиц: 1 мужское лицо.</p>	
<p>McGillFacesDatabase\mmdm2\video\sx372.avi</p> <p>Количество кадров: 400 Разрешение: 640x480, Кол-во различных лиц: 1 женское и 1 мужское лицо.</p>		<p>McGillFacesDatabase\mmdm2\video\sa1.avi</p> <p>Количество кадров: 325 Разрешение: 640x480, Кол-во различных лиц: 1 мужское лицо.</p>	

TR (%) - точность распознавания, FRR (%) –ложноположительная ошибка , FAR (%) –ложноотрицательная ошибка, для расчета данных показателей использовались следующие формулы:

$$TR = \frac{N_i}{N_t} \quad (13)$$

$$FAR = \frac{N_i - N_t}{N} \quad (14)$$

$$FRR = 1 - TR, \quad (15)$$

где  $N_t$  – количество кадров с лицом,  $N_i$  – количество кадров с лицом которое распознано,  $N$  – количество кадров в видеопоследовательности.

Алгоритм так же показывает достаточно высокую точность распознавания лиц, в ходе тестирования наилучшим результатом на видеопотоке было 100% , а наихудшим 88,7% , на ухудшение показателей распознавания повлияли следующие факторы:

- различные уровни освещенности (источники света, их цвет, яркость).

Например, когда освещение падает на одну часть лица, а другая часть лица темная из-за недостатка освещенности, то обнаружения лиц затрудняется, или становится невозможным;

- аксессуары, такие как очки, усы или борода, которые так же усложняют обнаружения лиц. Так же дополнительные погрешности в обнаружение лиц вносит закрытие частей лица, таким как волосы, шапки, шарфы и тд;

- зашумленность фото;

- эмоциональное выражение. В большинстве случаев эмоциональное выражение не влияет на результаты работы, однако некоторые из них способны вызвать трудности при распознавании, например, широкая улыбка или закрытые глаза.

Следующим шагом, для каждой выбранной видеопоследовательности, производилась процедура распознавания, описанная ранее TR, FRR, FAR представлены в таблице 4.

Таблица 4 – результаты вычисления

Видеопоток	Распознавание лиц		
	TR, %	FRR, %	FAR, %
YouTubeFaces\P1E_S1_C1	99,5	0,50	0,44
YouTubeFaces\P1E_S1_C3	100	0	0
YouTubeFaces\P1E_S1_C4	99,1	0,01	0,66
YouTubeFaces\P1E_S2_M1	97,5	2,50	1,10

#### Окончание таблицы 4

YouTubeFaces\P1E_S2_M3	96,2	4,00	3,80
YouTubeFaces\P1E_S2_M6	100	0	0
YouTubeFaces\P1E_S2_M13	87,9	12,1	11,7
YouTubeFaces\P1E_S2_M7	100	0	0
YouTubeFaces\P1E_S2_D3	100	0	0
YouTubeFaces\P1E_S2_D5	88,9	11,1	11,1
YouTubeFaces\P1E_S2_D6	98,2	1,80	1,78
db\faces\crglad\00002	100	0	0
YouTubeFaces\P1E_S1_K1	100	0	0
YouTubeFaces\P1E_S1_K2	100	0	0
McGillFacesDatabase\mmdm2\video\ si2028	100	0	0
db\faces\crglad\000046	99,3	0,70	0,58
McGillFacesDatabase\mmdm2\video\si142 5	100	0	0
db\faces\crglad\000050	100	0	0
McGillFacesDatabase\mmdm2\video\sx372	100	0	0
McGillFaces Database\mmdm2\video\ sa1	97,8	2,20	2,14

### 3.3 Выводы по 3 главе

Проведены экспериментальные исследования с применением нескольких наборов данных, расчет точности обнаружения и точности распознавания, так же ложноположительные и ложноотрицательные ошибки. Как показали результаты экспериментальных исследований, пол, и возраст людей не влияют на качество работы алгоритма обнаружения и распознавания лиц. На качество работы алгоритма оказывают влияние такие факторы как, уровень освещенности сцены,

разрешение видеоролика, скорость перемещения людей на сцене, угол поворота лица и степень открытости лица. Так, дополнительную погрешность в работу алгоритма обнаружения и распознавания лиц вносят надетые на лицо аксессуары, например, очки, шарфы, шапки. Негативное влияние оказывает также закрывание части лица волосами, наличие бороды или усов. Эмоциональное выражение лица в большинстве случаев не влияет на результаты работы алгоритма, однако оно способно вызвать трудности при распознавании, например, при широкой улыбке или закрытых глазах человека. Кроме того, при затенении части лица качество работы алгоритма может снижаться.

## ЗАКЛЮЧЕНИЕ

В ходе исследования был произведен обзор современных методов распознавания и обнаружения лиц, описан принцип работы, их особенности, недостатки.

Решение задачи распознавания лиц на сегодняшний день актуально для реализации различного рода практических задач. В настоящей работе для выделения лица использован алгоритм Виолы-Джонса, средняя точность обнаружения составляет 98,00%, для распознавания лица применен метод локальных бинарных шаблонов, средняя точность распознавания составила 94,3%. Экспериментальные исследования, проведенные на разнородных видеоданных, подтверждают эффективность предложенных методов.

Таким образом, предложенные в работе алгоритмы позволяют использовать их для решения практических задач идентификации лиц, как на служебных объектах, городских территориях, так и в сети Интернет на различных интернет ресурсах.

## СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Местецкий Л. М., «Математические методы распознавания образов», МГУ, ВМиК, Москва, 2002–2004., с. 42 – 44
2. Журавлев, Ю.И. Об алгебраическом подходе к решению задач распознавания и классификации / Ю.И. Журавлев // Проблемы кибернетики. — 1978. — Т. 33. С. 5–68
3. Bledsoe W.W. Some Results on Multicategory Pattern Recognition // J. ACM. 1966. No.13. P. 304-316.
4. Turk M., Pentland A. . Eigenfaces for recognition // J. Cognit. Neurosci. 1991. No. 3 (1). P. 71-86.
5. Belhumeur P.N., Hespanha J.P., Kriegman D.J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection // IEEE Trans. Pattern Anal. Mach. Intell. 1997. No. 19 (7). P. 711-720.
6. Li-Fang Zhou, Yue-Wei Du, Wei-Sheng Li, Jian-Xun Mi, Xiao Luan. Pose-robust face recognition with Huffman-LBP enhanced by Divide-and-Rule strategy // Pattern Recognition. 2018. Vol. 78. P. 43-55.
7. Taigman Y., Yang M., Ranzato M., Wolf L. Deepface: closing the gap to human-level performance in face verification // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014. P. 1701-1708.
8. Parkhi O.M., Vedaldi A., Zisserman A. Deep face recognition // Proceedings of the British Machine Vision Conference (BMVC). 2015. Vol. 1. P. 41.1-41.12.
9. Witham Claire L. Automated face recognition of rhesus macaques // Journal of Neuroscience Methods. 2018. Vol. 300. P. 157-165.
10. Hansen Mark F., Melvyn Smith L., Lyndon N., Smith N., Salter Michael G., Baxter Emma M., Farish Marianne, Grieve Bruce Towards on-farm pig face recognition using convolutional neural networks. Computers in Industry. 2018. No. 98. P. 145-152.



11. Wang Haiyan, He Zhongshi, Huang Yongwen, Chen Dingding, Zhou Zexun. Bodhisattva head images modeling style recognition of Dazu Rock Carvings based on deep convolutional network // Journal of Cultural Heritage. 2017. Vol. 27. P. 60-71.
12. Viola P., Jones M.J. Rapid Object Detection using a Boosted Cascade of Simple Features // Proceedings IEEE Conf. on Computer Vision and Pattern Recognition. 2001. Vol. 1. P. 511-518.
13. Nguyen, T., Hefenbrock, D., Oberg, J., Kastner, R., Baden, S. A software-based dynamic-warp scheduling approach for load-balancing the viola-jones face detection algorithm on gpus // J. Parallel Distrib. Comput. 2013. No. 73(5). P. 677–685.
14. Labeled Faces in the Wild Home database. URL: <http://www.cs.umass.edu/lfw/>. (дата обращения 09.06.2018).
15. Aberdeen dataset. URL: [http://pics.psych.stir.ac.uk/2D\\_face\\_sets.htm](http://pics.psych.stir.ac.uk/2D_face_sets.htm). (дата обращения 09.06.2018).
16. Ojala, T., Pietikainen, M., Maenpaa, M.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987 (2002).
17. Bledsoe W.W. Some Results on Multicategory Pattern Recognition // J. ACM. 1966. No.13. P. 304-316.
18. Turk M., Pentland A. . Eigenfaces for recognition // J. Cognit. Neurosci. 1991. No. 3 (1). P. 71-86.
19. Belhumeur P.N., Hespanha J.P., Kriegman D.J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection // IEEE Trans. Pattern Anal. Mach. Intell. 1997. No. 19 (7). P. 711-720. 4. Li-Fang Zhou, Yue-Wei Du, Wei-Sheng Li, Jian-Xun Mi, Xiao Luan. Pose-robust face recognition with Huffman-LBP enhanced by Divide-and-Rule strategy // Pattern Recognition. 2018. Vol. 78. P. 43-55.
20. Taigman Y., Yang M., Ranzato M., Wolf L. Deepface: closing the gap to human-level performance in face verification // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014. P. 1701-1708.
21. Parkhi O.M., Vedaldi A., Zisserman A. Deep face recognition // Proceedings of the British Machine Vision Conference (BMVC). 2015. Vol. 1. P. 41.1-41.12. 7.

Witham Claire L. Automated face recognition of rhesus macaques // Journal of Neuroscience Methods. 2018. Vol. 300. P. 157-165.

22. Hansen Mark F., Melvyn Smith L., Lyndon N., Smith N., Salter Michael G., Baxter Emma M., Farish Marianne, Grieve Bruce Towards on-farm pig face recognition using convolutional neural networks. Computers in Industry. 2018. No. 98. P. 145-152.

23. Wang Haiyan, He Zhongshi, Huang Yongwen, Chen Dingding, Zhou Zexun. Bodhisattva head images modeling style recognition of Dazu Rock Carvings based on deep convolutional network // Journal of Cultural Heritage. 2017. Vol. 27. P. 60-71.

24. Viola P., Jones M.J. Rapid Object Detection using a Boosted Cascade of Simple Features // Proceedings IEEE Conf. on Computer Vision and Pattern Recognition. 2001. Vol. 1. P. 511-518.

25. Irgens P., Bader C., Lé T., Saxena D., Ababei C. An efficient and cost effective FPGA based implementation of the Viola-Jones face detection algorithm // Hardware X. 2017. No.1. P. 68– 75.

26. Nguyen, T., Hefenbrock, D., Oberg, J., Kastner, R., Baden, S. A software-based dynamic-warp scheduling approach for load-balancing the viola-jones face detection algorithm on gpus // J. Parallel Distrib. Comput. 2013. No. 73(5). P. 677–685.

13. Labeled Faces in the Wild Home database. URL: <http://vis-www.cs.umass.edu/lfw/>.

27. Aberdeen dataset. URL: [http://pics.psych.stir.ac.uk/2D\\_face\\_sets.htm](http://pics.psych.stir.ac.uk/2D_face_sets.htm).

28. Huang, C.-C. An Extended Set of Haar-like Features for Bird Detection Based on AdaBoost / C.-C. Huang, C.-Y. Tsai, H.-C. Yang // Signal Processing, Image Processing and Pattern Recognition SE – 17 / ed. Kim T. et al. — Springer Berlin Heidelberg, 2011. — Vol. 260. — P. 160–169.

29. OpenCV [Электронный ресурс]. — Режим доступа: <http://opencv.org/>

30. Шепелев, К.В. Применение методов распознавания образов при модернизации существующих систем видеонаблюдения в системах безопасности / К.В. Шепелев // Современный мир опыт, проблемы и перспективы развития. — Ставрополь: Центр научного знания “Логос,” 2016. — № 2. — С. 9–12.

31. Норушис, А. Построение логических (древообразных) классификаторов методами нисходящего поиска (обзор) / А. Норушис // Статистические проблемы управления. — Вильнюс, 1990. — Т. 93. — С. 131–158.
32. OpenCV: Cascade Classifier Training [Электронный ресурс]. — Режим доступа: [http://docs.opencv.org/2.4/doc/user\\_guide/ug\\_traincascade.html](http://docs.opencv.org/2.4/doc/user_guide/ug_traincascade.html)
33. *Parkhi O.M., Vedaldi A., Zisserman A.* Deep face recognition // Proceedings of the British Machine Vision Conference (BMVC). 2015. Vol. 1. P. 41.1-41.12.
34. *Zhi H., Liu S.* Face recognition based on genetic algorithm // Journal of Visual Communication and Image Representation. 2019. Vol. 58. P. 495-502.
35. *Khan S.A., Ishtiaq M., Nazir M., Shaheen M.* Face recognition under varying expressions and illumination using particle swarm optimization // Journal of Computational Science. 2018. Vol. 28. P. 94-100.
36. *Nikan S., Ahmadi M.* A modified technique for face recognition under degraded conditions // Journal of Visual Communication and Image Representation. 2018. Vol. 55. P. 742-755.
37. *Ding C., Tao D.* Pose-invariant face recognition with homography-based normalization // Pattern Recognition. 2017. Vol. 66. P. 144-152.
38. *Liang Y., Zhang Y., Zeng X.X.* Pose-invariant 3D face recognition using half face // Signal Processing: Image Communication. 2017. Vol. 57. P. 84-90.
39. Labeled Faces in the Wild Home набор данных. <http://visumass.edu/lfw/> (дата обращения 09.06.2019).
40. *Wu S., Wang D.* Effect of subject's age and gender on face recognition results // Journal of Visual Communication and Image Representation. 2019. Vol. 60. P. 116-122.
41. *Muikudi P.B.L., Hills P.J.* The combined influence of the own-age, -gender, and -ethnicity bases on face recognition // Acta Psychologica. 2019. Vol. 194. P. 1-6.
42. *Segal S.C., Reyes B.N., Gobin K.C., Moulson M.C.* Children's recognition of emotion expressed by own-race versus other-race faces // Journal of Experimental Child Psychology. 2019. Vol. 182. P. 102-113.

43. *Пятаева А.В., Верхотурова М.В.* Детектирование лиц с использованием метода Виолы-Джонса // Материалы V Международной научной конференции «Региональные проблемы дистанционного зондирования Земли». Красноярск: ИКИТ СФУ, 2018. С. 188-191.

## СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ

*Индексируемые в базе Scopus:*

1. Video based face recognition method. //«CEUR Workshop Proceedings» (*в печати*). // Pyataeva A.V., Verkhoturova M.V.

*Российский индекс научного цитирования:*

2. Детектирование лиц с использованием метода Виолы – Джонса. «Материалы V Международной научной конференции «Региональные проблемы дистанционного зондирования земли» Красноярск, 11–14 сентября 2018 г.». 188-191 с. //Верхотурова М.В., Пятаева А.В.

3. Метод распознавания лиц по визуальным данным. «Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов (SDM-2019)» (*в печати*). // Пятаева А.В., Верхотурова М.В.

*Свидетельство о государственной регистрации программ для ЭВМ:*

4. Программа обнаружения лиц на базе метода Виолы – Джонса. (Дата государственной регистрации в Реестре программ для ЭВМ 18 марта 2019г.) // Верхотурова М.В., Пятаева А.В.

# ПРИЛОЖЕНИЕ А

## Плакаты презентации

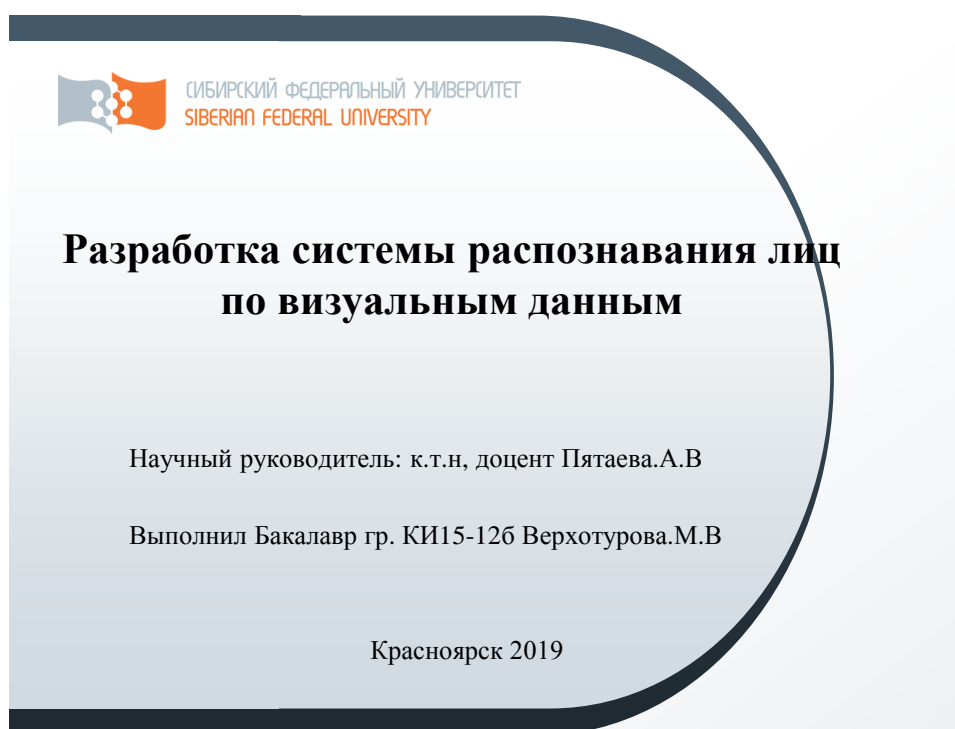


Рисунок А.1 – Слайд презентации № 1

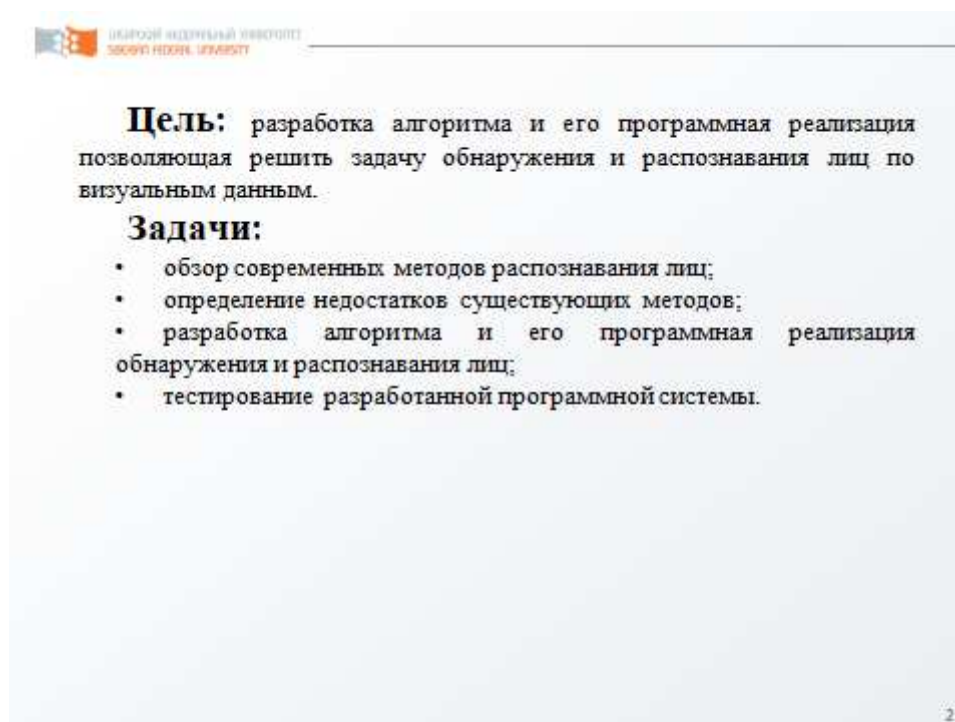


Рисунок А.2 – Слайд презентации № 2

## Подходы к обнаружению и распознаванию лиц

- эмпирический метод;
- методы характерных инвариантных признаков;
- активные модели внешнего вида;
- метод главных компонент;
- метод гибкого сравнения на графах;
- локальные бинарные шаблоны (ЛБШ);
- метод Виолы – Джонса.

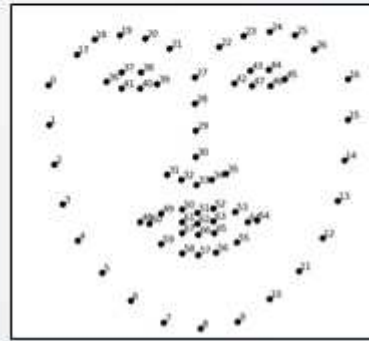
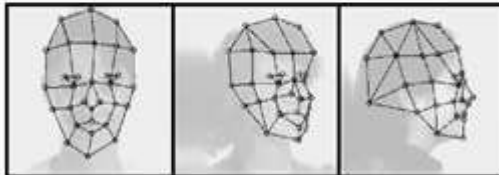


Рисунок А.3 – Слайд презентации № 3

## Задача распознавания лиц

Этап 1. Обнаружения лиц.



Этап 2. Распознавание лиц.



Рисунок А.4 – Слайд презентации № 4

## Этап 1. Метод Виолы-Джонса

Получение и преобразование видеопоследовательности в серию изображений.

Преобразование цветного изображения в полутоновое, для вычисления признаков Хаара.

Серия изображений взяты из набора данных Aberdeen.



Рисунок А.5 – Слайд презентации № 5

## Этап 1. Интегральное представление изображения

Элементы интегрального представления изображения рассчитываются по следующей формуле:

$$L(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j), \quad (1)$$

где  $I(i, j)$  – яркость текущего пиксела исходного изображения,  $(i, j)$  – координаты текущего пиксела.

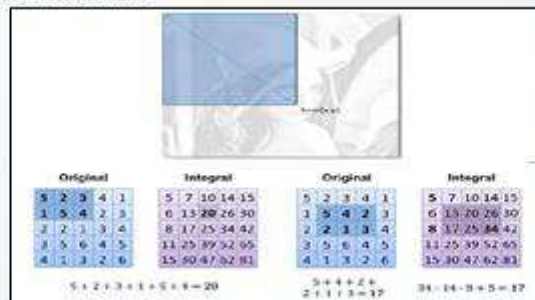


Рисунок А.6 – Слайд презентации № 6



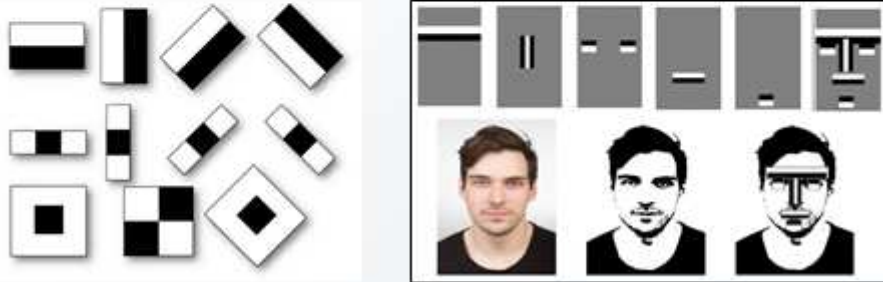
## Этап 1. Вычисление признаков Хаара

Поиск лица путем выделения признаков Хаара методом вычисления характерных перепадов яркостей пикселов.

Значение признака Хаара рассчитывается по формуле:

$$F = X - Y \quad (2)$$

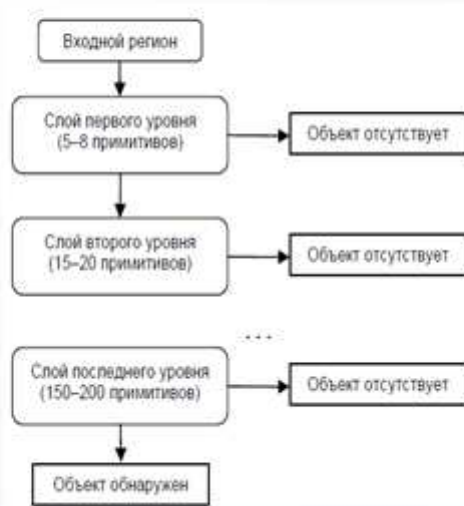
где  $X$  – сумма значений яркостей точек закрываемых светлой частью признака, а  $Y$  – сумма значений яркостей точек закрываемых темной частью признака.



7

Рисунок А.7 – Слайд презентации № 7

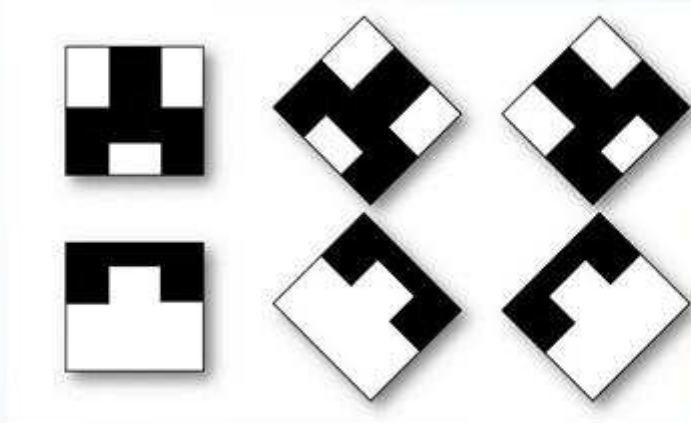
## Этап 1. Каскадная классификация



8

Рисунок А.8 – Слайд презентации № 8

## Этап 1. Дополнительные примитивы Хаара



9

Рисунок А.9 – Слайд презентации № 9

## Этап 2. Локальные бинарные шаблоны

При расчете шаблона изображение разбивается на непересекающиеся блоки. Формально обобщенный оператор ЛБШ можно представить в виде:

$$LBP_R(P) = \sum_{n=0}^{P-1} S(I_n - I_c) \cdot 2^n \quad (3)$$

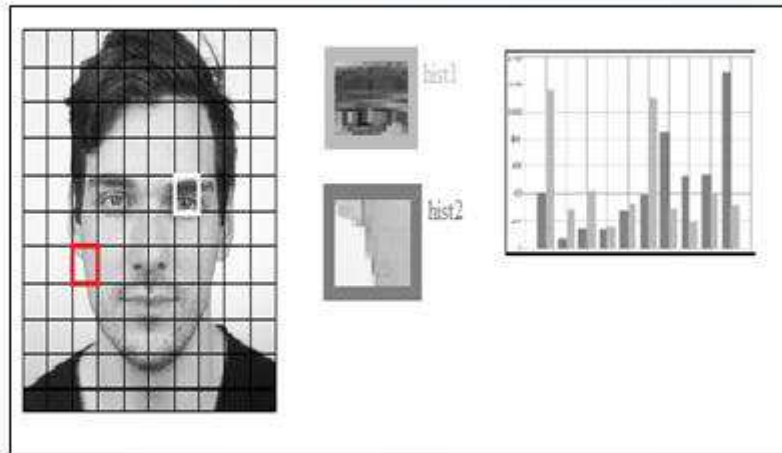
где  $P$  – количество пикселей в окрестности,  $R$  – радиус окрестности,  $s(x) = 1$ , если  $x \geq 0$ , и  $s(x) = 0$  в противном случае,  $I_n$  и  $I_c$  – значения яркости текущего и центрального пикселей, представляющие собой значение яркости  $Y$  из цветовой схемы YUV.



10

Рисунок А.10 – Слайд презентации № 10

## Этап 2. Построение гистограммы



11

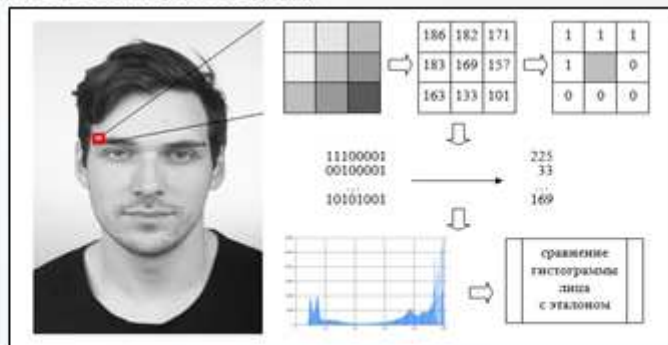
Рисунок А.11 – Слайд презентации № 11

## Этап 2. Различие гистограмм

В качестве меры различия гистограмм, вычисляется евклидово расстояние на основе данной формулы:

$$D = \sqrt{\sum_{i=1}^n (hist1_i - hist2_i)^2}, \quad (4)$$

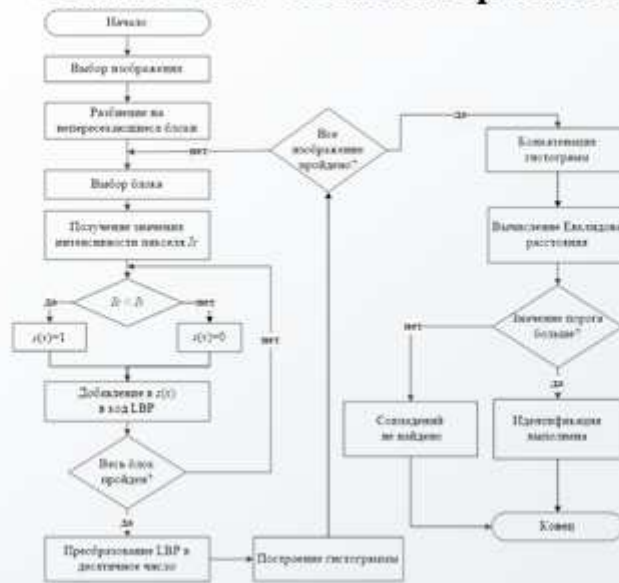
где,  $hist1_i$  – столбец с номером  $i$  гистограммы исследуемого изображения лица, где  $hist2_i$  – столбец с номером  $i$  гистограммы изображения из базы эталонов лиц,  $n$  – количество столбцов гистограмм.



12

Рисунок А.12 – Слайд презентации № 12

## Этап 2. Блок – схема алгоритма ЛБШ



13

Рисунок А.13 – Слайд презентации № 13

## Экспериментальные исследования

- Aberdeen, взятых из хранилища данных Labeled Faces in the Wild Home.
- «YouTube Faces (YTF) data set», база данных видео.
- «McGillFaces» база данных видео.
- «NRC-ИТ Лицевая Видео База Данных».

14

Рисунок А.14 – Слайд презентации № 14

## Экспериментальные исследования

Для проведения эксперимента использовались следующие показатели: TD (%) – точность обнаружения, FRR (%) – ложноположительные срабатывания, FAR (%) – ложноотрицательные срабатывания, для расчета данных показателей использовались следующие формулы:

$$TR = \frac{N_t}{N} \quad (4)$$

$$FAR = \frac{N_f - N_t}{N} \quad (5)$$

$$FRR = 1 - TR, \quad (6)$$

где  $N_t$  – количество кадров с лицом,  $N_f$  – количество кадров с лицом которое распознано,  $N$  – количество кадров в видеопоследовательности.



15

Рисунок А.15 – Слайд презентации № 15

## Экспериментальные исследования

Видеоролик	Обнаружение лиц			Распознавание лиц		
	TD,%	FRR,%	FAR,%	TR,%	FRR,%	FAR,%
YouTubeFaces/PIE_S1_C1	100	0,00	0,00	99,5	0,50	0,44
YouTubeFaces/PIE_S1_C3	100	0,00	0,00	100	0,00	0,00
YouTubeFaces/PIE_S1_C4	100	0,00	0,00	99,1	0,01	0,66
YouTubeFaces/PIE_S2_M1	100	0,00	0,00	97,5	2,50	1,10
YouTubeFaces/PIE_S2_M3	100	0,00	0,00	96,2	4,00	3,80
YouTubeFaces/PIE_S2_M6	100	0,00	0,00	100	0,00	0,00
YouTubeFaces/PIE_S2_M13	93,9	6,10	5,80	87,9	12,1	11,7
YouTubeFaces/PIE_S2_M7	100	0,00	0,00	100	0,00	0,00
YouTubeFaces/PIE_S2_D3	100	0,00	0,00	100	0,00	0,00
YouTubeFaces/PIE_S2_D5	95,3	6,66	4,70	88,9	11,1	11,1
YouTubeFaces/PIE_S2_D6	100	0,00	0,00	98,2	1,80	1,78
YouTubeFaces/PIE_S2_D8	100	0,00	0,00	100	0,00	0,00
<b>Среднее значение</b>	<b>98,01%</b>	<b>1,06%</b>	<b>0,88%</b>	<b>94,3%</b>	<b>2,66%</b>	<b>2,55%</b>

16

Рисунок А.16 – Слайд презентации № 16

## Заключение

Решение задачи распознавания лиц на сегодняшний день актуально для реализации различного рода практических задач. В настоящей работе для выделения лица использован алгоритм Виолы-Джонса, для распознавания лица применены локальные бинарные шаблоны. Экспериментальные исследования, проведенные на разнородных видеоданных, подтверждают эффективность предложенных методов.

17

Рисунок А.17 – Слайд презентации № 17

## Список публикаций по теме работы

*Индексируемые в базе Scopus:*

1. Video based face recognition method. //«CEUR Workshop Proceedings» (в печати). //Pyataeva A.V., Verkhoturova M.V.

*Российский индекс научного цитирования:*

2. Детектирование лиц с использованием метода Виолы – Джонса. «Материалы V Международной научной конференции «Региональные проблемы дистанционного зондирования земли» Красноярск, 11–14 сентября 2018 г.». 188-191 с. //Верхотурова М.В., Пятаева А.В.
3. Метод распознавания лиц по визуальным данным. «Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов (SDM-2019)» (в печати). // Пятаева А.В., Верхотурова М.В.

*Свидетельство о государственной регистрации программ для ЭВМ:*

4. Программа обнаружения лиц на базе метода Виолы – Джонса. (Дата государственной регистрации в Реестре программ для ЭВМ 18 марта 2019г.) // Верхотурова М.В., Пятаева А.В.

18

Рисунок А.18 – Слайд презентации № 18

**Спасибо за внимание!**

Рисунок А.19 – Слайд презентации № 19

Министерство науки и высшего образования РФ  
Федеральное государственное автономное  
образовательное учреждение  
высшего образования  
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»  
Институт космических и информационных технологий  
Кафедра систем искусственного интеллекта

УТВЕРЖДАЮ

Заведующий кафедрой СИИ  
Г. М. Цибульский

\_\_\_\_\_

подпись

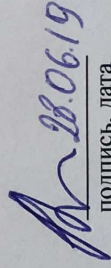
« \_\_\_\_\_ » \_\_\_\_\_ 2019 г.

## БАКАЛАВРСКАЯ РАБОТА

09.03.02 – «Информационные системы и технологии»

Разработка системы распознавания лиц по визуальным данным

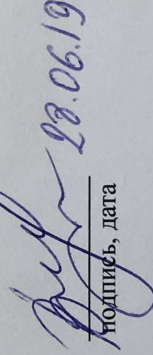
Руководитель



подпись, дата

к.т.н., доцент каф. СИИ А. В. Пятаева

Выпускник



подпись, дата

М.В.Верхотурова

Красноярск 2019