

Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Комсомольский-на-Амуре государственный университет»

На правах рукописи

Упская Ольга Константиновна

**Исследование и программная реализация  
алгоритмов распознавания лиц**

Направление подготовки  
02.04.03 «Математическое обеспечение  
и администрирование информационных систем»

**АВТОРЕФЕРАТ  
МАГИСТЕРСКОЙ ДИССЕРТАЦИИ**

2018

Работа выполнена в ФГБОУ ВО «Комсомольский-на-Амуре  
государственный университет»

Научный руководитель

кандидат физико-  
математических наук, доцент  
Лошманов Антон Юрьевич

Рецензент

кандидат технических наук,  
врио директора Института  
машиноведения и металлургии  
ДВО РАН  
Комаров Олег Николаевич

Защита состоится 22 июня 2018 года в 13 часов 00 мин на заседании государственной экзаменационной комиссии по направлению подготовки 02.04.03 «Математическое обеспечение и администрирование информационных систем» в Комсомольском-на-Амуре государственном университете по адресу: 681000, г. Комсомольск-на-Амуре, пр. Ленина, 27, ауд. 321/3.

Автореферат разослан 15 июня 2018 г.

Секретарь ГЭК

А.А. Сиротин

## ОБЩАЯ ХАРАКТЕРИСТИКА ДИССЕРТАЦИОННОЙ РАБОТЫ

*Актуальность темы.* Проблема формализации и автоматизации процесса распознавания человеческих лиц была затронута еще на самых ранних стадиях развития систем распознавания образов и остается актуальной до сих пор. Однако в течение последних десяти лет количество научных исследований и публикаций по данной тематике выросло в несколько раз, что свидетельствует о возрастании актуальности этой проблемы. Это объясняется тем, что современная вычислительная техника открывает новые возможности ее применения в самых различных сферах, и соответственно появилось множество прикладных задач, требующих своего скорейшего разрешения.

Одним из практических применений теории распознавания образов является распознавание лиц, в задачу которого входит автоматическая локализация лица на изображении и идентификация человека по лицу. Интерес к процедурам, лежащим в основе процесса локализации и распознавания лиц, довольно значителен в связи с разнообразием их практического применения в таких областях, как охранные системы, верификация, криминалистическая экспертиза, телеконференции, компьютерные игры и т. д.

Серьезной проблемой, стоящей перед системами идентификации по изображению лица, является большая изменчивость визуальных образов, связанная с изменениями освещенности, окраски, масштабов, ракурсов наблюдения, а также головные уборы и очки, которые носит человек.

В современном мире все большую популярность набирает использование квадрокоптеров в различных сферах деятельности. Одной из задач является управление квадрокоптером при помощи движений головы, для этого в первую очередь необходима реализация алгоритма распознавания лиц для идентификации человека. В данном случае критериями при выборе алгоритмов являются быстрое распознавание и устойчивость к изменениям.

*Целью магистерской диссертации* является программная реализация двух алгоритмов распознавания лиц и исследование (анализ) их эффективности и скорости работы.

Для достижения указанной цели поставлены следующие *задачи*:

- изучить предметную область;
- исследовать и сравнить существующие алгоритмы распознавания лиц;
- программно реализовать два алгоритма распознавания лиц;
- проанализировать эффективность и скорость работы данных алгоритмов.

*Объектом исследования* является процесс распознавания лиц.

*Предметом исследования* является метод Виолы-Джонса совместно с использованием локальных бинарных шаблонов и сверточная нейронная сеть для распознавания лиц.

Для решения поставленных задач использовались следующие *методы исследования*: теоретические (сравнение, анализ) и эмпирические (тестирование, изучение литературы и результатов деятельности).

*Научная новизна исследования* заключается в следующем:

- исследование применения совместного использования метода Виолы-Джонса для детектирования и локальных бинарных шаблонов для распознавания лиц;
- применение полученных результатов для управления квадрокоптером при помощи движений головы.

*Достоверность и обоснованность результатов исследования.*

Основные положения и выводы, полученные в диссертации, достаточно обоснованы и аргументированы. Сформулированная в диссертации научная задача, заключающаяся в анализе алгоритмов распознавания лиц и выявления алгоритма, подходящего для задачи управления квадрокоптером, была исследована и решена на основе совместного использования метода Виолы-Джонса и локальных бинарных шаблонов.

Достоверность основных выводов и результатов диссертации подтверждается:

- обоснованием выбора метода Виолы-Джонса и сверточной нейронной сети для распознавания лиц;
- использованием современного апробированного научно-методического аппарата для формализации и решения сформулированной в диссертации научной задачи;
- полнотой опубликования результатов исследования и их широкой апробацией.

*Практическая значимость.* Предложенные в диссертационной работе методы и алгоритмы предназначены для практического применения в системах видеонаблюдения и контроля доступа. На основе диссертационных исследований разработана библиотека программных модулей для создания систем локализации и распознавания изображений лиц при управлении квадрокоптером движениями головы.

*Апробация результатов.* Результаты работы докладывались на 48-ой научно-технической конференции студентов и аспирантов «Научно-техническое творчество аспирантов и студентов», Комсомольск-на-Амуре, апрель 2018 г.

*Публикации.* По результатам выполненных в диссертации исследований автором опубликовано 1 работа:

- сборник материалов 48-ой научно-технической конференции студентов и аспирантов «Научно-техническое творчество аспирантов и студентов» (Комсомольск-на-Амуре, апрель 2018 г.).

*Структура и объем.* Магистерская диссертация состоит из введения, трех глав, заключения, списка литературы и приложения. Объем работы – 61 страница, в том числе 29 рисунков, 1 таблица и 1 приложение.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Введение раскрывает актуальность темы, обоснование необходимости проведения работы, определяются цели и задачи исследования, объект, предмет, указываются научная новизна, практическая значимость, достоверность и обоснованность результатов исследования.

В первой главе дается понятие теории распознавания образов, ее основные задачи, рассматривается распознавание лиц, как одна из задач теории распознавания образов.

Теория распознавания образа — раздел информатики и смежных дисциплин, развивающий основы и методы классификации и идентификации предметов, явлений, процессов, сигналов, ситуаций и т. п. объектов, которые характеризуются конечным набором некоторых свойств и признаков.

Проблема распознавания образа приобрела выдающееся значение в условиях информационных перегрузок, когда человек не справляется с линейно-последовательным пониманием поступающих к нему сообщений, в результате чего его мозг переключается на режим одновременности восприятия и мышления, которому свойственно такое распознавание.

Распознавание лиц — практическое приложение теории распознавания образов, в задачу которого входит автоматическая локализация лица на фотографии и, в случае необходимости, идентификация персоны по лицу.

Во второй главе рассматриваются алгоритмы распознавания лиц, выявляются преимущества и недостатки каждого из алгоритмов, приводится их сравнительная характеристика.

*Метод главных компонент.* Входные векторы представляют собой отцентрированные и приведенные к единому масштабу изображения лиц. Собственные векторы, вычисленные для всего набора изображений лиц, называются собственными лицами. 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 55 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.

*Метод Виолы-Джонса совместно с использованием локальных бинарных шаблонов.*

Изображения используются в интегральном представлении, что позволяет вычислять быстро необходимые объекты. Интегральное представление – это матрица, равная размеру исходного изображения  $I$ . Каждый элемент матрицы можно рассчитать по формуле

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

где  $I(i, j)$  — яркость пиксела исходного изображения.

Каждый элемент матрицы  $L[x, y]$  представляет собой сумму пикселей в прямоугольнике от  $(0, 0)$  до  $(x, y)$ . Для вычисления суммы прямоугольной области в интегральном представлении изображения требуется всего 4 операции обращения к массиву и 3 арифметические операции, что позволяет быстро рассчитывать признаки Хаара для изображения.

Метод Виолы-Джонса использует признаки Хаара (рисунок 1), с помощью которых происходит обнаружения лица и его черт на изображении.

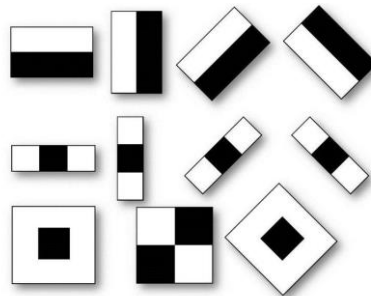


Рисунок 1 – Примитивы Хаара

Результаты детектирования лица очень быстры, но обучение классификаторов идет очень медленно, поэтому было решено объединить метод Вио-

лы-Джонса с локальными бинарными шаблонами (ЛБШ), которые быстро справляются с задачей распознавания лица.

Локальный бинарный шаблон (ЛБШ) представляет собой описание окрестности пикселя изображения в двоичном представлении. Исходное изображение разбивается на блоки 3x3 пикселя. Для вычисления ЛБШ в некоторой точке изображения используется восемь пикселей ее окрестности, а значение интенсивности центрального пикселя принимается в качестве порога. Пиксели, которые имеют значения больше, чем центральный пиксель (или равное ему), принимают значения 1, те, которые меньше центрального, принимают значения 0. Так мы получим восьмиразрядный бинарный код, который описывает окрестность этого пикселя, например, 11100001, как показано на рисунке 2.

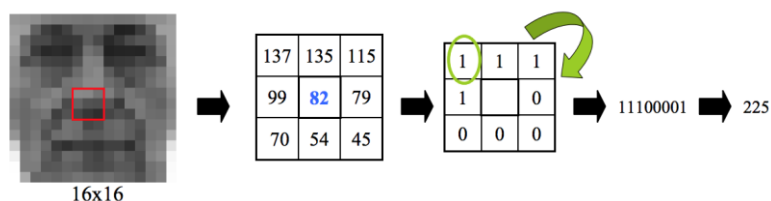


Рисунок 2 – Получение локального бинарного шаблона пикселя

ЛБШ вычисляется для каждого блока изображения, после этого вычисляются гистограммы каждого блока и конкатенируются в общую гистограмму особенностей изображения человеческого лица. Пример разбиения изображения на блоки и расчет гистограмм показан на рисунке 3.

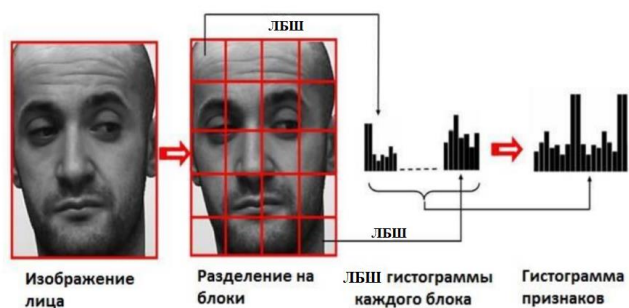


Рисунок 3 – Разбиение изображения и формирование гистограммы



Распознавание происходит за счет сравнения гистограмм лиц из тренировочной базы данных и гистограммы распознаваемого лица.

*Метод сравнения эластичных графов.* Лица представляются 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 подгонки к первому. В вершинах графа вычисляются значения признаков, чаще всего используют комплексные значения фильтров Габора или их упорядоченных наборов – Габоровских вейвлет, которые вычисляются в некоторой локальной области вершины графа локально путем свертки значений яркости пикселей с фильтрами Габора. Ребра графа взвешиваются расстояниями между смежными вершинами. Различие 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. Результат распознавания системы – эталон с наилучшим значением ценовой функции деформации.

*Сверточная нейронная сеть.* Отличительными особенностями СНС являются локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов), общие веса (обеспечивают детектирование некоторых черт в любом месте изображения) и иерархическая организация с пространственными сэмпингом (spatial subsampling). Благодаря этим нововведениям СНС обеспечивает частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и прочим искажениям.

На рисунке 4 представлена архитектура сверточной нейронной сети, которая состоит из входа, сверточных и подвыборочных слоев (которые чередуются между собой), слоя обычных нейронов и выхода.

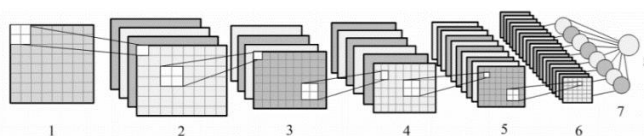


Рисунок 4 – Архитектура сверточной нейронной сети

Сравнительная характеристика вышеописанных алгоритмов представлена в таблице 1.

Таблица 1 – Сравнительная характеристика алгоритмов

Свойство \ Алгоритм	Метод главных компонент	Метод Виолы-Джонса совместно с использованием ЛБШ	Метод сравнения эластичных графов	Свёрточная нейронная сеть
Быстрое распознавание	+	+	–	+
Устойчивость к изменениям освещения	–	+	+	+
Устойчивость к изменениям эмоционального выражения лица	–	+	+	+
Устойчивость к поворотам лица	–	+	+	+
Высокая эффективность	–	+	+	+

Третья глава посвящена программной реализации метода Виолы-Джонса совместно с использованием локальных бинарных шаблонов и сверточной нейронной сети.

После сравнительного анализа алгоритмов, приведенных в главе 2, были выбраны два алгоритма: метод Виолы-Джонса совместно с локальными бинарными шаблонами и сверточная нейронная сеть. Критериями при выборе алгоритмов были быстрое распознавание и устойчивость к изменениям.

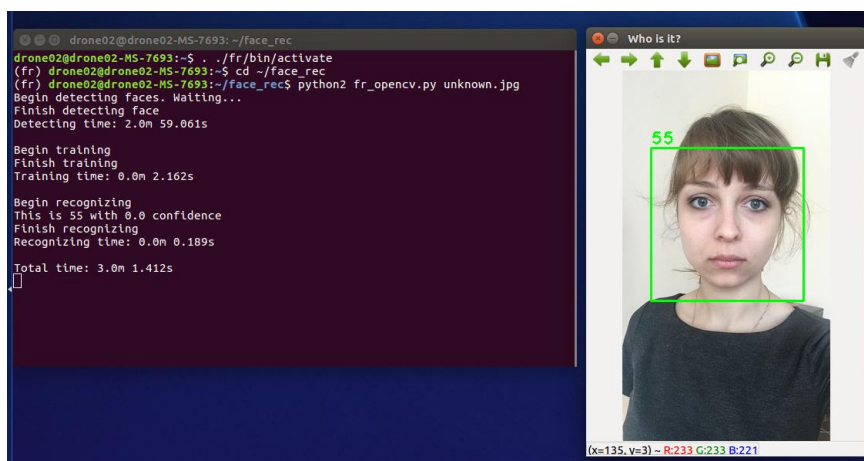


Рисунок 5 – Распознавание лица на изображении (метод Виолы-Джонса совместно с использованием локальных бинарных шаблонов)

Результат распознавания выполнен верно: это человек под номер 55 из базы изображений «train\_data». Общее время работы: 3 минуты 1.4 секунды. Параметр confidence равен 0, т.е. уверенность алгоритма в том, что это именно этот человек, равна 100 %.

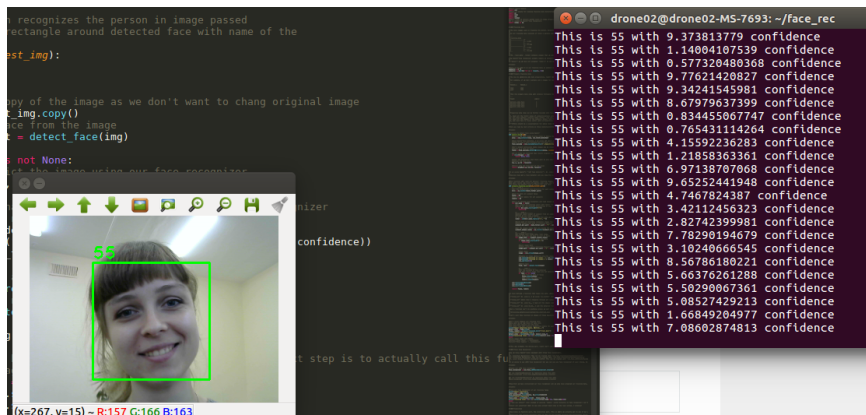


Рисунок 6 – Распознавание лица на видеопотоке (метод Виолы-Джонса совместно с использованием локальных бинарных шаблонов)

Результат распознавания выполнен верно: это человек под номер 55 из базы изображений «train\_data». Общее время работы: 3 минуты 1.4 секунды. Параметр confidence варьируется от 0.57 до 9.65, т.е. уверенность алгоритма в том, что это именно этот человек, варьируется от 99.43 до 90.35 % в зависимости от положения лица в камере.

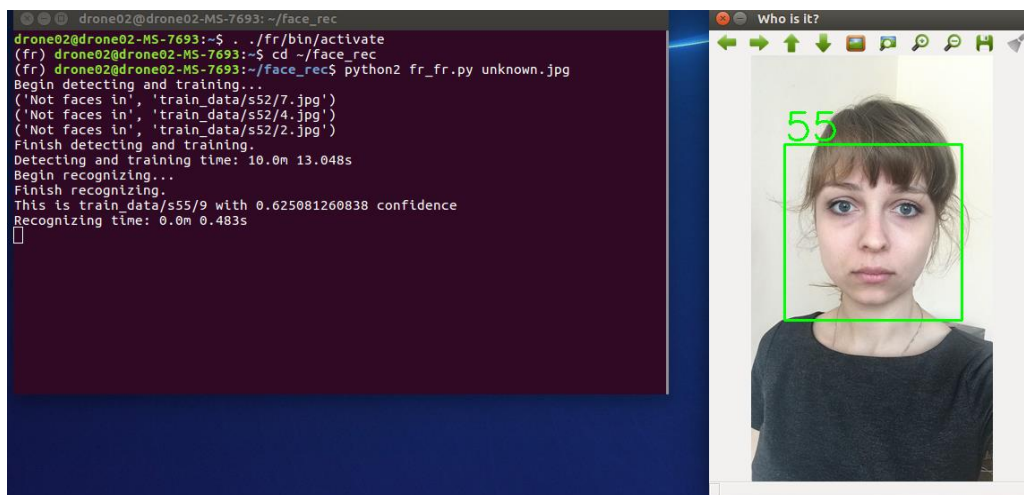


Рисунок 7 – Распознавание лица на изображении (сверточная нейронная сеть)

Результат распознавания выполнен верно: это человек под номер 55 из базы изображений «train\_data». Общее время работы: 10 минут 13 секунд.

Параметр confidence равен 0,625 т.е. уверенность алгоритма в том, что это именно этот человек, равна 62.5 %. В ходе обучения и обнаружения лиц на трех изображениях лица не обнаружены.

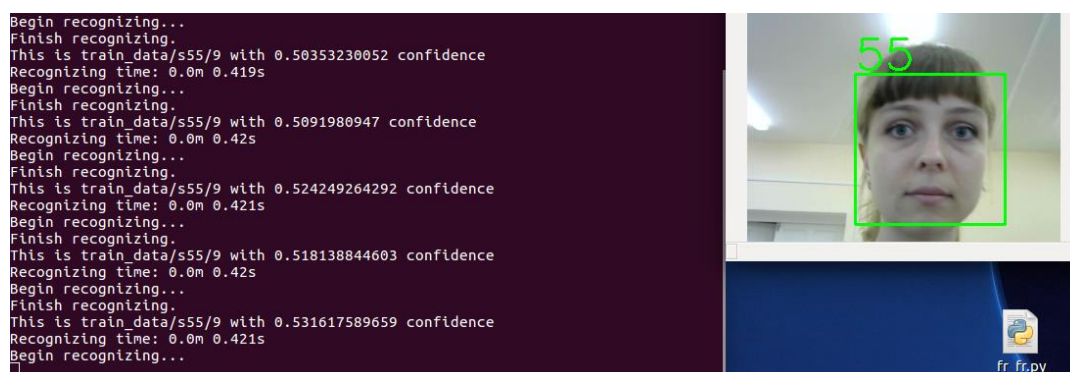


Рисунок 8 – Распознавание лица на видеопотоке (сверточная нейронная сеть)

Результат распознавания выполнен верно: это человек под номер 55 из базы изображений «train\_data». Общее время работы: 9 минут 46 секунд. Параметр confidence варьируется от 0.5 до 0.53, т.е. уверенность алгоритма в том, что это именно этот человек, варьируется от 50 до 53 % в зависимости от положения лица в камере.

В заключении подводятся итоги исследования, формируются окончательные выводы по рассматриваемой теме.

## СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ИССЛЕДОВАНИЯ

1 Упская, О. К. Исследование алгоритмов распознавания лиц / О.К. Упская, А. Ю. Лошманов // Научно-техническое творчество аспирантов и студентов: материалы всероссийской научно-технической конференции студентов и аспирантов. Комсомольск-на-Амуре, 09-20 апреля 2018г.: в 2 ч. – 2018. – Ч.2. – С. 245-247.