

ЯРОСЛАВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
им. П.Г. ДЕМИДОВА

На правах рукописи



Шмаглит Лев Александрович

**Разработка и анализ алгоритмов распознавания лиц
на телевизионных изображениях для биометрической
идентификации**

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата технических наук
по специальности 05.12.04 «Радиотехника, в том числе системы и
устройства телевидения»

Научный руководитель - д.т.н., доцент Приоров Андрей Леонидович

Ярославль – 2014

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	4
1. АЛГОРИТМИЧЕСКИЕ ОСНОВЫ РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ТЕЛЕВИЗИОННЫХ ИЗОБРАЖЕНИЯХ.....	13
1.1. Вводные замечания.....	13
1.2. Алгоритмы детектирования лиц на телевизионных изображениях.....	20
1.2.1. Алгоритм на базе адаптивного бустинга.....	21
1.2.2. Алгоритм на базе разреженной сети просеивающих элементов.....	26
1.2.3. Алгоритм на базе метода опорных векторов.....	31
1.3. Построение бинарных классификаторов для распознавания пола людей по детектированной области лица.....	34
1.3.1. Алгоритм на базе линейного дискриминантного анализа.....	34
1.3.2. Применение метода опорных векторов в задаче гендерной классификации.....	36
1.3.3. Алгоритм на базе адаптивных признаков и метода опорных векторов.....	37
1.4. Методика и результаты сравнительного анализа существующих алгоритмов распознавания лиц.....	40
1.4.1. Тестирование алгоритмов детектирования лиц на телевизионных изображениях.....	40
1.4.2. Тестирование алгоритмов распознавания пола людей по детектированной области лица.....	42
1.5. Краткие выводы.....	46
2. ДЕТЕКТИРОВАНИЕ ЛИЦ НА ТЕЛЕВИЗИОННЫХ ИЗОБРАЖЕНИЯХ ПРИ ИЗМЕНЕНИИ УСЛОВИЙ ОСВЕЩЕННОСТИ.....	48
2.1. Вводные замечания.....	48
2.2. Модели изменения условий освещенности.....	49
2.2.1. Модель изменения яркости.....	51
2.2.2. Модель изменения контраста.....	52
2.2.3. Модель неравномерной освещенности.....	54
2.3. Исследование работы существующих алгоритмов детектирования лиц при изменении условий освещенности.....	55
2.4. Модификация этапа предобработки изображений при детектировании лиц.....	59
2.4.1. Эквализация гистограммы изображения.....	60
2.4.2. Локальная предобработка.....	61
2.5. Результаты тестирования алгоритмов детектирования лиц с предложенными модификациями.....	64
2.6. Краткие выводы.....	74

3. РАСПОЗНАВАНИЕ ПОЛА ЛЮДЕЙ ПО ИЗОБРАЖЕНИЮ ЛИЦА ПРИ ИЗМЕНЕНИИ УСЛОВИЙ ОСВЕЩЕННОСТИ	76
3.1. Вводные замечания.....	76
3.2. Исследование работы алгоритма на базе метода опорных векторов при изменении условий освещенности	78
3.3. Предлагаемые модификации базового алгоритма на основе метода опорных векторов	82
3.3.1. Эквализация гистограммы изображения	83
3.3.2. Алгоритм локального выравнивания яркости и контраста	84
3.3.3. Алгоритм локальной предобработки с сегментацией	87
3.4. Результаты тестирования предложенных модификаций.....	90
3.5. Краткие выводы	96
4. МОДИФИКАЦИЯ ПРИЗНАКОВОГО ПРОСТРАНСТВА АЛГОРИТМА РАСПОЗНАВАНИЯ ПОЛА НА БАЗЕ МЕТОДА ОПОРНЫХ ВЕКТОРОВ ..	97
4.1. Вводные замечания.....	97
4.2. Процедура вычисления локальных бинарных шаблонов.....	98
4.3. Результаты тестирования предложенных модификаций признакового пространства.....	100
4.4. Краткие выводы	104
ЗАКЛЮЧЕНИЕ	105
СПИСОК ЛИТЕРАТУРЫ.....	108
ПРИЛОЖЕНИЕ 1. АКТЫ ВНЕДРЕНИЯ РЕЗУЛЬТАТОВ РАБОТЫ	119

ВВЕДЕНИЕ

Актуальность темы. Задача распознавания объектов [1-5] на телевизионных изображениях заключается в определении на основе многочисленных характеристик некоторого объекта одной или нескольких наиболее существенных его характеристик, недоступных для непосредственного определения, в частности, его принадлежности к определенному классу объектов. Для решения этой задачи необходимо применение комплекса алгоритмов, направленных на получение, обработку и анализ видеоданных [6-10]. При этом конечной целью анализа является выработка решений, позволяющих получить новую информацию, обнаружить интересующие объекты, распознать их свойства, отследить и предсказать их дальнейшее поведение [11, 12]. Информация такого рода широко востребована и используется в различных областях, начиная от контроля качества за производственными процессами на предприятии и заканчивая созданием искусственного интеллекта и интерфейсов взаимодействия человек-компьютер [13-16].

В настоящее время быстро развиваются технологии получения и передачи телевизионных изображений, увеличивается емкость передаваемой информации. В связи с этим растет потребность в интеллектуальных системах анализа видеоданных. Так миллиарды камер охранного телевидения – теперь неотъемлемая часть современного цивилизованного мира [17, 18]. Изображения сменяют друг друга на мониторах, многочисленные операторы обеспечивают безопасность в аэропортах, на вокзалах и в других общественных местах. Видеонаблюдение сегодня считается едва ли не самым главным техническим инструментом обеспечения безопасности. При этом для того чтобы контролировать тысячи видеоканалов в режиме реального времени необходимы эффективные алгоритмы, способные без участия человека автоматически оценивать внешнюю среду и выполнять в ней те или иные действия.

Среди всех возможных анализируемых объектов наибольший интерес представляют человеческие лица. Задача детектирования и классификации человеческих лиц [19-25] находит применение в различных сферах человеческой деятельности, в первую очередь, в системах безопасности. Сферами применения распознавания лиц являются, например, системы охранного телевидения, сравнение фотографий на паспортах или водительских удостоверениях, контроль доступа к безопасным компьютерным сетям и оборудованию в учреждениях, наблюдение за аэропортами и вокзалами для предотвращения террористических актов и др.

В данной работе решаются две подзадачи распознавания лиц – детектирование лиц на телевизионных изображениях и распознавание пола людей по детектированной области лица. Алгоритм автоматического детектирования лиц по произвольному изображению на входе определяет, имеются ли на этом изображении лица, и если да, то указывает, где находится каждое лицо и каков его размер. Алгоритм распознавания пола разделяет все анализируемые фрагменты на два класса – «мужчины» и «женщины». Информация, полученная в результате подобного анализа, широко востребована в системах биометрической идентификации, решающих задачу идентификации личности человека по изображению его лица. Так, успешное распознавание пола (при одинаковом количестве мужчин и женщин в базе) вдвое сокращает область поиска по базе лиц при идентификации личности. Другой сферой применения рассматриваемых алгоритмов являются системы автоматизированных отношений с клиентами, решающие задачу анализа состава аудитории с целью индивидуального подхода к предоставлению рекламы и услуг клиентам различной гендерной принадлежности.

Следует отметить, что задачи детектирования лиц и распознавания пола являются составной частью системы интеллектуального анализа видеоданных и решаются совместно с рядом других задач, таких как сопровождение объектов, идентификация личности, анализ и накопление

данных. При этом алгоритмы детектирования и трекинга лиц на изображениях [26-33] являются источником входных данных для алгоритмов распознавания. Современные алгоритмы позволяют детектировать лицо человека на реальных видеоизображениях со сложным фоном с вероятностью, превышающей 90%. Среди них самым известным является алгоритм на основе бустинга [34], предложенный Р. Viola и М. Jones в 2001 г.

Рассмотрим задачу распознавания пола человека по изображению лица. С точки зрения принятия решения (разбиения всех анализируемых фрагментов на два класса: «мужчины» и «женщины») – это классическая задача бинарной классификации, имеющая равный вес ошибок первого и второго рода. Такая симметричность обусловлена приблизительно равным количеством лиц мужчин и женщин в видеопотоке. Исходя из этого предположения, при обучении классификатора формируются равные по количеству фрагментов обучающие выборки для каждого класса. Основной проблемой при построении классификатора с целью автоматического распознавания пола человека является нахождение признаков, по которым можно различить мужчин и женщин, лица которых обладают существенным сходством. Другим важным вопросом является устойчивость алгоритма к изменениям условий освещенности в реальных условиях эксплуатации, а также к наличию искажений, связанных с помехами в радиотехнических устройствах и сжатием двумерного сигнала.

Эффективное решение обозначенных проблем требует применения специализированных алгоритмов цифровой обработки изображений, а также современных методов машинного обучения. Среди них можно выделить нелинейные алгоритмы фильтрации и восстановления изображений, вейвлет-обработку, системы, построенные на нечеткой логике, генетических алгоритмах, нейронных сетях.

В область разработки алгоритмов цифровой обработки телевизионных изображений внесли вклад как отечественные ученые – Ю.Б. Зубарев [35], М.И. Кривошеев, В.П. Дворкович [36-40], А.В. Дворкович, М.К. Чобану [41],

А.С. Крылов, Ю.С. Бехтин, Ю.С. Радченко, А.Л. Приоров, так и зарубежные – R. Gonzalez [42-44], R. Woods [45], A. Bovik, Y. Neuvo, J. Astola, K. Egiazaryan, M. Nikolova.

Первые работы по классификации динамических объектов были опубликованы еще в конце 1980-х гг. Следует отметить особый вклад советских и российских ученых, работающих в данной области, таких как Ю.И. Журавлев [46-49], Ю.В. Гуляев, В.А. Сойфер [50], А.И. Галушкин, В.Н. Вапник [51], Н.Г. Загоруйко [52, 53], А.А. Потапов [54].

На сегодняшний день достигнуты значительные успехи при построении систем детектирования и распознавания объектов на телевизионных изображениях, однако, существуют нерешенные проблемы при классификации динамических образов, возникающие из-за сложности и многообразия поведения объектов реального мира [55]. Существует отдельные задачи (видеонаблюдение в закрытых помещениях, в местах большого скопления людей, управление движением мобильных робототехнических комплексов, наблюдение за движением транспортных средств и т.д.), в которых устойчивость к наличию искажений и изменению условий освещенности, а также способность работы в реальном времени приобретают решающее значение. Поэтому на современном этапе развития науки и техники разработка и анализ алгоритмов распознавания лиц на телевизионных изображениях представляют собой актуальную задачу.

Основной целью работы является улучшение характеристик алгоритмов детектирования и распознавания лиц людей на телевизионных изображениях для систем биометрической идентификации.

Для достижения указанной цели в диссертационной работе решаются следующие **задачи**:

- усовершенствование работы алгоритмов детектирования лиц на телевизионных изображениях при изменении условий освещенности;
- сравнительный анализ работы алгоритмов гендерной классификации с использованием единой базы обучающих и тестовых изображений;

- исследование устойчивости базового классификатора на основе метода опорных векторов к изменению условий освещенности при распознавании пола человека по изображению лица;
- модификация классификатора на базе метода опорных векторов путем применения существующих и разработки новых алгоритмов предварительной обработки изображений;
- усовершенствование работы классификатора на базе метода опорных векторов за счет применения локальных бинарных шаблонов и комбинированного пространства признаков.

Объектом исследования являются алгоритмы распознавания объектов, применяемые в радиотехнических системах обработки и анализа изображений и современных системах телевидения.

Предметом исследования являются модификация и разработка алгоритмов с целью повышения эффективности функционирования видеоинформационных систем биометрической идентификации.

Методы исследования. При решении поставленных задач использовались современные методы цифровой обработки изображений, машинного обучения, распознавания образов, математического анализа, теории вероятностей и математической статистики. Для практической реализации алгоритмов применялись современные численные методы и методы объектно-ориентированного программирования на языке C#.

Научная новизна

В рамках данной работы получены следующие новые научные результаты:

1. Исследовано влияние условий освещенности на работу алгоритмов детектирования лиц.
2. Разработаны алгоритмы локальной предобработки изображений и локальной предобработки с предварительной сегментацией по цвету кожи.

3. Разработаны три модификации гендерного классификатора на базе метода опорных векторов с применением существующих и предлагаемых алгоритмов предобработки изображений.
4. Разработана модификация гендерного классификатора с применением локальных бинарных шаблонов и комбинированного пространства признаков.

Практическая значимость

1. Проведен анализ работы трех алгоритмов детектирования лиц на телевизионных изображениях, предложены алгоритмы предобработки, позволяющие повысить устойчивость тестируемых алгоритмов к изменению условий освещенности.
2. Проведен анализ работы алгоритма гендерной классификации на базе метода опорных векторов при изменении условий освещенности, предложены модификации, позволяющие повысить средний уровень верного распознавания на 10-15% в зависимости от рассматриваемых условий освещенности.
3. Разработана модификация алгоритма гендерной классификации на базе комбинированного пространства признаков и метода опорных векторов, уровень верного распознавания которой составляет порядка 94%, что на 3,3% превышает показатели лучшего из известных до этого алгоритмов.

Разработанные алгоритмы требуют для их практической реализации относительно небольших вычислительных ресурсов, что позволяет использовать их для обработки телевизионных изображений в системах реального времени.

Результаты работы внедрены в соответствующие разработки ООО «АйДата», г. Ярославль; ООО «ДиММ», г. Ярославль. Отдельные результаты диссертационной работы внедрены в учебный процесс ЯрГУ им. П.Г. Демидова в рамках дисциплин «Цифровая обработка изображений» и «Компьютерное зрение», а также в научно-исследовательские работы при

выполнении исследований в рамках гранта РФФИ №10-08-01186. Все результаты внедрения подтверждены соответствующими актами. Получено свидетельство на программный продукт, зарегистрированное в Реестре программ для ЭВМ [114].

Достоверность полученных научных результатов обусловлена применением адекватного математического аппарата, подтверждается их согласованностью с результатами проведенного компьютерного моделирования и сопоставлением полученных результатов с научными данными, известными из российской и зарубежной литературы.

Апробация работы. Результаты работы докладывались и обсуждались на следующих научно-технических конференциях и семинарах.

- 11, 14-16-я международные конференции «Цифровая обработка сигналов и ее применение», Москва, 2009, 2012-2013.
- Двадцатая международная конференция по компьютерной графике и зрению «ГрафиКон'2010». Санкт-Петербург, 2010.
- 67-я научная сессия, посвященная Дню радио, Москва, 2012.
- 18-я международная научно-техническая конференция «Радиоэлектроника, электротехника и энергетика», Москва, 2012.
- 10-я всероссийская научная конференция «Нейрокомпьютеры и их применение», Москва, 2012.
- 14-я всероссийская научно-техническая конференция «Нейроинформатика». Москва, 2012.
- World Congress on Engineering and Computer Science, Berkeley, USA, 2012.
- International MultiConference of Engineers and Computer Scientists (imaging engineering section), Hong Kong, 2013.
- 14th Conference of Open Innovation Association FRUCT, Helsinki, Finland, 2013.
- 23rd International Conference on Computer Graphics and Vision, GraphiCon'2013, Vladivostok, 2013.

Публикации. По теме диссертации опубликовано 17 научных работ, из них 3 статьи в журналах, рекомендованных ВАК («Проектирование и технологии электронных средств», «Optical Memory and Neural Networks», «Электромагнитные волны и электронные системы»), 1 статья в журнале «IAENG International Journal of Computer Science» и 13 докладов на научных конференциях.

Структура и объем работы. Диссертация состоит из введения, четырех разделов, заключения, списка использованных источников, содержащего 115 наименований. Она изложена на 119 страницах машинописного текста, содержит 56 рисунков и 18 таблиц.

Основные научные положения и результаты, выносимые на защиту

1. Алгоритм локальной предобработки изображений и алгоритм локальной предобработки с предварительной сегментацией по цвету кожи.
2. Модификация алгоритма распознавания пола человека по изображению лица на базе предложенных алгоритмов предобработки.
3. Модификация алгоритма распознавания пола на основе метода опорных векторов путем расширения пространства признаков классификатора.

Благодарности. Автор выражает слова глубокой благодарности своему научному руководителю доценту Приорову Андрею Леонидовичу за помощь на всех этапах выполнения данной работы. Отдельная благодарность преподавателям ЯрГУ Л.Н. Казакову, А.Н. Кренину, В.А. Тимофееву, К.С. Артемову, Т.К. Артемовой, А.Н. Тараканову, работы которых оказали значительное влияние на формирование взглядов автора в данном научном направлении. Эти взгляды формировались также в совместной работе с коллегами по лаборатории «Цифровые цепи и сигналы» кафедры динамики электронных систем ЯрГУ, среди которых особенно хочется отметить Ю. Лукашевича, А. Ганина, М. Голубева, И. Апалькова, Д. Куйкина,

Д. Зараменского, В. Волохова, Е. Аминову, Д. Завьялову, Д. Тарасову, Т. Тихонову.

Особая благодарность – профессору Юрию Александровичу Брюханову и доценту Хрящеву Владимиру Вячеславовичу за постоянную поддержку во время обучения в аспирантуре и подготовки данной диссертации.

1. АЛГОРИТМИЧЕСКИЕ ОСНОВЫ РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ТЕЛЕВИЗИОННЫХ ИЗОБРАЖЕНИЯХ

1.1. Вводные замечания

Визуальный анализ динамических сцен, особенно с участием людей, сегодня одно из наиболее активно развивающихся направлений исследований в области технического зрения и искусственного интеллекта. Оно имеет широкий спектр приложений для обеспечения общественного порядка и безопасности, включая контроль доступа, сбор статистики о движении толпы и анализ очередей, детектирование и анализ человеческого поведения и т.д. [56-60]. Системы подобного класса включают в себя комплекс алгоритмов, решающих разные задачи на разных этапах анализа. В центре внимания данной работы находятся алгоритмы детектирования и распознавания объектов, основанные на методах машинного обучения. В качестве объекта анализа выбрано лицо человека.

Рассмотрим пример системы автоматизированного видеонаблюдения, в которой объектом анализа выступает лицо человека (рис. 1.1) [114, 115].



Рис 1.1. Схема биометрической идентификации на основе анализа области лица на изображении

Телевизионные изображения поступают по радиоканалу в центр обработки информации, где поэтапно осуществляется их анализ.

На первом этапе решается задача обнаружения интересующего объекта, в данном случае – автоматическое детектирование лиц [61-64]. Алгоритм автоматического детектирования лиц решает следующую задачу: по произвольному изображению на входе определить, имеются ли на этом изображении лица, и если да, то указать, где находится каждое лицо и каков его размер. Работа большинства алгоритмов детектирования лиц на изображениях заключается в сканировании входного изображения окном, имеющим определенную форму и различный масштаб, и в определении к какому классу относится изображение внутри этого окна («лицо» либо «не лицо»). Таким образом, задача детектирования лиц на изображениях сводится к построению классификатора, эффективно разделяющего классы «лиц» и «не лиц» [22].

За последние несколько лет было предложено множество алгоритмов обнаружения лиц, использующих различные подходы. Основные методы детектирования лиц на изображениях можно разделить на четыре категории [19].

1. Методы, основанные на знаниях. Сначала осуществляется поиск определенных черт лица (глаза, нос, рот) на входном изображении. Затем найденные кандидаты проверяются на соответствие закодированным законам, которые используют человеческие знания о том, что собой представляет типичное человеческое лицо. Законы описывают взаимозависимости между чертами лица, представленные их положением и расстоянием между ними. Например, лицо на изображении обычно имеет два глаза, расположенных симметрично относительно носа и рта.

2. Методы на основе инвариантных свойств. Эти алгоритмы стараются найти инвариантные свойства, которыми обладают области изображения, где находится лицо, даже при изменении условий освещения, выражения лица и его положения по отношению к камере. Сначала с помощью глобальных

свойств, таких как цвет кожи, размер и форма лица, находятся кандидаты. Затем осуществляется проверка отобранных кандидатов с помощью таких черт лица как брови, глаза, нос, рот и волосы. Черты лица обычно находятся с помощью краевых детекторов.

3. Методы на основе сравнения с шаблоном. В этом случае выбирается несколько шаблонов лиц или отдельных черт лиц, чтобы в дальнейшем определить местонахождение лица, посчитав корреляции между входным изображением и этими шаблонами. Шаблоны лиц (обычно фронтальных) задаются вручную непосредственно или в виде параметров некоторой функции. Для того чтобы справляться с разнообразием форм и размеров, применяются масштабируемые, деформируемые шаблоны, шаблоны с переменным разрешением.

4. Методы на основе обучения. Алгоритмы на базе методов обучения используют математические модели, которые обучаются с помощью набора тренировочных изображений. Затем обученные модели используются для решения задачи обнаружения лиц на изображениях.

Методы, входящие в первые три категории, имеют существенные недостатки. Так, недостатком методов, основанных на знаниях, является то, что очень сложно преобразовать человеческие знания в хорошо определенные законы. Если составить очень детальные (строгие) законы, то система будет отбрасывать лица, которые не удовлетворяют им полностью. Если законы будут слишком общими, то это приведет к большому числу неверных обнаружений лиц. К тому же тяжело составить правила для обнаружения лиц в различных позах, поскольку необходимо предусмотреть все возможные случаи.

Проблемой методов на базе неизменных характерных черт является то, что эти черты могут быть серьезно повреждены шумами, а также при засвечивании и затемнении изображения. Лицо может иметь еле заметную границу, в то время как тени могут дать многочисленные и четкие края, что приведет к некорректной работе алгоритма.

Методы на основе сравнения с шаблоном демонстрируют низкий уровень выделения лиц, так как они не могут эффективно справляться с разнообразием форм, поз и размеров.

Методы на основе обучения лишены перечисленных выше недостатков, и поэтому считаются более эффективными. Среди них наибольшее распространение в настоящее время получили такие подходы как метод главных компонент, линейный дискриминантный анализ, искусственные нейронные сети, метод опорных векторов [26-30]. Сравнительный анализ алгоритмов выделения лиц на изображениях на основе обучения приведен в работе [65]. Один из лучших результатов демонстрирует подход, предложенный П. Виолой и М. Джонсом [66], основанный на процедуре адаптивного бустинга [67]. Данный алгоритм реализован в библиотеке компьютерного зрения OpenCV [68]. Другой алгоритм, описанный в работе [29], базируется на обучающей сети SNoW (Sparse Network of Winnows), что в переводе означает разреженная сеть просеивающих элементов (РСПЭ). Третий алгоритм, представленный в работе [28], основан на методе опорных векторов (МОВ). Результаты тестирования перечисленных алгоритмов в условиях наличия шумов и искажений, характерных для радиотехнических устройств, приведены в работах [69, 113]. Однако устойчивость трех выбранных детекторов к изменению условий освещенности ранее не исследовалась. Этому вопросу посвящен второй раздел данной диссертационной работы. Структура и алгоритмические особенности исследуемых алгоритмов детектирования лиц приведены в разделе 1.2.

Следует отметить, что этап детектирования объекта имеет важнейшее значение, поскольку возможные неточности в определении положения лица способны оказать значительное влияние на работоспособность всей системы в целом [70].

На втором этапе в системе интеллектуального анализа видеоданных решается задача сопровождения объектов [71-76]. Данная операция отвечает

за отслеживание объектов между моментами детектирования. Движения людей имеют непредсказуемый характер и не во всех кадрах видеопоследовательности лицо человека отчетливо видно, и может быть детектировано. Кроме того, поскольку поиск лиц производится по всему изображению и при различных масштабах, операция детектирования является вычислительно затратной и вносит задержки в работу всей системы. Поэтому использовать детектирование в каждом кадре нецелесообразно. В силу описанных выше причин сопровождение лиц становится важным компонентом системы для непрерывного отслеживания положения и размеров объекта [77].

В настоящее время существует несколько подходов к реализации сопровождения объектов на видеопоследовательностях. Для применения в задачах реального времени часто используются методы, основанные на вычислении оптического потока. В общем случае под оптическим потоком можно понимать двумерную проекцию движения объекта в пространстве на плоскость изображения в виде траекторий пикселей. Оптический поток может вычисляться как на основе слежения за всеми пикселями изображения (полный оптический поток), так и на основе слежения за точечными особенностями изображения (разреженный оптический поток) [78].

Основным недостатком полного оптического потока является низкая устойчивость к шумам. В силу этого более практичным является вычисление разреженного потока. Одним из алгоритмов, основанных на вычислении оптического потока, является алгоритм, предложенный В. Lucas и Т. Kanade [79]. Он был использован для вычисления координат ключевых точек объекта в текущем кадре на основе их координат в предыдущем кадре. Определение нового положения и размеров лица осуществлялось на основе алгоритма, предложенного в работе [80]. Исследование устойчивости используемого алгоритма сопровождения лиц к шумам в радиотехнических устройствах приведено в работе [81]. Следует отметить, что в данной диссертационной работе не исследуется возможность модификации этапа

сопровождения лиц, поскольку этот этап служит повышению быстродействия системы, и оказывает меньшее влияние на результат распознавания по сравнению с этапом детектирования лиц.

На третьем этапе анализа видеоданных детектированные и отобранные алгоритмом сопровождения фрагменты поступают на вход алгоритмов распознавания. Задача распознавания объектов на телевизионных изображениях заключается в определении на основе многочисленных характеристик некоторого объекта одной или нескольких наиболее существенных его характеристик, недоступных для непосредственного определения, в частности, его принадлежности к определенному классу объектов. Другими словами, алгоритмы распознавания разделяют анализируемые фрагменты на категории с тем, чтобы получить новую информацию о свойствах интересующего объекта.

В данной диссертационной работе рассматривается задача распознавания пола человека по изображению его лица. Распознавание пола – бинарная задача разделения всех фрагментов на два класса («мужчины» и «женщины»). Сравнение современных методов обучения классификатора применительно к задаче распознавания пола приведено в работе [82]. Наиболее эффективными методами признаны метод опорных векторов (МОВ) [83], алгоритм на базе линейного дискриминантного анализа KDDA (Kernel Direct Discriminant Analysis) [84] и алгоритм на базе адаптивных признаков и метода опорных векторов (АП-МОВ) [85]. Эти алгоритмы выбраны в качестве базовых для проведения первоначального анализа работы классификатора на собственной базе лиц. Описание принципов их работы приведено в разделе 1.3. Данные о результатах работы выбранных алгоритмов, имеющиеся в литературе, получены на различных обучающих и тестовых наборах, что и вызвало необходимость тестирования алгоритмов в равных условиях на одинаковой базе изображений (раздел 1.4.2).

В третьем разделе данной диссертационной работы проводится исследование и модификация базового алгоритма распознавания пола на

основе метода опорных векторов для повышения его устойчивости к изменению условий освещенности. Четвертый раздел посвящен усовершенствованию базового алгоритма при работе в нормальных условиях освещенности путем расширения пространства признаков классификатора.

Заключительный этап рассмотренной системы анализа видеоданных заключается в идентификации личности человека по детектированной области его лица [14]. Алгоритмы подобного класса [21, 25] предполагают поиск соответствия анализируемого лица определенному человеку из некоторой заранее заданной базы лиц людей. В данной работе предлагается использовать информацию о гендерной принадлежности человека, полученную на этапе распознавания пола, для сужения области поиска при идентификации личности. Исходя из предположения о приблизительно равном количестве мужчин и женщин в базе лиц, успешное распознавание пола позволяет вдвое сократить время работы алгоритма идентификации личности и исключить возможность ошибки идентификации в пользу лиц, чей пол противоположен распознанному.

Следует отметить, что непосредственно сами алгоритмы идентификации личности не рассматриваются в данной работе. Внимание уделено алгоритмам детектирования лиц и распознавания пола, чей результат служит алгоритмам идентификации входными данными. На рис. 1.2 приведен визуальный пример системы автоматического анализа данных, поступающих с видеокамеры, установленной в помещении. Отобранные объекты выделены рамками, цвет которых условно показывает распознанный пол (синий – «мужчины», красный – «женщины»).

Кроме точечных данных для каждого захваченного объекта система способна получать статистические данные, такие как подсчет общего количества людей каждого пола в зоне видимости камеры за определенный промежуток времени, скорость захваченных объектов и время их пребывания в зоне видимости. Подобная информация востребована в системах

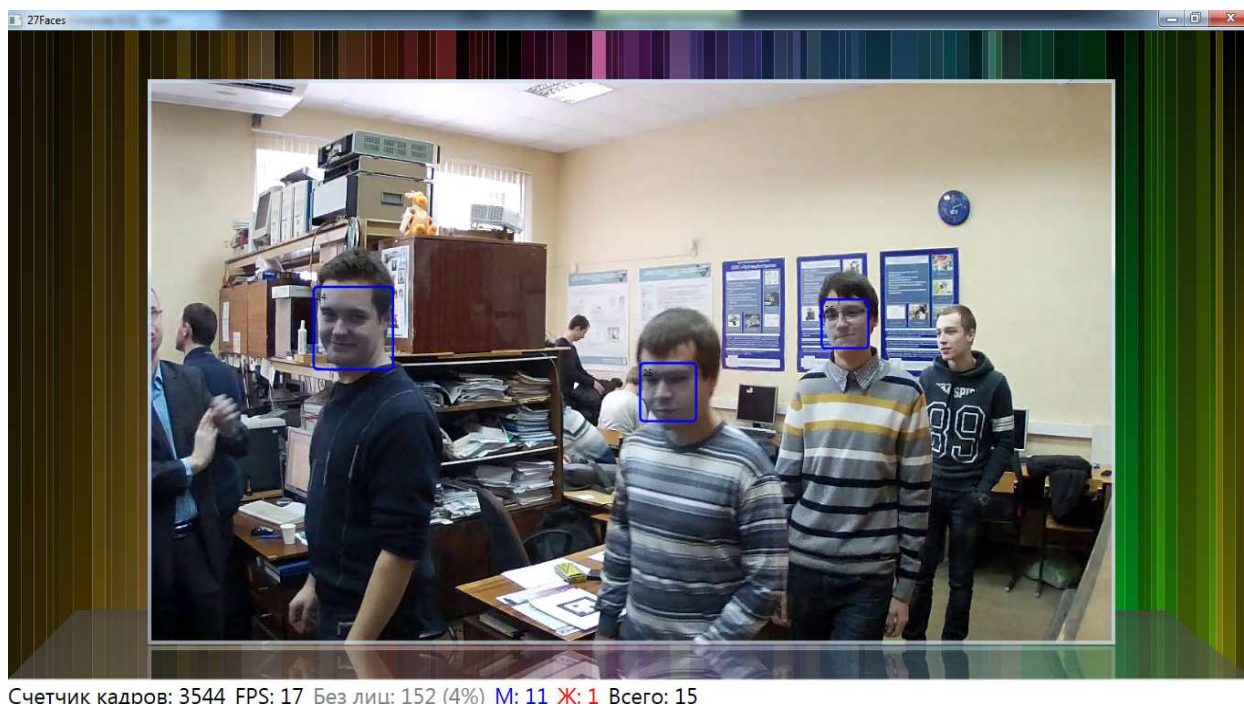


Рис. 1.2. Визуальный пример работы алгоритмов распознавания лиц
на телевизионных изображениях

автоматизированного отношения с клиентами, решающих задачу анализа гендерного состава аудитории, например, с целью индивидуального подхода к предоставлению рекламы и услуг клиентам в зависимости от их гендерной принадлежности.

Перейдем к рассмотрению принципов работы алгоритмов распознавания лиц, выбранных в качестве базовых по результатам анализа литературы. Рассмотрим методику и результаты тестирования рассмотренных алгоритмов по предлагаемой собственной базе обучающих и тестовых изображений.

1.2. Алгоритмы детектирования лиц на телевизионных изображениях

Рассмотрим структуру и принцип работы трех современных алгоритмов детектирования лиц на базе обучения. Как уже отмечалось, первый алгоритм, предложенный П. Виолой и М. Джонсом в работе [66], использует процедуру обучения, основанную на бустинге. Второй алгоритм, описанный в работе [29], базируется на обучающей сети SNoW (Sparse

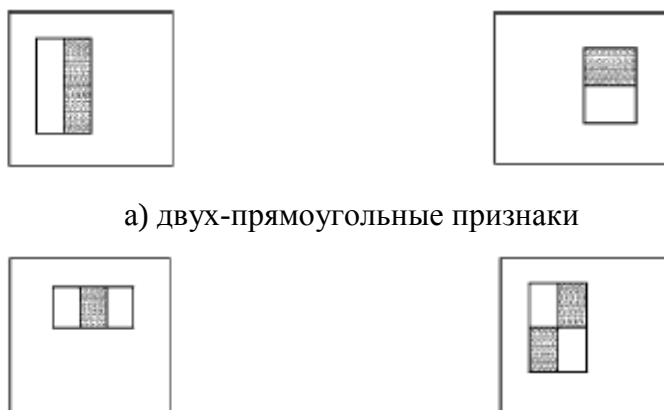
Network of Winnows). Третий алгоритм, представленный в работе [28], основан на методе опорных векторов (МОВ).

1.2.1. Алгоритм на базе адаптивного бустинга

Данный алгоритм состоит из трех этапов.

1. *Переход к интегральному изображению.* Интегральное изображение позволяет быстро вычислять признаки изображения, которые используются классификатором. Основная причина использования признаков вместо значений пикселей в том, что признаки позволяют закодировать полученную в результате обучения информацию. К тому же системы, работающие с признаками, гораздо быстрее систем, оперирующих со значениями пикселей.

Предлагаемый алгоритм использует три вида простых признаков. Значение двух-прямоугольного признака вычисляется как разность между суммами значений пикселей, принадлежащих двум прямоугольным областям. Области имеют одинаковую длину и ширину и ориентированы вертикально или горизонтально, как показано на рис. 1.3.



а) двух-прямоугольные признаки

б) трех-прямоугольный признак в) четырех-прямоугольный признак

Рис. 1.3. Примеры прямоугольных признаков и их расположения внутри окна детектора

Значение трех-прямоугольного признака вычисляется как разность между суммой значений пикселей в двух внешних прямоугольниках и суммой значений пикселей в центральном прямоугольнике. И, наконец,

четырёх-прямоугольный признак находится как разность между суммами значений пикселей, расположенных в диагональных парах прямоугольников.

Прямоугольные признаки могут быть быстро вычислены с помощью промежуточного представления изображения, которое носит название – интегральное изображение. Интегральное изображение в точке (x, y) содержит сумму пикселей, расположенных слева и сверху над этой точкой, и определяется по формуле:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'), \quad (1.1)$$

где $ii(x, y)$ – интегральное изображение, $i(x, y)$ – исходное изображение. Для вычисления используется следующая пара рекурсий:

$$\begin{aligned} s(x, y) &= s(x, y-1) + i(x, y) \\ ii(x, y) &= ii(x-1, y) + s(x, y) \end{aligned} \quad (1.2)$$

где $s(x, y)$ – кумулятивная строка суммы, $s(x, -1) = 0$, $ii(-1, y) = 0$.

Следовательно, интегральное изображение может быть вычислено за один проход по первоначальному изображению. Используя интегральное изображение, вычисление признаков одинакового вида, но с разными геометрическими параметрами проходит за одинаковое время.

2. Построение классификатора на основе алгоритма бустинга. Сложный классификатор можно рассматривать как суперпозицию простых классификаторов порогового типа. Обозначим простой классификатор как:

$$h_j(x) = \begin{cases} 1, & \text{если } p_j f_j(x) < p_j \theta_j \\ 0 & \text{в остальных случаях} \end{cases}, \quad j = 1 \dots N, \quad (1.3)$$

где p_j – показывает направление знака неравенства, θ_j – значение порога, $f_j(x)$ – вычисленное значение признака, x – входное изображение, N – общее количество признаков.

Рассмотрим алгоритма бустинга [67], с помощью которого осуществляется отбор простых классификаторов и формирование на их основе сложного классификатора. Зададим обучающий набор $X_p = \{x_{p1} \dots x_{pL}\}$, содержащий L изображений лиц разрешением 24×24 пикселя и обучающий

набор $X_n = \{x_{n1} \dots x_{nM}\}$, содержащий M изображений не лиц того же разрешения. Пусть $w_{1,i} = \frac{1}{2M}$, $v_{1,i} = \frac{1}{2L}$, где $w_{1,i}$ и $v_{1,i}$ – начальные веса для изображений обучающего набора «лиц» и «не лиц» соответственно. Тогда в цикле $t=1 \dots T$, где T – число отбираемых простых классификаторов, произведем следующие операции:

а) нормализация весов:

$$\begin{aligned} w_{t,i} &\leftarrow \frac{w_{t,i}}{\sum_{j=1}^M w_{t,j} + \sum_{j=1}^L v_{t,j}}, & i=1 \dots M; \\ v_{t,i} &\leftarrow \frac{v_{t,i}}{\sum_{j=1}^M w_{t,j} + \sum_{j=1}^L v_{t,j}}, & i=1 \dots L; \end{aligned} \quad (1.4)$$

б) вычисление ошибки классификации:

$$\sigma_t(f, p, \theta) = \sum_{i=1}^M w_{t,i} h(x_{ni}, f, p, \theta) + \sum_{i=1}^L v_{t,i} [1 - h(x_{pi}, f, p, \theta)]; \quad (1.5)$$

в) выбор классификатора с минимальной ошибкой:

$$h_t(x) = h(x, f_t, p_t, \theta_t), \quad (1.6)$$

где f_t , p_t и θ_t – параметры, при которых функция ошибки $\sigma_t(f, p, \theta)$ принимает минимальное значение;

г) адаптивное обновление весов:

$$\begin{aligned} w_{t+1,i} &\leftarrow w_{t,i} \beta_t^{1-h_t(x_{ni}, f_t, p_t, \theta_t)}, & i=1 \dots M; \\ v_{t+1,i} &\leftarrow v_{t,i} \beta_t^{h_t(x_{pi}, f_t, p_t, \theta_t)}, & i=1 \dots L; \end{aligned} \quad (1.7)$$

где коэффициент обновления весов β_t определяется по следующей формуле:

$$\beta_t = \frac{\sigma_t(f_t, p_t, \theta_t)}{1 - \sigma_t(f_t, p_t, \theta_t)}. \quad (1.8)$$

В результате работы алгоритма бустинга на каждой итерации формируется простой классификатор, который имеет минимальную ошибку по отношению к текущим значениям весов, задействованных в процедуре обучения для определения ошибки. После процедуры обновления весов, подчеркиваются те образцы, которые были неправильно классифицированы.

Тем самым, на текущей итерации обучения следующий простой классификатор будет сформирован таким образом, чтобы не допустить ошибку на этих элементах обучающей выборки.

На рис. 1.4 представлены первый и второй признаки, отобранные в результате работы алгоритма бустинга, на фоне окна детектора и на фоне типичного лица, используемого для обучения.

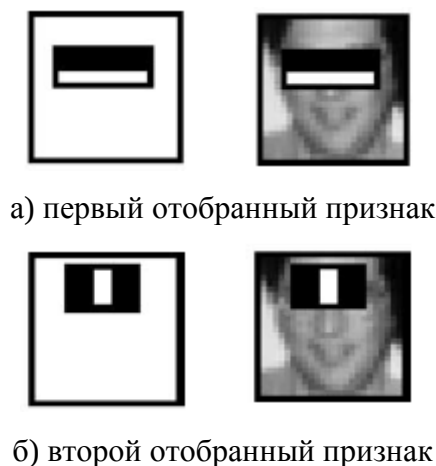


Рис. 1.4. Первый и второй признаки, отобранные алгоритмом бустинга

Первый отобранный признак основан на том свойстве, что область глаз чаще всего темнее области носа и щек. Этот признак имеет большой размер по сравнению с окном детектора, и потому не чувствителен к размеру и положению лица. Второй отобранный признак опирается на то, что глаза темнее спинки носа. Алгоритм использует окно детектора разрешением 24×24 пикселя и коэффициент масштабирования окна 1,25.

Отобранные простые классификаторы формируют сложный классификатор по следующему правилу:

$$C(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq Q \\ 0 & \text{в остальных случаях} \end{cases}, \quad (1.9)$$

где α_t – коэффициенты линейной комбинации, а Q – оптимальный порог классификатора, вычисляющиеся по формулам:

$$\alpha_t = \log \frac{1}{\beta_t}, \quad Q = \frac{1}{2} \sum_{t=1}^T \alpha_t. \quad (1.10)$$

3. Комбинирование классификаторов в каскадную структуру.

Структура каскадного детектора приведена на рис. 1.5. Каскад состоит из слоев, которые представляют собой классификаторы, обученные с помощью процедуры бустинга.

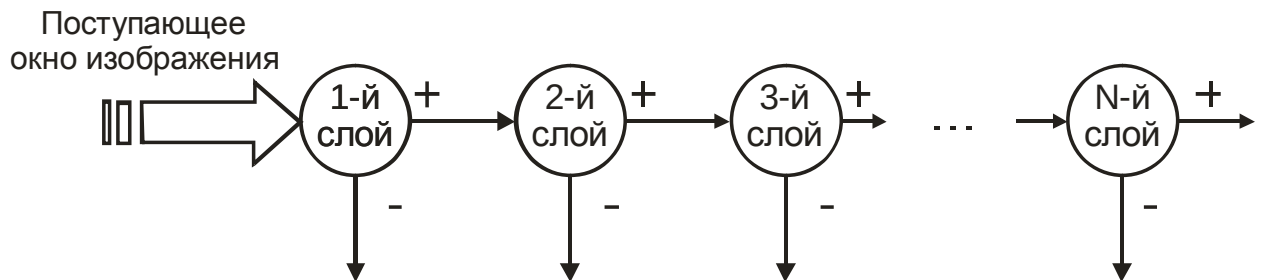


Рис. 1.5. Структура каскадного детектора алгоритма на базе бустинга

На вход каскада подается скользящее окно изображения, которое попадает на первый слой. Если текущий слой классифицирует окно как негативное (не являющееся лицом), то происходит выход, и анализ последующими слоями не выполняется. Если же текущий слой определяет окно как позитивное (являющееся лицом), то окно анализируется следующим слоем. Такой процесс происходит до тех пор, пока не будут пройдены все слои при положительном решении предыдущих слоев или не произойдет выход по негативному решению одного из слоев. Учитывая тот факт, что на изображении количество фоновых скользящих окон во много раз превышает количество окон, на которых присутствуют лица, то выигрыш очевиден. Скорость обработки изображения при этом возрастает в 10-20 раз в зависимости от изображения, но при этом не теряется качество распознавания.

При конструировании каскада, его структура определяется исходя из того, какие целевые требования к нему предъявляются.

В результате обучения определяются следующие параметры каскада:

- количество слоев;

- количество простых классификаторов в каждом слое;
- пороговое значение каждого слоя.

Задаются требуемые значения целевых параметров для каждого слоя. Каждый слой формируется с помощью процедуры бустинга, в процессе которой количество простых классификаторов увеличивается до тех пор, пока не будут достигнуты требуемые значения коэффициентов. Во время процедуры обучения данные разделяются на две выборки: обучающую и контрольную. Формирование простых классификаторов происходит с использованием обучающей выборки, а определение требуемых параметров – с использованием контрольной выборки. По ходу формирования слоев, каждый новый слой обучается на тех негативных экземплярах, которые были ошибочно классифицированы предыдущими слоями, благодаря этому каскад обладает высокими обобщающими способностями и низким коэффициентом ошибки. При тестировании использовался каскад «default» из библиотеки OpenCV с параметрами, заданными по умолчанию.

1.2.2. Алгоритм на базе разреженной сети просеивающих элементов

Алгоритм на базе РСПЭ также можно условно разбить на три этапа.

1. *Переход от значений пикселей к локальным SMQT признакам.* SMQT [106] (Successive Mean Quantization Transform) – это преобразование, которое позволяет извлечь из локальной области изображения составляющую, не зависящую от освещенности. Оно заключается в квантовании области изображения с порогом квантования, равным среднему значению пикселей, входящих в эту область. Локальная область может быть задана различными способами: как блок пикселей определенного разрешения; как набор интерполированных значений точек, находящихся на заданном расстоянии от фиксированного центра; либо каким-то другим произвольным образом. В любом случае, если обозначить значение одного пикселя как x , то локальная область изображения может быть представлена в виде вектора $D(x)$, содержащего набор значений пикселей. Тогда результатом SMQT

преобразования будет вектор той же размерности $M(x)$, содержащий новый набор значений. Вычисление вектора $M(x)$ осуществляется в три этапа. Сначала находится среднее значение пикселей в области $D(x)$:

$$\bar{D} = \frac{1}{N} \sum_{i=1}^N D(x_i). \quad (1.11)$$

Затем производится операция квантования по следующему правилу:

$$M(x_i) = \begin{cases} 1, & D(x_i) > \bar{D} \\ 0, & \text{иначе} \end{cases}. \quad (1.12)$$

В дальнейшем производится представление вектора в виде индекса:

$$m = \sum_{i=1}^N M(x_i) \cdot 2^{i-1}. \quad (1.13)$$

Разработчики алгоритма задают локальную область $D(x)$ как блок разрешением 3×3 пикселя. Для нахождения лиц на изображении используется окно детектора разрешением 32×32 пикселя. Сканирование осуществляется с шагом в один пиксель, как по горизонтали, так и по вертикали. Для того чтобы выделять лица разного размера, изображение многократно масштабируется с коэффициентом масштабирования 1,2. С целью избежания краевых эффектов при возможной фильтрации, к окну детектора применяется овальная маска, содержащая 648 пикселей (рис. 1.6). Каждому из этих 648 пикселей соответствует свой SMQT признак.



а) изображение из тренировочного набора



б) соответствующие этому изображению SMQT признаки

Рис. 1.6. Маскировка пикселей изображения и соответствующих им SMQT признаков

Переход к SMQT признакам позволяет алгоритму адаптироваться к изменению освещения объекта, так как SMQT признаки извлекают из изображения не зависящие от освещения компоненты.

2. *Классификатор на базе обучающей сети SNoW.* Обучающая архитектура SNoW (Sparse Network of Winnows) представляет собой разреженную сеть линейных элементов в пространстве признаков [107]. Большим достоинством сети SNoW является возможность создания весовых таблиц для классификации. Пусть W – это набор SMQT признаков $M(x)$, тогда классификатор

$$\theta = \sum_{x \in W} h_x^{nonface}(M(x)) - \sum_{x \in W} h_x^{face}(M(x)) \quad (1.14)$$

может быть получен посредством использования весовой таблицы для не лиц $h_x^{nonface}$ и весовой таблицы для лиц h_x^{face} , и определения порога для θ . Поскольку обе таблицы работают в одном домене, их можно объединить в одну весовую таблицу:

$$h_x = h_x^{nonface} - h_x^{face}. \quad (1.15)$$

Для обучения классификатора использовалась база обучающих изображений лиц и не лиц. Изображения лиц были получены с помощью web-камеры. Далее на них вручную были отмечены три точки: правый глаз, левый глаз и центральная точка внешней границы верхней губы. После этого лицо деформировалось в блок 32×32 пикселя с различным расположением опорных точек (рис. 1.7).



Рис. 1.7. Получение набора обучающих изображений лиц

Деформация необходима для того, чтобы имитировать различное положение лица по отношению к камере. В результате описанной процедуры был получен обучающий набор лиц емкостью порядка 1 миллиона образцов.

База обучающих изображений не лиц первоначально состояла из случайным образом сгенерированных изображений. После того, как она была дополнена ошибками классификации, ее емкость также составила порядка 1 миллиона образцов.

Каждому обучающему изображению соответствует набор SMQT признаков. Пусть тренировочная база состоит из $i=1,2,...,K$ наборов SMQT признаков $M_i(x)$ и соответствующих каждому набору значений классов c_i («лицо» или «не лицо»). Тогда весовые таблицы лиц и не лиц могут быть обучены с помощью закона обновления весов, получившего название Winnow Update Rule [108]. Изначально обе таблицы содержат нули. При первом обращении к элементу таблицы в процессе обучения ему присваивается значение 1. Три параметра системы задаются вручную: порог γ , повышающий коэффициент $\alpha > 1$ и понижающий коэффициент $\beta < 1$. Их значения были заданы следующим образом: $\alpha = 1.005$, $\beta = 0.995$, $\gamma = 200$. Если

$\sum_{x \in W} h_x^{face}(M_i(x)) \leq \gamma$ и c_i – лицо, то веса обновляются согласно формуле:

$$h_x^{face}(M_i(x)) = \alpha h_x^{face}(M_i(x)), \quad \forall x \in W. \quad (1.16)$$

Если $\sum_{x \in W} h_x^{face}(M_i(x)) > \gamma$ и c_i – лицо, то значение весов уменьшается:

$$h_x^{face}(M_i(x)) = \beta h_x^{face}(M_i(x)), \quad \forall x \in W. \quad (1.17)$$

Эта процедура повторяется до достижения неизменного значения весов. Обучение весовой таблицы для не лиц происходит точно таким же образом. И, наконец, общая весовая таблица находится по формуле:

$$h_x = h_x^{nonface} - h_x^{face}. \quad (1.18)$$

3. *Организация каскадной структуры.* С целью увеличения быстродействия алгоритма полный классификатор SNoW разбивается на более слабые классификаторы, которые организуются в каскадную структуру (алгоритм Split-up SNoW). При этом не требуется дополнительного обучения слабых классификаторов. Пусть все возможные значения одного признака ограничиваются множеством $P_i, i=1,2,\dots,2^N$, тогда вклад каждого признака в разделение классов можно оценить по формуле:

$$v_x = \sum_{i=1}^{2^N} |h_x(P_i)|, \quad \forall x \in W. \quad (1.19)$$

Отсортировав признаки по соответствующим им значениям v_x , можно получить рейтинг лист признаков (importance list). Пусть $W' \in W$ – это подпространство, содержащее определенное количество признаков с наибольшим рейтингом. Тогда θ' можно определить как слабый классификатор, который выделяет все лица из тренировочной базы, но имеет больший по сравнению с полным классификатором уровень неверного выделения:

$$\theta' = \sum_{x \in W'} h_x(M(x)). \quad (1.20)$$

Увеличить количество слабых классификаторов в каскаде можно за счет тех же самых операций, выбирая подпространства следующим образом: $W' \in W'' \in \dots \in W$. Разбиения происходят при 20, 50, 100, 200 и 648 признаках, как показано на рис. 1.8.

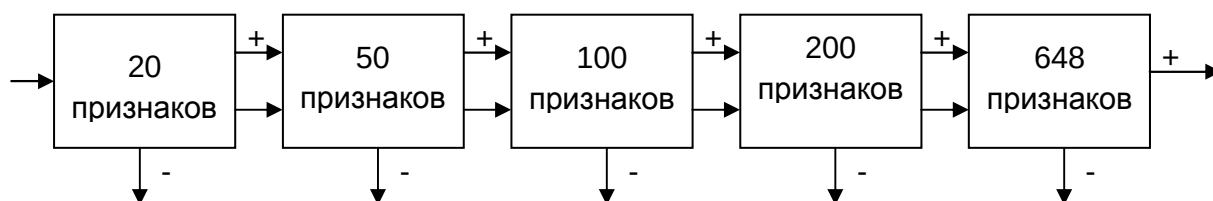


Рис. 1.8. Каскадная структура алгоритма на базе РСПЭ

Каскад состоит из 5 слоев. Признаки в каждый слой каскада отбираются согласно их рейтингу (importance list). Каждый слой отбирает кандидатов для следующего слоя и передает ему значения уже посчитанных для этих кандидатов SMQT признаков.

1.2.3. Алгоритм на базе метода опорных векторов

Этот алгоритм также удобно разбить на три этапа.

1. *Построение классификатора на базе метода опорных векторов (МОВ).* Применение метода опорных векторов [109] к задаче обнаружения лиц заключается в поиске оптимальной гиперплоскости в признаковом пространстве, отделяющей класс изображений лиц от изображений не лиц. В качестве признаков служат значения пикселей изображения, представленные в виде n -мерного вектора. Оптимальной считается гиперплоскость, которая максимизирует ширину полосы между классами (рис. 1.9).

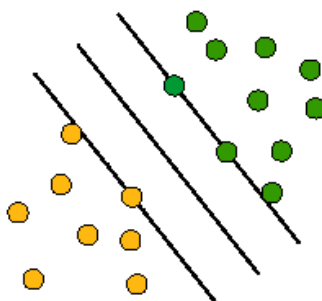


Рис. 1.9. Пример разделяющей полосы классификатора на базе МОВ

Разделяющая гиперплоскость определяется как линейная комбинация небольшого набора тренировочных векторов, называемых опорными векторами. Обозначим набор собственных векторов как $\{X_1, \dots, X_m\}$, а соответствующие им коэффициенты линейной комбинации – $\{\alpha_1, \dots, \alpha_m\}$. МОВ – это линейный классификатор, поэтому для разделения линейно неразделимых классов применяется неявное проецирование векторов-признаков в пространство потенциально намного более высокой размерности (еще выше, чем пространство изображений), в котором классы могут оказаться линейно разделимыми (рис. 1.10).

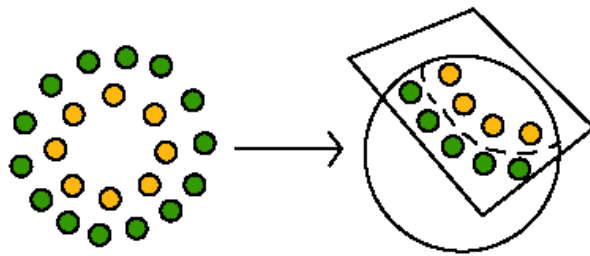


Рис. 1.10. Пример разделения линейно неразделимых классов с помощью перехода к пространству более высокой размерности

Проецирование осуществляется с помощью аппарата ядерных функций. Неявное проецирование с помощью ядерных функций не приводит к усложнению вычислений, что позволяет успешно использовать линейный классификатор для линейно неразделимых классов. Формально решающее правило классификатора на базе МОВ для входного изображения X может быть записано следующим образом:

$$f(X) = \text{sgn} \left(\sum_{i=1}^m y_i \alpha_i k(X_i, X) + b \right), \quad (1.21)$$

где $k(\cdot, \cdot)$ – ядерная функция, а b – смещение. Для нахождения разделяющей гиперплоскости использовался тренировочный набор, содержащий около 13 тысяч изображений лиц с разрешением 19×19 пикселей и около 36 тысяч изображений не лиц. Для увеличения размера окна детектора применялось

масштабирование с коэффициентом 3/4. В качестве ядерной функции использовалась радиальная базисная функция Гаусса $k(x, x') = \exp(-\frac{\|x - x'\|^2}{2\sigma^2})$ при $\sigma = 10$. Обученный классификатор содержал $m = 6910$ опорных векторов.

2. *Вычисление аппроксимаций для классификатора МОВ.* Для ускорения принятия решения обученный классификатор на базе МОВ, содержащий $m = 6910$ опорных векторов, аппроксимируется классификатором с меньшим набором опорных векторов $\{Z_1, \dots, Z_{m'}\}$, где $m' < m$. При этом быстродействие алгоритма увеличивается в m/m' раз [110, 111]. Решающее правило такого классификатора определяется по формуле:

$$f'(X) = \text{sgn}\left(\sum_{i=1}^{m'} \beta_i k(Z_i, X) + b\right). \quad (1.22)$$

Для нахождения β_i и Z_i необходимо решить уравнение

$$\min \left\| \sum_{i=1}^m y_i \alpha_i k(X_i, \cdot) - \sum_{i=1}^{m'} \beta_i k(Z_i, \cdot) \right\|^2. \quad (1.23)$$

На практике β_i и Z_i находятся при помощи градиентного метода оптимизации [112].

3. *Комбинирование классификаторов в каскадную структуру.* С целью увеличения быстродействия алгоритма и уменьшения ошибки классификации классификатор на базе МОВ организуется в каскадную структуру, аналогичную каскадной структуре алгоритма на базе РСПЭ (рис. 1.11).

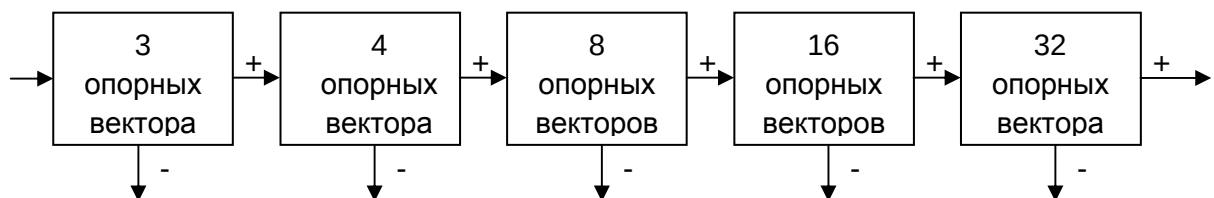


Рис. 1.11. Каскадная структура алгоритма на базе МОВ

Каскад состоит из 5 слоев, в которых классификатор МОВ аппроксимируется уменьшенным набором из 3, 4, 8, 16 и 32 опорных векторов соответственно.

1.3. Построение бинарных классификаторов для распознавания пола людей по детектированной области лица

1.3.1. Алгоритм на базе линейного дискриминантного анализа

Рассмотрим принцип построения классификатора KDDA, основанного на линейном дискриминантном анализе. Основная идея заключается в том, чтобы представить входные обучающие изображения в виде n -мерных векторов, а затем найти такое подпространство, в котором проекции векторов, принадлежащих разным классам, были бы расположены как можно дальше друг от друга, а проекции векторов, принадлежащих одному и тому же классу, наоборот, – ближе друг к другу.

Однако в случае анализа лиц классы могут оказаться линейно неразделимыми. Для решения проблемы линейной неразделимости классов алгоритм KDDA использует неявное проецирование векторов-признаков в пространство потенциально намного более высокой размерности (еще выше, чем пространство изображений), в котором классы могут оказаться линейно разделимыми (рис. 1.10).

В качестве ядра используется радиальная базисная функция Гаусса, определяемая следующим выражением:

$$k(z_1, z_2) = \exp\left(\frac{-\|z_1 - z_2\|^2}{\sigma^2}\right). \quad (1.24)$$

В полученном пространстве признаков считается, что объекты, принадлежащие различным классам, линейно разделимы, и применяется D-LDA (Direct Linear Discriminant Analysis) [86]. Такой выбор связан с тем, что применение классического LDA в задачах распознавания невозможно. К примеру, если используется обучающая выборка, состоящая из изображений

размерностью $40 \times 40 = 1600$ пикселей, значит, размерность матриц межклассовой и внутриклассовой дисперсии (S_{BTW} и S_{WTH}) составляет 1600×1600 . Кроме того, что работа непосредственно с матрицей S_{WTH} затруднительна из-за её размерности, обучающая выборка должна состоять из 1600 изображений для каждого класса, иначе матрица S_{WTH} будет вырожденной. Для решения данной проблемы в LDA использовано предварительное уменьшение размерности с помощью метода главных компонент. Таким образом, размер матрицы S_{WTH} уменьшается за счет удаления близких к нулю собственных значений.

В алгоритме D-LDA применяется другой способ для решения данной проблемы. Основная идея алгоритма заключается в том, что в «нулевом» пространстве S_{WTH} может содержаться значимая информация для дискриминантного анализа, если проекция S_{BTW} не нулевая в данном направлении, и поэтому полезная информация будет утеряна. В отличие от классического LDA, при поиске матрицы для отображения пространства изображений на пространство признаков вначале диагонализировать матрицу S_{BTW} , при этом уменьшается размер матрицы за счет удаления нулевых и близких к нулю собственных значений, которые не содержат полезной информации. Затем диагонализировать матрицу S_{WTH} . Последовательность диагонализации матриц имеет значение, только если матрица S_{WTH} вырожденная, что в нашем случае имеет место быть. Матрица A для проецирования пространства изображения на пространство признаков в алгоритме D-LDA выбирается из следующего условия:

$$A = \operatorname{argmax}_A \frac{|A^T S_{BTW} A|}{|A^T S_{BTW} A + A^T S_{WTH} A|}. \quad (1.25)$$

Пространство признаков строится на основе собственных векторов матрицы A на этапе обучения. На этапе тестирования для принятия

решения, к какому классу относится вектор в пространстве признаков, используется метод ближайших соседей.

1.3.2. Применение метода опорных векторов в задаче гендерной классификации

Применение метода опорных векторов к задаче разделения анализируемых изображений на два класса заключается в поиске оптимальной гиперплоскости в признаковом пространстве, отделяющей один класс от другого. В качестве признаков служат значения пикселей изображения, представленные в виде n -мерного вектора. Оптимальной считается гиперплоскость, которая максимизирует ширину полосы между классами (рис. 1.9).

Как и в случае применения метода опорных векторов к задаче детектирования лиц классификатор формально описывается выражением (1.21). В качестве ядерной функции, как и для алгоритма KDDA, используется радиальная базисная функция Гаусса, описываемая выражением (1.24). Принципиальным отличием обучения классификатора на базе МОВ в задаче гендерной классификации является одинаковая цена ошибки для обоих классов. Используются равные наборы по количеству обучающих изображений для классов «мужчина» и «женщина». Параметры ядерной функции оптимизируются с использованием механизма кроссвалидации, чтобы избежать эффекта переобучения.

Каскадная структура, описанная в разделе 1.2.3, не используется, поскольку число анализируемых фрагментов в задаче гендерной классификации относительно мало, и нет необходимости отсева фрагментов одного из классов.

1.3.3. Алгоритм на базе адаптивных признаков и метода опорных векторов (АП-МОВ)

Основой для построения классификатора является метод опорных векторов с применением ядерного преобразования для внесения нелинейности, который наряду с искусственными нейронными сетями является на сегодняшний день наиболее эффективным с точки зрения качества распознавания [87]. Принципиальным отличием алгоритма АП-МОВ является наличие перехода от значений пикселей в пространство признаков и применения метода опорных векторов уже в этом новом пространстве [88]. При этом признаки не определяются заранее заданным детерминированным законом. Они генерируются на этапе обучения в результате процедуры оптимизации, которая позволяет извлекать из изображения только ту информацию, которая необходима для разделения классов. Схема работы алгоритма АП-МОВ приведена на рис. 1.12.

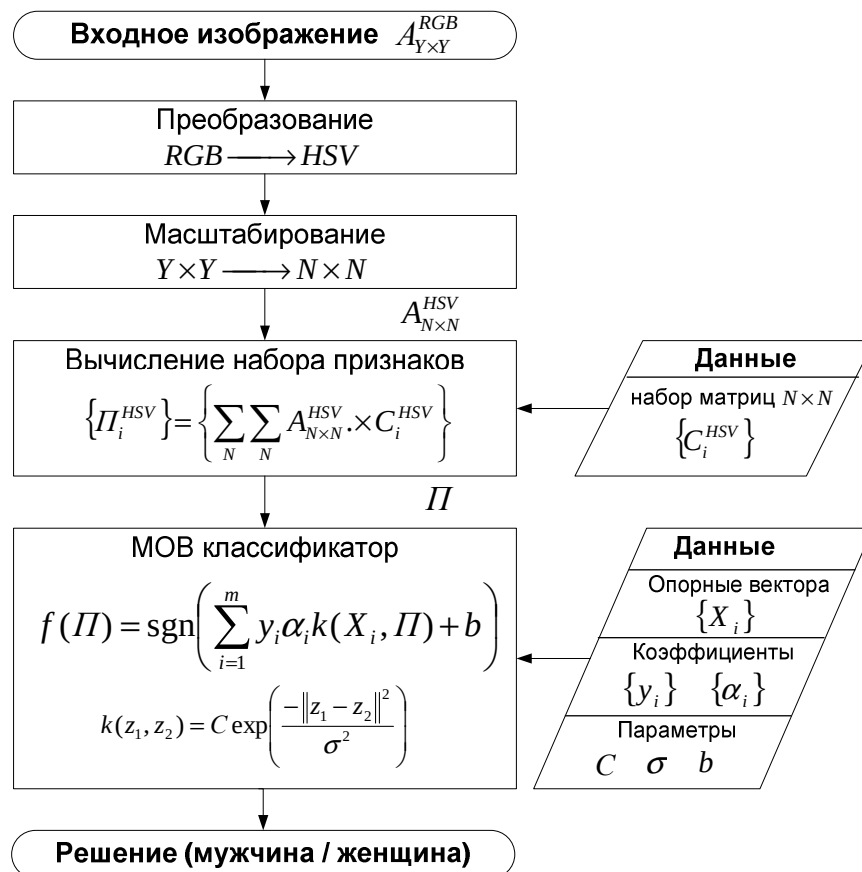


Рис. 1.12. Схема работы алгоритма АП-МОВ

Алгоритм АП-МОВ включает в себя следующие шаги: преобразование цветового пространства, масштабирование, вычисление набора адаптивных признаков и классификация методом опорных векторов с применением ядерного преобразования. Данные, необходимые для вычисления признаков и классификации, получаются в результате обучения классификатора. Обучение классификатора АП-МОВ содержит в себе два независимых этапа:

- генерация признаков;
- формирование и оптимизация классификатора на базе МОВ.

Процедура генерации признаков содержит следующие основные шаги:

– преобразование цветового пространства обучающих фрагментов RGB → HSV (далее все операции выполняются независимо для каждой цветовой компоненты);

– масштабирование обучающих фрагментов до единого разрешения $N \times N$;

– генерация по случайному закону матрицы коэффициентов C_i^{HSV} ;

– вычисление признака для каждого обучающего фрагмента как суммы по строкам и столбцам поэлементного произведения матрицы изображения на матрицу C_i^{HSV} :

$$\Pi_i^{HSV} = \sum_N \sum_N A_{N \times N}^{HSV} \times C_i^{HSV}; \quad (1.26)$$

– вычисление целевой функции оптимизации как отношения квадрата разности средних значений признака, вычисленных для наборов обучающих изображений «мужчин» и «женщин», к сумме их дисперсий:

$$F = \frac{\left(\langle \{ \Pi_i^{HSV} \}_M \rangle - \langle \{ \Pi_i^{HSV} \}_Ж \rangle \right)^2}{\sigma \{ \Pi_i^{HSV} \}_M + \sigma \{ \Pi_i^{HSV} \}_Ж}; \quad (1.27)$$

– далее в цикле итеративно (до достижения фиксированного максимального числа итераций): генерация по случайному закону матрицы $\tilde{C}_i^{HSV}$ внутри фиксированной окрестности матрицы C_i^{HSV} , вычисление

признака $\tilde{P}_i^{HSV}$ для каждого обучающего фрагмента и вычисление целевой функции $\tilde{F}$, переход в новую точку ($F \rightarrow \tilde{F}, C \rightarrow \tilde{C}$), если $\tilde{F} > F$;

- сохранение по истечении некоторого заданного числа итераций матрицы коэффициентов C_i^{HSV} ;
- переход к генерации следующего $i + 1$ признака.

Целью этапа генерации признаков является получение признаков, которые обладают большей разделяющей способностью, то есть имеют высокое значение целевой функции F . Настраиваемыми параметрами на этом этапе являются:

- разрешение обучающих изображений;
- число обучающих изображений;
- максимальное число итераций.

Второй этап обучения состоит из следующих шагов:

- вычисление набора признаков, сгенерированных на предыдущем этапе, для каждого из обучающих фрагментов;
- нормализация признаков;
- обучение МОВ классификатора на обучающем наборе изображений с различными параметрами ядерной функции;
- вычисление уровня верного распознавания (D) по валидационной выборке;
- нахождение оптимальных параметров ядра (максимизирующих D);
- обучение итогового классификатора МОВ с найденными оптимальными параметрами ядра.

Целью этапа формирования классификатора является нахождение решения, обладающего большей обобщающей способностью, то есть имеющего минимальную ошибку классификации. Настраиваемыми параметрами на этом этапе являются:

- число признаков;
- число обучающих изображений;

- параметр ядра σ ;
- параметр ядра C .

1.4. Методика и результаты сравнительного анализа существующих алгоритмов распознавания лиц

Адекватное сравнение алгоритмов распознавания лиц, описанных в разделах 1.2 и 1.3, возможно только при условии использования достаточно большого количества тестовых изображений, одинаковых для всех тестируемых алгоритмов. В работе используются собственные базы изображений как для тестирования алгоритмов детектирования лиц, так и для обучения и тестирования алгоритмов распознавания пола. Рассмотрим параметры используемых баз, метрики тестирования и результаты оценки работы тестируемых алгоритмов.

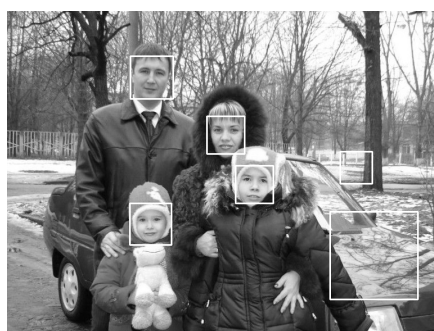
1.4.1. Тестирование алгоритмов детектирования лиц на телевизионных изображениях

Для проведения экспериментов была составлена база данных из 50 цветных изображений, разрешения 768×576 пикселей, суммарно содержащая 213 лиц. На рис. 1.13 представлены примеры изображений из этой базы, на которых с помощью тестируемых алгоритмов выделены лица.

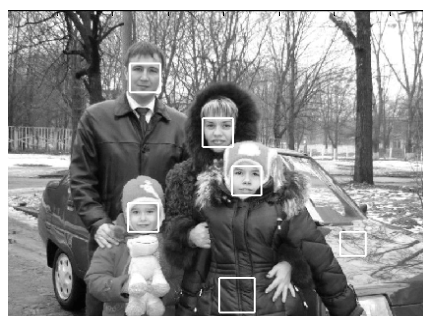
Приведенные примеры иллюстрируют два типа ошибок, возникающих в результате работы системы детектирования лиц: не выделение лица и ложное обнаружение (детектирование объекта, который лицом не является). В связи с наличием двух типов ошибок, существует два основных параметра, характеризующих эффективность работы алгоритмов обнаружения лиц: *уровень детектирования* (detection rate), показывающий процент обнаруженных лиц, и *уровень ложного детектирования* (false positive rate), равный общему числу ложных обнаружений на всем тестовом наборе.

Из рис. 1.13 видно, что одно и то же лицо, выделенное разными алгоритмами, имеет разный размер. Так алгоритм на базе бустинга выделяет

лицо полностью, захватывая лоб, подбородок и щеки. Алгоритмы на базе РСПЭ и МОВ выделяют только глаза, нос и рот, причем алгоритм на базе МОВ выделяет лицо более узким окном. Эти различия вызваны тем, что в процессе создания алгоритмов использовались различные наборы обучающих изображений для построения классификатора. Следует отметить, что алгоритм на базе бустинга более адекватно отображает истинный размер лица.



а)



б)



в)

Рис. 1.13. Примеры изображений, обработанных с помощью: а) алгоритма на базе бустинга; б) алгоритма на базе РСПЭ; в) алгоритма на базе МОВ

В табл. 1.1 представлены результаты сравнения тестируемых алгоритмов между собой по двум параметрам: уровню детектирования и уровню ложного детектирования.

Таблица 1.1

Результаты тестирования алгоритмов детектирования лиц

Алгоритм Параметр	Алгоритм на базе бустинга	Алгоритм на базе РСПЭ	Алгоритм на базе МОВ
Уровень детектирования, %	71,8	87,3	62,4
Количество ложных детектирований	12	34	73

Лидирует по уровню детектирования алгоритм на базе РСПЭ – 87,3%, на втором месте алгоритм на базе бустинга – 71,8%. Однако при этом последний имеет втрое меньше ложных детектирований на тестовом наборе по сравнению с алгоритмом на базе РСПЭ. Алгоритм на базе МОВ уступает как по уровню детектирования, так и по количеству ошибок, занимая последнее место среди тестируемых алгоритмов.

1.4.1. Тестирование алгоритмов распознавания пола людей по детектированной области лица

Для осуществления обучения и тестирования гендерных классификаторов, описанных в разделе 1.3, необходима достаточно объемная база полноцветных изображений, содержащая лица мужчин и женщин. Существует общепринятая база данных для распознавания лиц, носящая название FERET [89]. В ней содержится 14 051 изображение лиц 1 199 людей, которые сняты при различном освещении, либо в разные промежутки времени. Изображения также отличаются выражением лица и его положением по отношению к камере. Цветная версия базы FERET содержит 11 338 изображений лиц 994 людей. Однако база FERET не может быть использована для задачи обучения и тестирования гендерного

классификатора, поскольку содержит недостаточное количество лиц различных людей каждого пола. В связи с этим была составлена собственная база изображений, набранная из различных источников, в соответствии с требованиями, приведенными в табл. 1.2.

Таблица 1.2

Параметры собственной базы изображений

Параметр	Значение
Общее количество изображений	8 654
Число лиц мужчин	5 250
Число лиц женщин	5 250
Минимальное разрешение изображений	640×480
Формат цветового пространства	RGB
Выражение лица	естественное
Положение по отношению к камере	фронтальное
Условия освещения и фон	требования не предъявляются
Возраст людей	от 18 до 65 лет
Раса	европеоидная

На каждом изображении из собственной базы алгоритмом на базе бустинга были выделены лица. Затем вручную были удалены ложные выделения и сформирована база выделенных фрагментов, содержащая 10 500 изображений (по 5 250 на каждый класс). Эта база была разделена на три независимых выборки: обучающую, контрольную и тестовую. Обучающая выборка использовалась для генерации признаков и построения классификаторов на базе MOB и KDDA. Контрольная выборка необходима для того, чтобы избежать эффекта переобучения в процессе оптимизации параметров ядерной функции. Проверка работы обученных классификаторов осуществлялась с использованием тестовой выборки. Пример изображений из собственной базы лиц приведен на рис. 1.14.



Рис. 1.14. Примеры изображений мужчин и женщин из собственной базы лиц

Для представления результатов классификации использовалась так называемая кривая ошибок или ROC-кривая (Receiver Operator Characteristic). Поскольку классов два, один из них называется классом с положительными исходами, второй – с отрицательными исходами. ROC-кривая показывает зависимость количества верно классифицированных положительных примеров от количества неверно классифицированных отрицательных примеров при варьировании порога решающего правила [90]. Преимуществом ROC-кривой является её инвариантность относительно отношения цены ошибки I и II рода.

Результаты сравнительного анализа существующих алгоритмов классификации пола человека по изображению лица приведены на рис. 1.15 и в табл. 1.3.

Таблица 1.3

Уровень распознавания существующих алгоритмов классификации пола

Алгоритм	KDDA		АП-МОВ		МОВ	
Параметр	Верно	Неверно	Верно	Неверно	Верно	Неверно
Уровень распознавания мужчин, %	75,8	24,2	90,6	9,4	89,0	11,0
Уровень распознавания женщин, %	65,5	34,5	91,0	9,0	81,3	19,7
Общий уровень распознавания, %	69,7	30,3	90,8	9,2	85,1	14,9

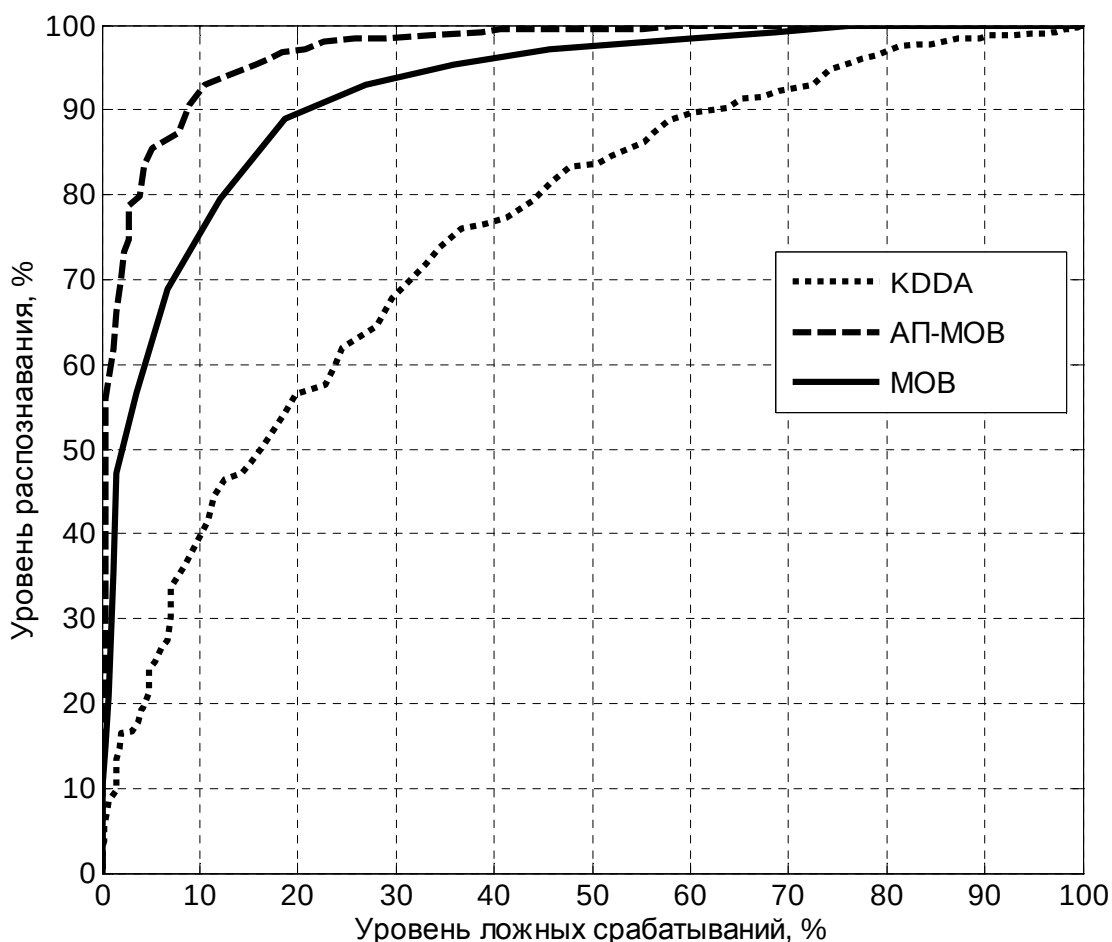


Рис. 1.15. ROC-кривые рассматриваемых алгоритмов классификации пола

Сравнение алгоритмов показывает, что классификаторы, построенные на базе MOB, имеют значительное преимущество по сравнению с классификатором, обученным методом линейного дискриминантного анализа. При этом использование адаптивных признаков дает дополнительное преимущество порядка 5,7%.

Существует небольшой перекося между уровнями распознавания мужчин и женщин, а именно, вероятность верного распознавания мужчин несколько выше, что может быть объяснено большей вариативностью внешности внутри класса женщин, однако, результаты тестирования показывают, что переход в признаковое пространство и оптимизация итогового порога классификатора способны устранить имеющуюся несимметричность. Следует отметить, что неисследованным остается вопрос

устойчивости классификатора на базе МОВ к изменению условий освещенности, а также к наличию искажений и помех на изображениях.

1.5. Краткие выводы

Результаты проведенной работы позволяют сделать следующие основные выводы.

- Рассмотрена система автоматизированного визуального наблюдения, в которой объектом анализа выступает лицо человека. Сформулированы задачи автоматического детектирования лиц и распознавания пола по детектированной области лица на изображении.

- Приведены алгоритмические основы построения классификаторов на базе методов машинного обучения для задач детектирования лиц и распознавания пола.

- Разработана база для тестирования алгоритмов детектирования лиц, состоящая из 50 изображений, суммарно содержащих 213 лиц.

- Разработана собственная база, содержащая более 10 тысяч изображений лиц людей, для проведения обучения и тестирования алгоритмов распознавания пола по детектированной области лица.

- Проведено тестирование трех алгоритмов детектирования лиц на изображениях. Алгоритм на базе РСПЭ показал уровень детектирования 87,3%; алгоритм на базе бустинга – 71,8%; алгоритм на базе МОВ – 62,4%.

- Проведено тестирование трех современных алгоритмов распознавания пола по детектированной области лица. Лучший результат демонстрируют алгоритмы, в основе которых лежит классификатор на базе метода опорных векторов. Уровень верного распознавания алгоритма АП-МОВ составил 90,8%; алгоритма МОВ – 85,1%; алгоритма KDDA – 69,7%.

Таким образом, проведенный в первом разделе анализ позволил сформулировать следующие **основные задачи диссертации**:

- исследование устойчивости алгоритмов детектирования лиц к изменению условий освещенности;
- модификация алгоритмов детектирования лиц путем применения алгоритмов предобработки;
- исследование устойчивости базового классификатора на основе метода опорных векторов к изменению условий освещенности при распознавании пола человека по детектированной области лица;
- модификация гендерного классификатора на основе метода опорных векторов путем применения алгоритмов предварительной обработки изображений, а также путем расширения признакового пространства классификатора.

2. ДЕТЕКТИРОВАНИЕ ЛИЦ НА ТЕЛЕВИЗИОННЫХ ИЗОБРАЖЕНИЯХ ПРИ ИЗМЕНЕНИИ УСЛОВИЙ ОСВЕЩЕННОСТИ

2.1. Вводные замечания

Эффективность алгоритмов детектирования лиц на цифровых изображениях во многом зависит от наличия таких искажающих факторов, как шумы, характерные для функционирования радиотехнических устройств, искажения, возникающие при сжатии видеоданных, а также изменение условий освещенности. Результаты исследований некоторых современных алгоритмов детектирования лиц в условиях наличия гауссовского и импульсного шумов, при размытии, а также сжатии алгоритмом JPEG-2000 приведены в работах [69, 113]. Влияние условий освещенности на работу алгоритмов детектирования и распознавания лиц на сегодняшний день менее изучено, и потому исследование этого вопроса является актуальной задачей.

Данный раздел посвящен исследованию влияния условий освещенности на работу трех выбранных для тестирования алгоритмов детектирования лиц на телевизионных изображениях. Это алгоритм на базе метода опорных векторов (МОВ), алгоритм на базе адаптивного бустинга и алгоритм на базе разреженной сети просеивающих элементов (РСПЭ). Описание алгоритмических особенностей тестируемых алгоритмов приведено в разделе 1.2.

Исследование проводилось по схеме, представленной на рис. 2.1. Для проведения тестирования была составлена база изображений, параметры которой приведены в разделе 1.4.1. Изображения из тестовой базы подвергались преобразованию с тем, чтобы смоделировать следующие условия освещенности: изменение яркости изображения, изменение контраста, неравномерное освещение. Параметры используемых моделей приведены в разделе 2.2.

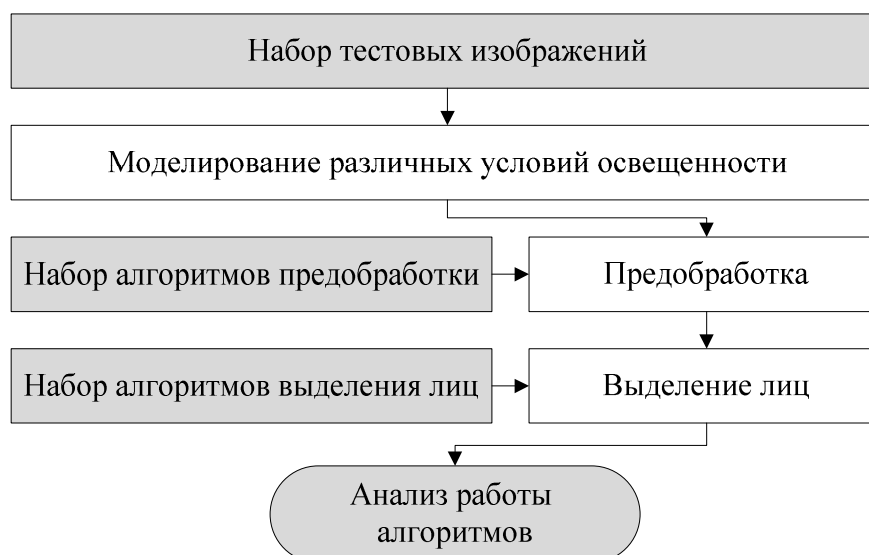


Рис. 2.1. Схема исследования алгоритмов детектирования лиц

На полученных при моделировании условий освещенности изображениях с помощью тестируемых алгоритмов детектирования были выделены лица. Было предложено использовать этап предобработки для того, чтобы повысить устойчивость алгоритмов детектирования лиц к изменению условий освещенности. На этапе предобработки использовались следующие алгоритмы: эквализация гистограммы изображения; алгоритм локальной предобработки. Их описание приведено в разделе 2.4.

Работа тестируемых алгоритмов оценивалась по двум параметрам: уровень верного детектирования, показывающий процент обнаруженных лиц от общего числа лиц в тестовом наборе; и число ложно детектированных объектов, равное общему количеству ошибок детектирования на всем тестовом наборе.

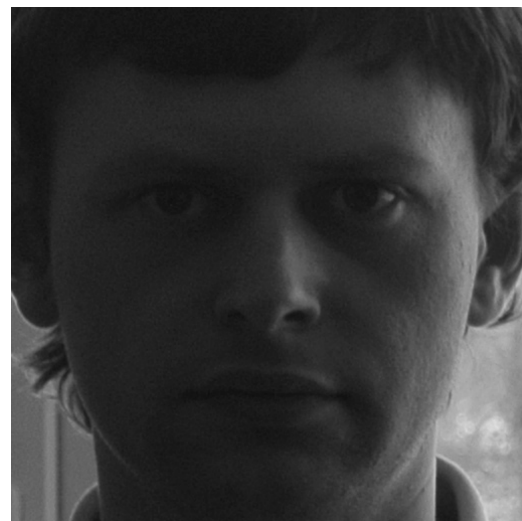
2.2. Модели изменения условий освещенности

Не смотря на то, что на настоящий момент достигнут достаточно высокий уровень верного распознавания в задачах детектирования области лица на изображении (более 80%) и определения пола человека по изображению лица (порядка 90%), в практической ситуации, обусловленной особенностями функционирования систем охранного телевидения, процент

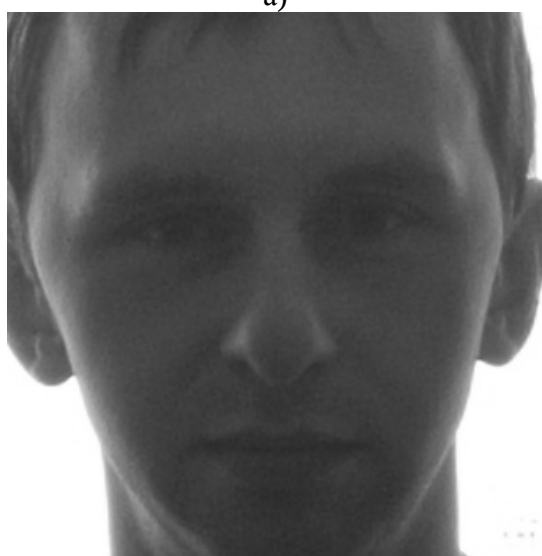
верного распознавания может оказаться значительно ниже. Следует отметить, что сравнительный анализ существующих алгоритмов, результаты которого приведены в разделе 1.4, проводился на изображениях с контролируемым равномерным освещением и высоким контрастом. Изображения с такими же яркостными характеристиками использовались и при обучении классификаторов. Ряд одиночных экспериментов с видеопоследовательностями, полученными с камеры видеонаблюдения в условиях эксплуатации, приближенных к реальным, выявили уязвимость тестируемых алгоритмов к изменению условий освещенности.



а)



б)



в)



г)

Рис. 2.2. Примеры изображений лиц в условиях неконтролируемой освещенности: а) засвечивание; б) затемнение; в) низкий контраст; г) неравномерная освещенность

Значительное падение вероятности верной классификации наблюдалось в следующих случаях (рис. 2.2):

- увеличение средней яркости области лица (засвечивание) вследствие наличия фронтально расположенного источника света;
- уменьшение контраста и средней яркости области лица при ведении видеосъемки против источника света;
- наличие неравномерного освещения левой и правой сторон лица при боковом расположении источника света.

Для проведения исследования алгоритмов детектирования лиц и распознавания пола в различных условиях освещенности требуется создание моделей внесения искажений в оригинальное изображение, соответствующих той или иной реальной ситуации. Эти модели должны позволять изменять степень выраженности различных условий освещенности с некоторым шагом. В данной диссертационной работе предлагаются следующие три модели внесения искажений в исходное изображение, имитирующие вышеописанные ситуации неконтролируемой освещенности:

- модель изменения яркости;
- модель контрастности изображения;
- модель неравномерной освещенности.

Предлагаемые модели носят интуитивный характер, отражая физический смысл рассматриваемого явления. Так, изменение яркости может интерпретироваться, как изменение мощности сигнала, а изменение контраста – как уменьшение динамического диапазона. Некоторые из используемых операций преобразования яркости являются общепринятыми в области цифровой обработки изображений и определены соответствующими стандартами [98].

2.2.1. Модель изменения яркости

Модель изменения яркости описывается выражением:

$$\hat{Y} = Y \cdot \left(1 + \frac{dY}{100} \%\right), \quad (2.1)$$

где Y – яркостная компонента изображения, вычисляемая как взвешенная сумма цветовых компонент исходного пространства RGB:

$$Y = 0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B, \quad (2.2)$$

а dY – внесенное искажение яркости, выраженное в процентах от Y . При проведении тестирования алгоритмов классификации пола величина dY изменялась от -70% до 70% с шагом 10%. Визуальный пример области лица на тестовом изображении, искаженном с различной величиной dY , приведен на рис. 2.3. Рассматриваемое изменение яркости приводит к потере части полезной информации в результате усечения (при достижении максимального значения яркости), а также из-за эффектов квантования (при уменьшении значения яркости). Эта потеря информации может привести к уменьшению вероятности верного распознавания. Изображения, полученные с помощью модели, визуально соответствуют изображениям, полученным с камеры видеонаблюдения в условиях затемнения и засвечивания.

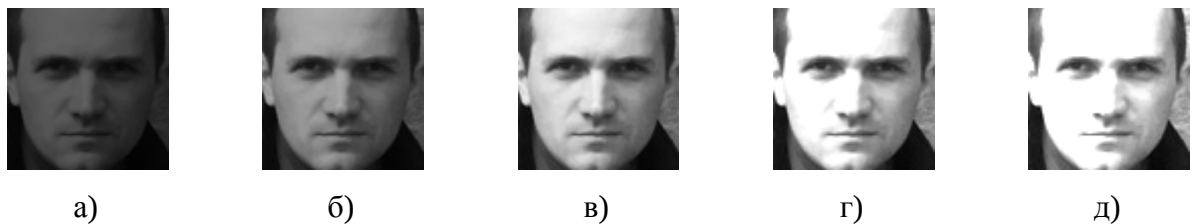


Рис. 2.3. Тестовое изображение лица с различным изменением яркости dY : а) $dY = -60\%$; б) $dY = -30\%$; в) $dY = 0\%$ (неискаженное изображение); г) $dY = 30\%$; д) $dY = 60\%$

2.2.2. Модель изменения контраста

Рассмотрим модель контрастности изображения. Полный контраст исходного изображения задается как разность между максимальным и минимальным значением яркости:

$$K = Y_{max} - Y_{min}. \quad (2.3)$$

Тогда уменьшенное значение контраста может быть представлено следующим выражением:

$$\hat{K} = K \cdot \frac{\alpha}{100} \%, \quad (2.4)$$

где α – коэффициент, выраженный в процентах, показывающий какую долю от исходного контраста составляет контраст искаженного изображения. Для того, чтобы значения яркости искаженного изображения оказались в заданных рамках, применялось следующее преобразование (рис. 2.4):

$$\hat{Y} = \left(Y - Y_{min} - \frac{K}{2} \right) \cdot \frac{\alpha}{100} \% + Y_{min} + \frac{K}{2}. \quad (2.5)$$

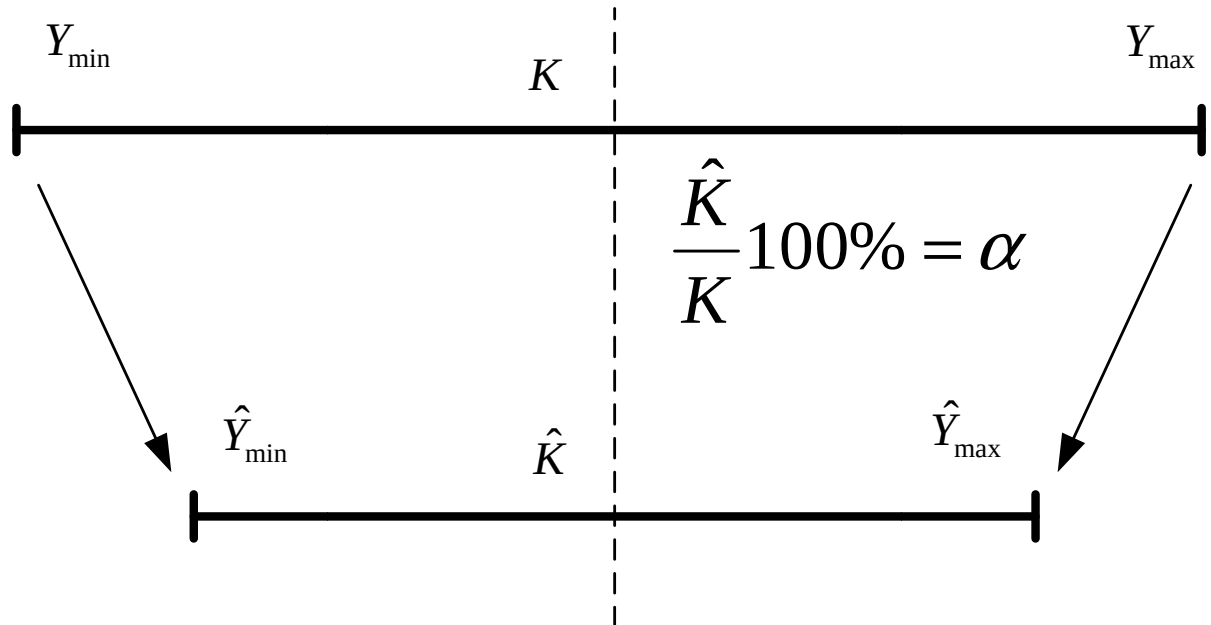


Рис. 2.4. Иллюстрация преобразования контраста изображения

В процессе тестирования использовались значения α от 10% до 100% с шагом 10%. Визуальный пример области лица на тестовом изображении, искаженном с различной величиной α , приведен на рис. 2.5.

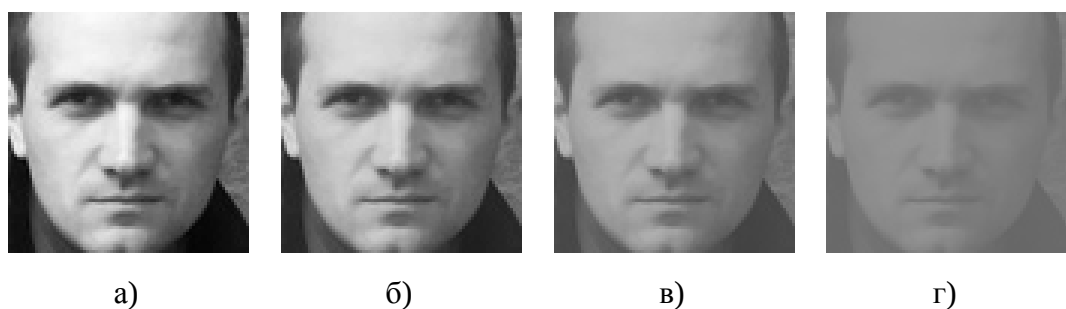


Рис. 2.5. Тестовое изображение лица с различным изменением контраста:

а) $\alpha = 100\%$ (неискаженное изображение); б) $\alpha = 80\%$; в) $\alpha = 50\%$; г) $\alpha = 20\%$

Визуальный пример показывает, что изображения, полученные с помощью модели изменения контраста, соответствуют изображениям, полученным с камеры видеонаблюдения при ведении съемки против источника света, либо на ярком фоне.

2.2.3. Модель неравномерной освещенности

Перейдем к рассмотрению последней модели искажений – модели неравномерной освещенности. Чтобы симитировать боковой источник света применяется взвешенное изменение яркости в зависимости от номера столбца изображения (j):

$$\hat{Y}_j = Y_j \left(1 + \beta \cdot \left(\frac{j}{N} - \frac{1}{2} \right) \right), \quad (2.6)$$

где N – разрешение изображения в количестве пикселей (число столбцов); β – коэффициент, определяющий амплитуду вносимого искажения; j – номер столбца изображения; а $\hat{Y}_j$ – результат преобразования (искаженное изображение). В работе использовалось значение $N = 80$ и β от -2,4 до 2,4 с шагом 0,16. Отрицательные значения β соответствуют имитации источника света, расположенного слева, а положительные – справа (рис. 2.6). Изображения, полученные с помощью модели, визуально соответствуют изображениям, полученным с камеры видеонаблюдения при боковом освещении.

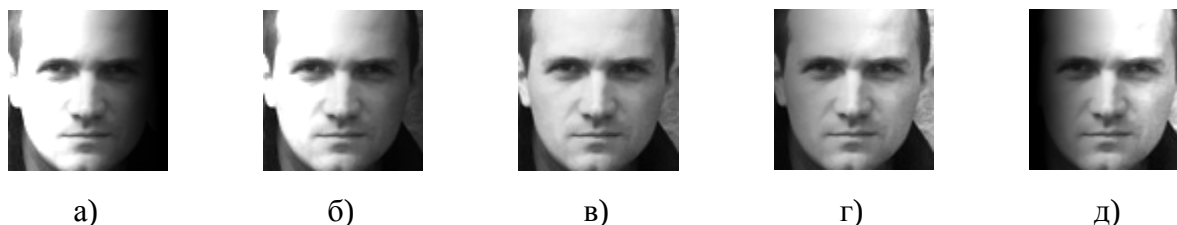


Рис. 2.6. Тестовое изображение лица, искаженное согласно модели неравномерной освещенности с различной β : а) $\beta = -2,4$; б) $\beta = -1,2$; в) $\beta = 0$ (неискаженное изображение); г) $\beta = 1,2$; д) $\beta = 2,4$

2.3. Исследование работы существующих алгоритмов детектирования лиц при изменении условий освещенности

Рассмотрим результаты тестирования алгоритмов детектирования лиц при изменении условий освещенности. Зависимость уровня верного детектирования тестируемых алгоритмов при изменении яркости приведена на рис. 2.7, а число ложно детектированных объектов в табл. 2.1.

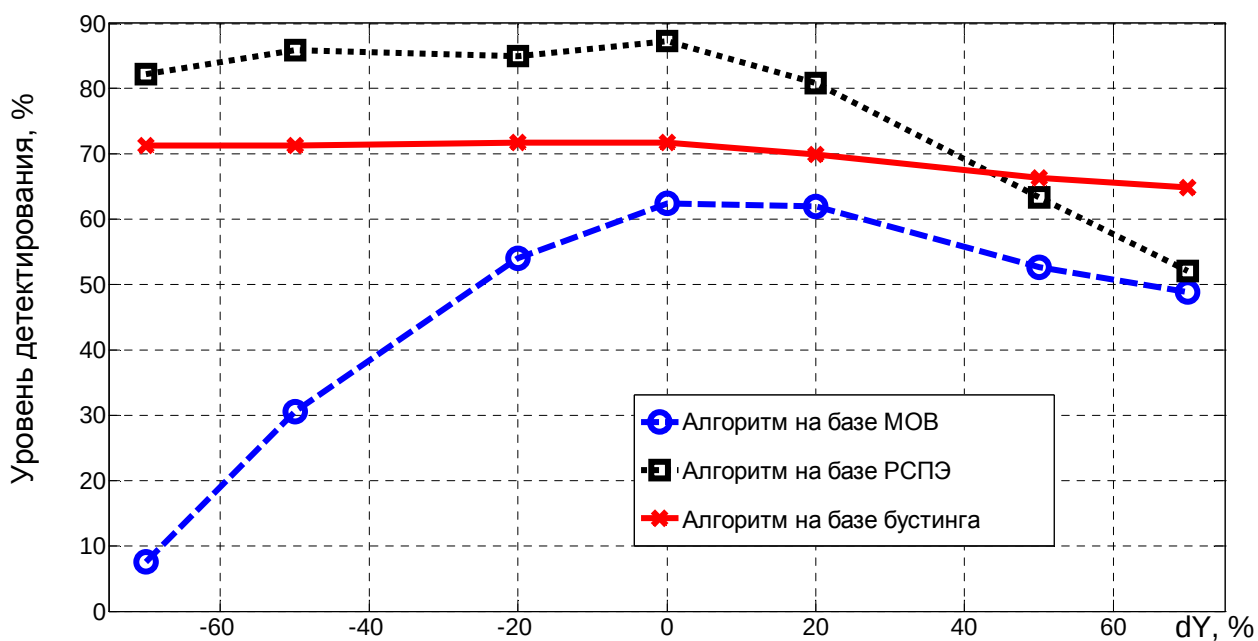


Рис. 2.7. Зависимость уровня верного детектирования тестируемых алгоритмов при изменении яркости изображений

Таблица 2.1.

Число ложно детектированных объектов тестируемыми алгоритмами при изменении яркости изображений

Алгоритм \ Изменение яркости dY , %	Изменение яркости dY , %						
	-70	-50	-20	0	20	50	70
Алгоритм на базе МОВ	8	42	77	73	75	75	86
Алгоритм на базе РСПЭ	15	32	31	34	25	20	17
Алгоритм на базе бустинга	12	12	12	12	11	9	6

Изменение яркости по-разному влияет на результат работы тестируемых алгоритмов. Алгоритм на базе МОВ не справляется с задачей детектирования при уменьшении яркости изображений, показывая уровень верного детектирования менее 10% при уменьшении яркости на 70%. В среднем уровень верного детектирования алгоритма на базе МОВ при изменении яркости составляет 45,4%. Уровень верного детектирования на базе РСПЭ значительно снижается при увеличении яркости и в среднем составляет 76,7%. Алгоритм на базе бустинга демонстрирует устойчивость к изменению яркости (уровень детектирования снижается в среднем на 2,2%).

Анализ числа ложно детектированных объектов показывает, что на тех участках изменения яркости, где наблюдается уменьшение уровня детектирования, количество ложных обнаружений также уменьшается. Алгоритм на базе МОВ детектирует в среднем 62 ложных объекта на всем тестовом наборе, алгоритм на базе РСПЭ – 25, алгоритм на базе бустинга – 10.

На рис. 2.8. и в табл. 2.2 приведены результаты работы тестируемых алгоритмов при изменении контраста изображения. Можно сделать вывод, что алгоритм на базе бустинга и алгоритм на базе РСПЭ эффективно справляются с уменьшением контраста изображения. Их уровень верного детектирования изменяется незначительно и составляет в среднем 71,9% и 86,0% соответственно.

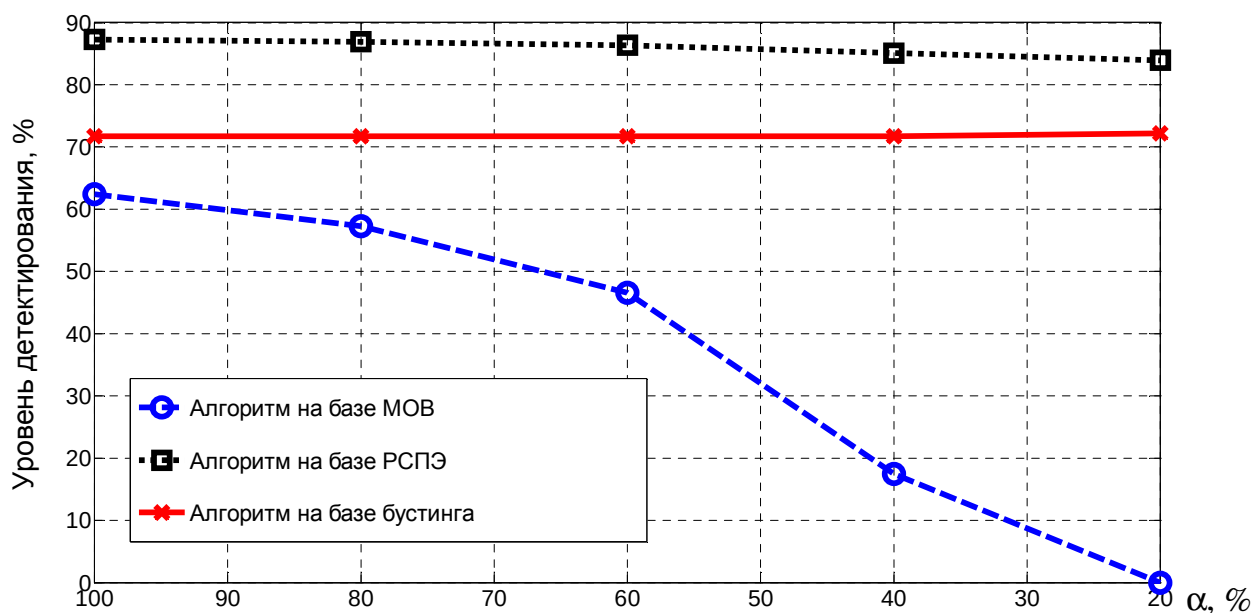


Рис. 2.8. Зависимость уровня верного детектирования тестируемых алгоритмов при изменении контраста изображений

Таблица 2.2.

Число ложно детектированных объектов тестируемыми алгоритмами при изменении контраста изображений

Алгоритм \ Изменение контраста α , %	Изменение контраста α , %				
	100	80	60	40	20
Алгоритм на базе МОВ	73	69	64	28	1
Алгоритм на базе РСПЭ	34	32	26	11	5
Алгоритм на базе бустинга	12	12	12	12	11

Следует отметить, что в условиях низкого контраста изображений алгоритм на базе РСПЭ допускает намного меньше ошибок ложного детектирования, что может быть вызвано тем, что объекты, которые в нормальных условиях освещенности были ложно детектированы, при уменьшении контраста имеют меньшее сходство с чертами лица.

Результаты тестирования показывают, что при уменьшении контраста изображений уровень верного детектирования алгоритма на базе МОВ

значительно снижается (практически до нуля при контрасте изображений в 20% от исходного) и составляет в среднем 36,7%. Та же зависимость наблюдается и для числа ложных обнаружений.

Рассмотрим результаты работы тестируемых алгоритмов при неравномерной освещенности (рис. 2.9 и табл. 2.3).

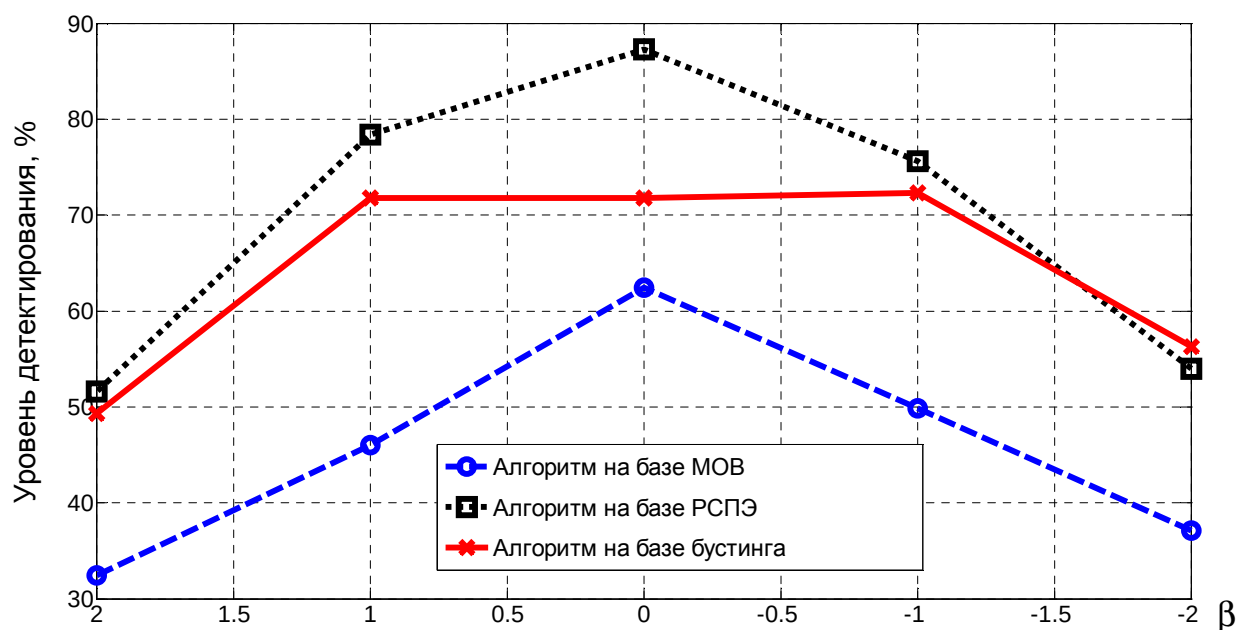


Рис. 2.9. Зависимость уровня верного детектирования тестируемых алгоритмов в условиях неравномерной освещенности изображений

Таблица 2.3.

Число ложно детектированных объектов тестируемыми алгоритмами в условиях неравномерной освещенности изображений

Алгоритм \ Амплитуда неравномерной освещенности β					
	-2	-1	0	1	2
Алгоритм на базе МОВ	47	56	73	56	55
Алгоритм на базе РСПЭ	7	20	34	25	20
Алгоритм на базе бустинга	2	10	12	6	1

Наибольшую устойчивость к неравномерной освещенности демонстрирует алгоритм на базе бустинга. При небольшой амплитуде вносимых искажений его уровень верного детектирования остается неизменным и составляет в среднем 64,3%. Уровень верного детектирования алгоритма на базе МОВ и алгоритма на базе РСПЭ в условиях неравномерной освещенности снижается до среднего уровня в 45,5% и 69,4% соответственно. Также наблюдается снижение и числа ложно детектированных объектов для всех рассматриваемых алгоритмов.

Таким образом, проведенное исследование показало, что в ряде случаев изменение условий освещенности негативно сказывается на работе алгоритмов детектирования лиц на телевизионных изображениях, и существует необходимость повышения устойчивости рассмотренных алгоритмов к этим условиям. Для решения этой задачи предлагается модифицировать этап предобработки изображений.

2.4. Модификация этапа предобработки изображений при детектировании лиц

Для того чтобы улучшить яркостные характеристики изображений и таким образом скомпенсировать влияние условий освещенности на работу тестируемых алгоритмов детектирования лиц предлагается использовать два алгоритма предобработки: алгоритм эквализации гистограммы изображения и алгоритм локальной предобработки.

Эквализация гистограммы является общепринятой процедурой [50], часто используемой в задачах повышения качества изображений. Алгоритм локальной предобработки предлагается в данной диссертационной работе. Он основан на поблочном выравнивании яркости и контраста на основе статистических характеристик отдельных участков изображения. Его применение к задаче распознавания пола по изображению лица рассмотрено в разделе 3.3.2.

2.4.1. Эквиализация гистограммы изображения

Эквиализация гистограммы является общепринятым и наиболее часто используемым инструментом для приведения яркостных характеристик изображения к стандартному виду. Этот алгоритм осуществляет выравнивание количества пикселей на изображении каждой конкретной интенсивности. Формальная запись преобразования интенсивностей пикселей при эквиализации гистограммы приведена в разделе 3.3.1.

На рис. 2.10 представлен визуальный пример изображения из тестовой базы в различных условиях освещенности до и после его преобразования алгоритмом эквиализации гистограммы.



а)



б)



в)



г)

Рис. 2.7. Пример изображения из тестовой базы в различных условиях освещенности (слева) и его преобразования алгоритмом эквализации гистограммы (справа): а) исходное изображение; б) изменение яркости $dY = 70\%$; в) изменение контраста $\alpha = 20\%$; г) неравномерная освещенность амплитудой $\beta = 1$

Визуальный пример показывает, что эквализация гистограммы увеличивает контраст изображения, делая черты лица более яркими по сравнению с окружающим фоном.

2.4.2. Локальная предобработка

В связи с тем, что яркостные характеристики изображения на его различных областях могут существенно отличаться, предлагается обрабатывать изображение поблочно, выравнивая яркостные характеристики для каждого конкретного блока. Такой подход позволяет контролировать среднюю яркость и контраст не только для всего изображения, но и для каждого его отдельного участка.

В работе изображение разбивалось на блоки разрешением 48×48 пикселей с перекрытием в 16 пикселей между двумя соседними блоками. На основе статистики каждого блока осуществлялось преобразование яркости и контраста. Для преобразования яркости блока использовалось выражение:

$$Y_k = Y - \langle Y_k \rangle + 0,5, \quad (2.1)$$

где Y – интенсивность пикселей исходного изображения, представленная в формате double (по шкале от 0 до 1); $\langle Y_k \rangle$ – среднее значение интенсивности пикселей исходного изображения, входящих в блок k . Для преобразования контраста значения всех пикселей внутри блока сортируются по возрастанию. Определенный небольшой процент самых ярких и самых темных пикселей отбрасывается, после чего яркость оставшихся пикселей рассчитывается по формуле:

$$Y_k = \frac{Y - 0,5}{K} + 0,5, \quad (2.2)$$

где $K = \max\{(1 - 2 \cdot Y_{\min}), (2 \cdot Y_{\max} - 1)\}$ – коэффициент, показывающий контраст входного изображения и определяющий степень увеличения контраста обработанного изображения.

В результате разбиения изображения на блоки для каждого из полученных блоков образуется 8 зон перекрытия разрешением 16×16 пикселей. Четыре зоны – перекрытие между двумя соседними блоками, еще четыре – перекрытие на стыке четырех прилегающих друг к другу блоков. Для получения итогового значения интенсивности пикселей, находящихся в зонах перекрытия, производится взвешенное суммирование значений интенсивности пикселей перекрывающихся блоков в зависимости от их отдаления от центра каждого блока. Такой подход позволяет избежать появления артефактов блочности при обработке изображения.

На рис. 2.11 представлен визуальный пример изображения из тестовой базы в различных условиях освещенности до и после его преобразования алгоритмом локальной предобработки.



а)



б)



в)



г)

Рис. 2.8. Пример изображения из тестовой базы в различных условиях освещенности (слева) и его преобразования алгоритмом локальной предобработки (справа): а) исходное изображение; б) изменение яркости $dY = 70\%$; в) изменение контраста $\alpha = 20\%$; г) неравномерная освещенность амплитудой $\beta = 1$

Визуальный пример показывает, что в результате обработки изображения средняя яркость и контраст его отдельных участков выравниваются.

2.5. Результаты тестирования алгоритмов детектирования лиц с предложенными модификациями

Рассмотрим результаты тестирования алгоритмов детектирования лиц с предложенными модификациями на этапе предобработки. На рис. 2.12 и в табл. 2.4 приведены результаты тестирования алгоритма на базе МОВ при изменении яркости изображений.

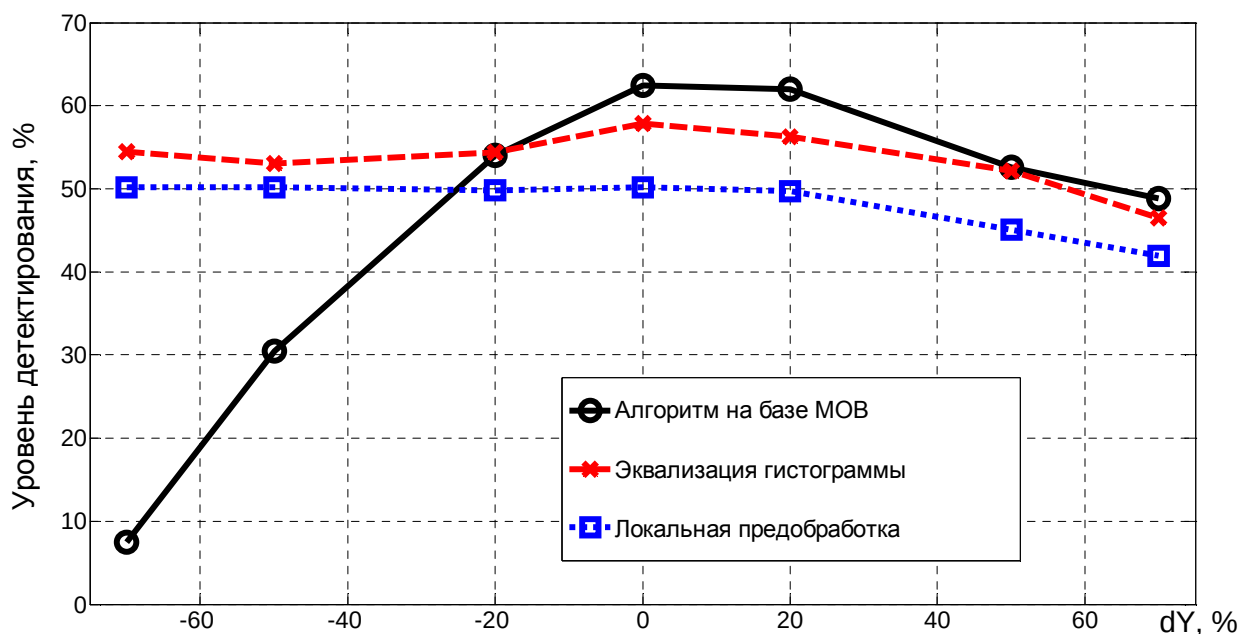


Рис. 2.12. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе МОВ при изменении яркости изображений

Эквализация гистограммы и локальная предобработка позволяют повысить уровень верного детектирования алгоритма на базе МОВ при уменьшении яркости изображений. Выигрыш в уровне детектировании при уменьшении яркости на 70% составляет более 40%. Средний уровень детектирования при использовании эквализации гистограммы составляет

53,5% (выигрыш в 8,1%), при использовании локальной предобработки – 48,2% (выигрыш в 2,8%).

Таблица 2.4.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе МОВ при изменении яркости изображений

Изменение яркости dY , %	-70	-50	-20	0	20	50	70
Алгоритм							
Алгоритм на базе МОВ	8	42	77	73	75	75	86
Эквализация гистограммы	92	93	95	87	89	84	69
Локальная предобработка	116	98	92	99	86	112	89

Число ложно детектированных объектов при эквализации гистограммы и при локальной предобработке увеличивается и составляет в среднем 87 и 99 объектов соответственно.

На рис. 2.13 и в табл. 2.5 приведены результаты тестирования алгоритма на базе МОВ при изменении контраста изображений.

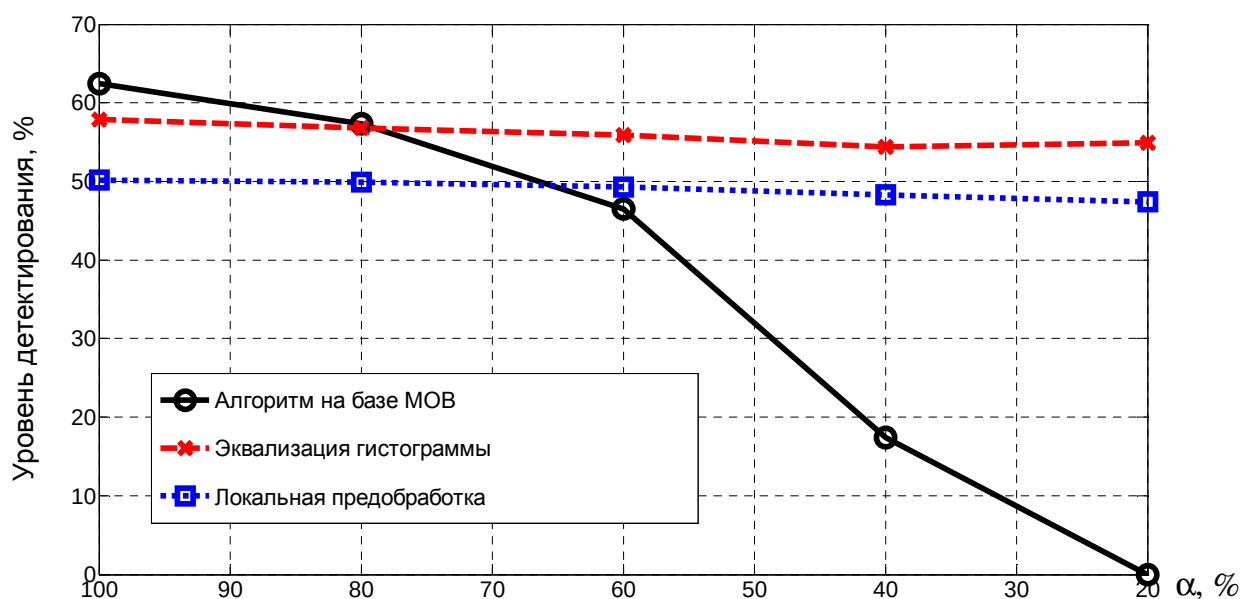


Рис. 2.13. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе МОВ при изменении контраста изображений

Таблица 2.5.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе МОВ при изменении контраста изображений

Изменение контраста α , %	100	80	60	40	20
Алгоритм					
Алгоритм на базе МОВ	73	69	64	28	1
Эквализация гистограммы	87	85	89	92	98
Локальная предобработка	99	92	88	92	93

Удается значительно повысить средний уровень детектирования алгоритма на базе МОВ в условиях изменения контраста. При использовании эквализации гистограммы этот уровень составляет 56%, что на 19,3% больше по сравнению с результатами без использования предобработки. Средний уровень детектирования при локальной предобработке составляет 49,0% (выигрыш в 12,3%). Наблюдается небольшой рост числа ложно детектированных объектов – порядка 92 объектов в среднем при использовании обоих алгоритмов предобработки.

На рис. 2.14 и в табл. 2.6 приведены результаты тестирования алгоритма на базе МОВ в условиях неравномерной освещенности.

Таблица 2.6.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе МОВ в условиях неравномерной освещенности

Амплитуда неравномерной освещенности β	-2	-1	0	1	2
Алгоритм					
Алгоритм на базе МОВ	47	56	73	56	55
Эквализация гистограммы	41	62	87	72	44
Локальная предобработка	62	76	99	87	63

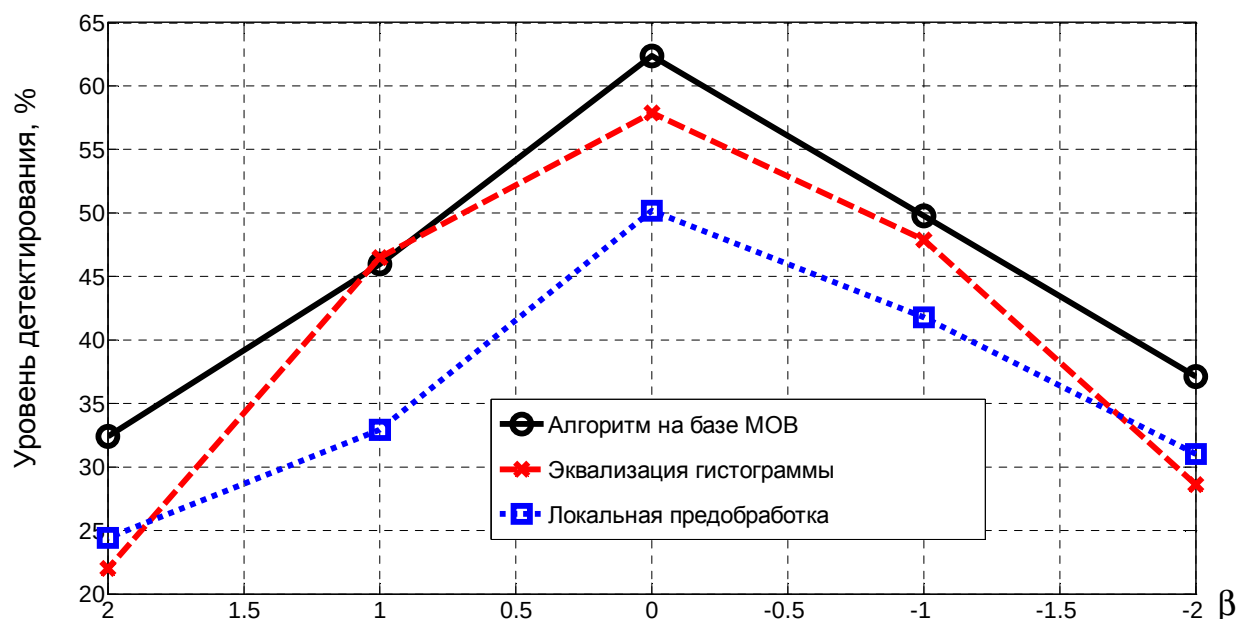


Рис. 2.14. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе МОВ в условиях неравномерной освещенности изображений

Результаты показывают, что в условиях неравномерной освещенности уровень верного детектирования алгоритма на базе МОВ снижается при использовании предобработки. Для эквализации гистограммы он составляет в среднем 40,6%, для локальной предобработки – 36,0%. Число ложных детектирований незначительно возрастает до среднего уровня в 61 и 77 объектов при использовании эквализации гистограммы и локальной предобработки соответственно.

Перейдем к рассмотрению результатов тестирования алгоритма на базе РСПЭ, чьи показатели в условиях изменения яркости приведены на рис. 2.15 и в табл. 2.7. Применение алгоритма локальной предобработки позволяет добиться улучшения уровня детектирования в условиях увеличения яркости изображений. В среднем при изменении яркости уровень верного детектирования алгоритма на базе РСПЭ увеличивается на 2% и составляет 78,7%. Эквализация гистограммы приводит к ухудшению результатов со средним уровнем детектирования 74,1%.

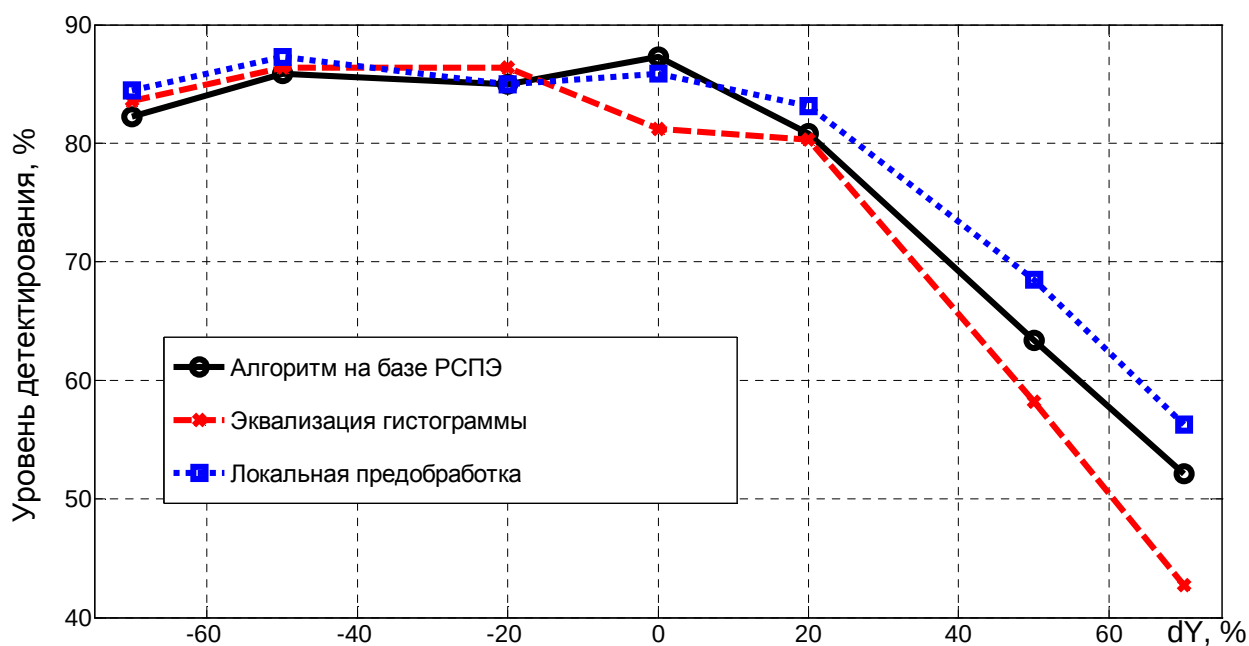


Рис. 2.15. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе РСПЭ при изменении яркости изображений

Таблица 2.7.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе РСПЭ при изменении яркости изображений

Алгоритм \ Изменение яркости dY , %							
	-70	-50	-20	0	20	50	70
Алгоритм на базе РСПЭ	15	32	31	34	25	20	17
Эквализация гистограммы	22	37	36	14	37	30	25
Локальная предобработка	17	29	43	54	44	27	25

Число ложных срабатываний при применении эквализации гистограммы составляет в среднем 29 объектов, при применении локальной предобработки – 34 объекта.

На рис. 2.16 и в табл. 2.8 приведены результаты тестирования алгоритма на базе РСПЭ при изменении контраста изображений. Поскольку уровень верного детектирования алгоритма на базе РСПЭ изменяется незначительно при изменении контраста изображений, не удастся достичь его улучшения в результате использования алгоритмов предобработки.

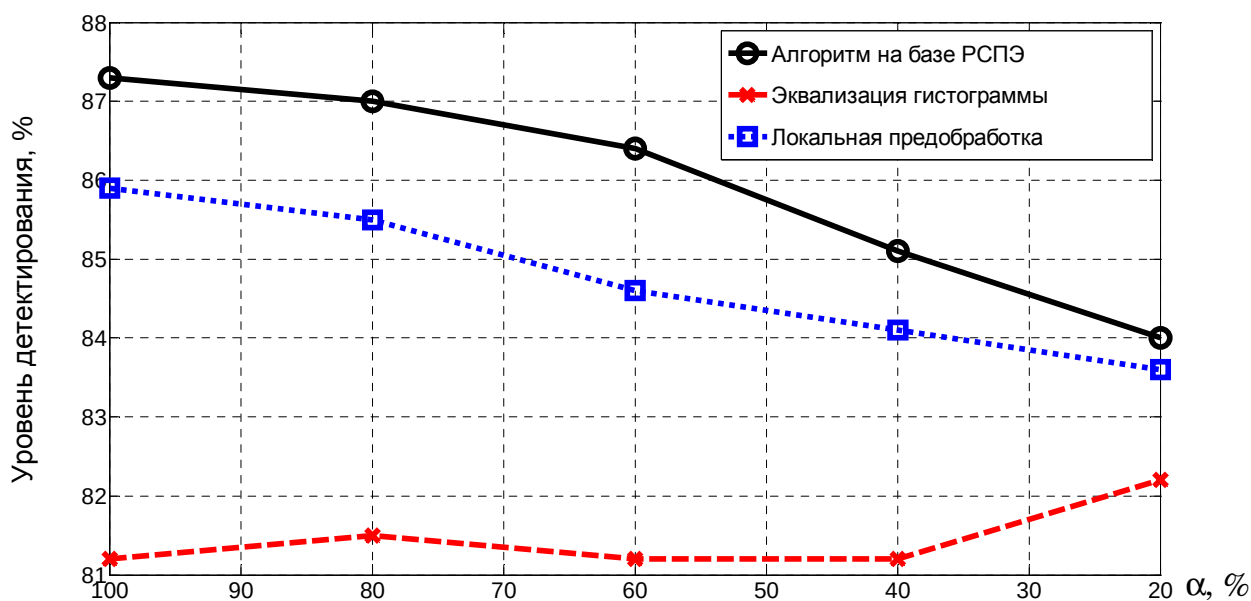


Рис. 2.16. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе РСПЭ при изменении контраста изображений

Таблица 2.8.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе РСПЭ при изменении контраста изображений

Алгоритм	Изменение контраста α , %				
	100	80	60	40	20
Алгоритм на базе РСПЭ	34	32	26	11	5
Эквализация гистограммы	14	12	11	10	9
Локальная предобработка	54	41	31	25	16

Средний уровень детектирования при эквализации гистограммы составил 81,5%, при локальной предобработке – 84,7%. Число ложно детектированных объектов алгоритма на базе РСПЭ составляет в среднем 11 и 33 объекта при использовании эквализации гистограммы и локальной предобработки соответственно.

Рассмотрим результаты тестирования алгоритма на базе РСПЭ в условиях неравномерной освещенности, приведенные на рис. 2.17 и в табл. 2.9.

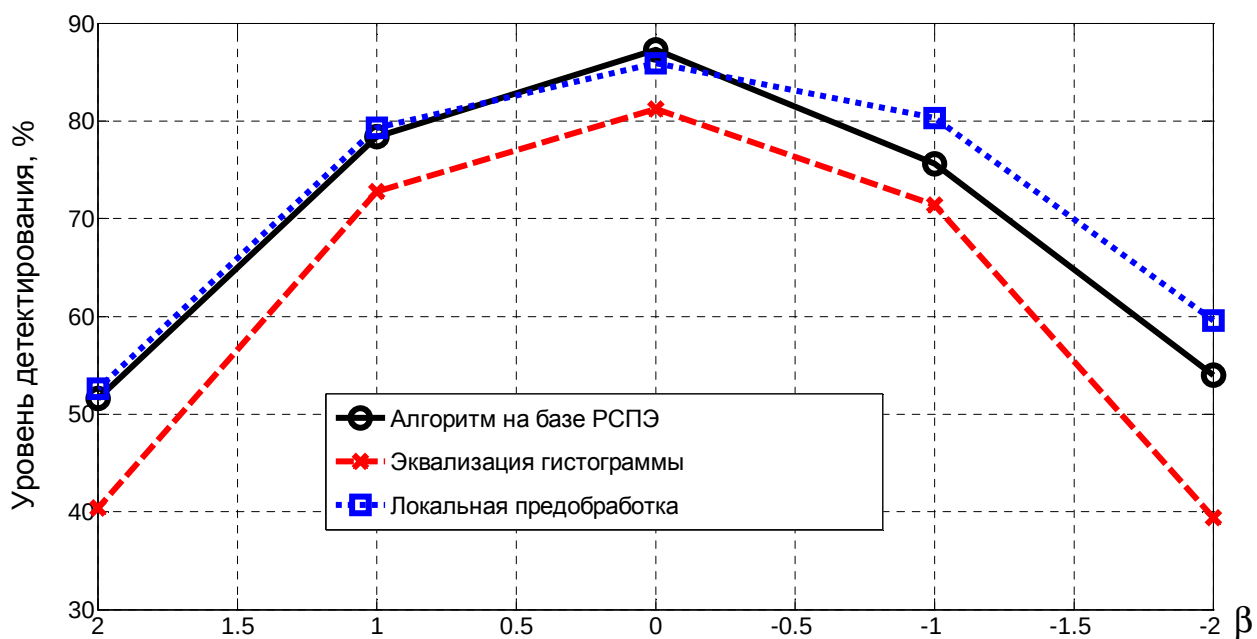


Рис. 2.17. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе РСПЭ в условиях неравномерной освещенности изображений

Таблица 2.9.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе РСПЭ в условиях неравномерной освещенности

Алгоритм \ Амплитуда неравномерной освещенности β					
	-2	-1	0	1	2
Алгоритм на базе РСПЭ	7	20	34	25	20
Эквализация гистограммы	2	12	14	19	4
Локальная предобработка	15	34	54	36	28

Локальная предобработка позволяет увеличить на 2,1% средний уровень верного детектирования алгоритма на базе РСПЭ в условиях неравномерной освещенности. Эквализация гистограммы, напротив, приводит к ухудшению результата в среднем на 8,4%. Число ложных обнаружений при использовании локальной предобработки увеличивается и составляет в среднем 33 объекта. Тот же показатель для эквализации гистограммы составляет 10 объектов.

Рассмотрим результаты тестирования последнего из рассмотренных алгоритмов детектирования лиц на цифровых изображениях – алгоритма на базе бустинга. На рис. 2.18 и в табл. 2.10 приведены его показатели при изменении яркости изображений.

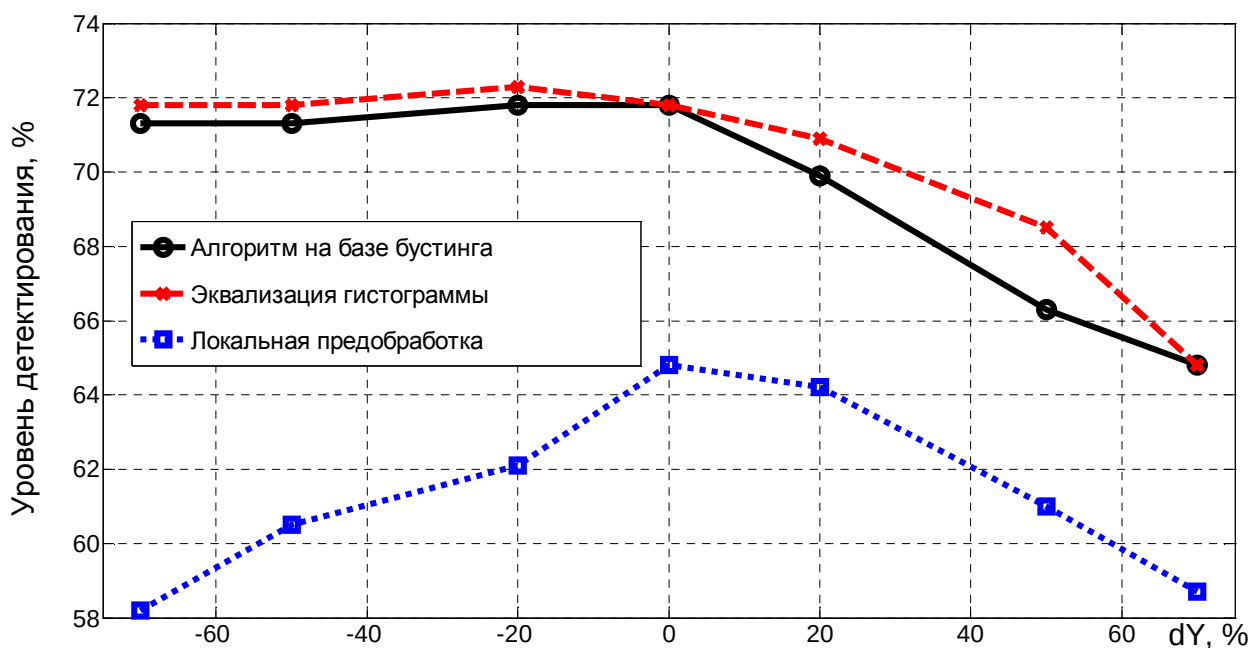


Рис. 2.18. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе бустинга при изменении яркости изображений

Таблица 2.10.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе бустинга при изменении яркости изображений

Алгоритм \ Изменение яркости dY , %							
	-70	-50	-20	0	20	50	70
Алгоритм на базе бустинга	12	12	12	12	11	9	6
Эквиализация гистограммы	10	11	12	12	10	7	6
Локальная предобработка	30	17	14	9	10	13	14

Использование эквиализации гистограммы дает преимущество по уровню верного детектирования в среднем порядка 0,7%. Число ложно детектированных объектов при этом практически не изменяется. Локальная

предобработка вызывает ухудшение работы алгоритма на базе бустинга – уровень детектирования уменьшается в среднем на 8,2%, число ложно детектированных объектов возрастает в среднем до 15.

На рис. 2.19 и в табл. 2.11 приведены результаты тестирования алгоритма на базе бустинга при изменении контраста изображений.

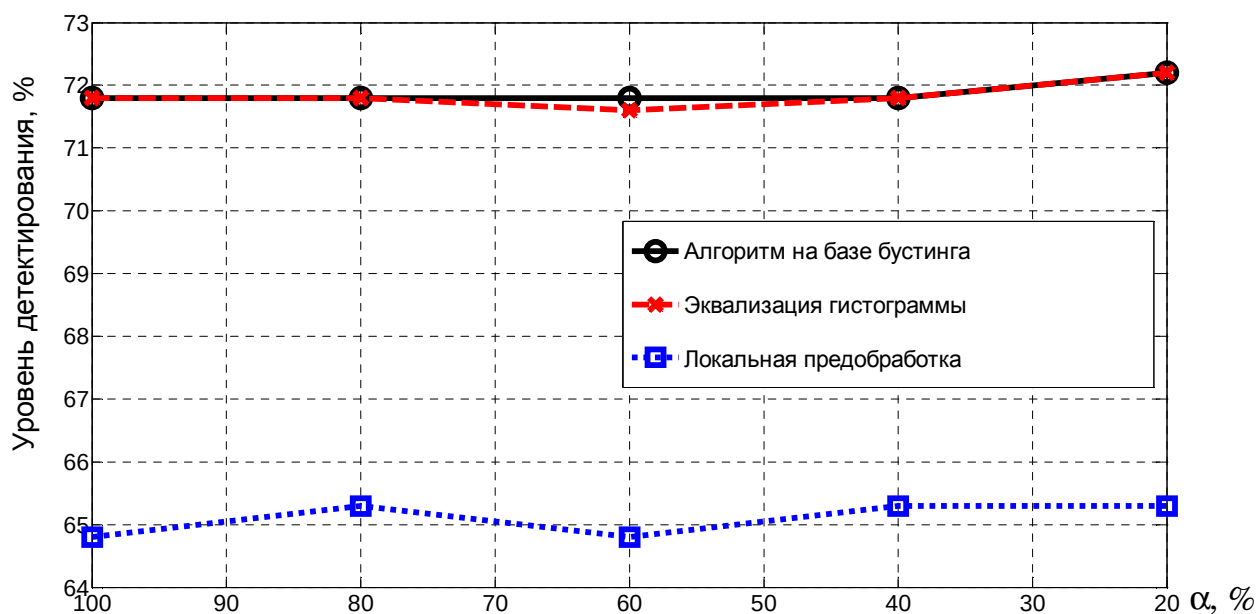


Рис. 2.19. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе бустинга при изменении контраста изображений

Таблица 2.11.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе бустинга при изменении контраста изображений

Алгоритм \ Изменение контраста α , %	Изменение контраста α , %				
	100	80	60	40	20
Алгоритм на базе бустинга	12	12	12	12	11
Эквализация гистограммы	12	12	10	12	11
Локальная предобработка	9	10	9	10	10

Идентичные результаты достигаются без использования предобработки и с использованием эквализации гистограммы. Число ложных детектирований остается неизменным. Это объясняется чрезвычайной устойчивостью алгоритма на базе бустинга к изменению контраста изображений. Локальная предобработка приводит к уменьшению уровня верного детектирования при изменении контраста в среднем на 6,8%.

На рис. 2.20 и в табл. 2.12 приведены результаты тестирования алгоритма на базе бустинга в условиях неравномерной освещенности.

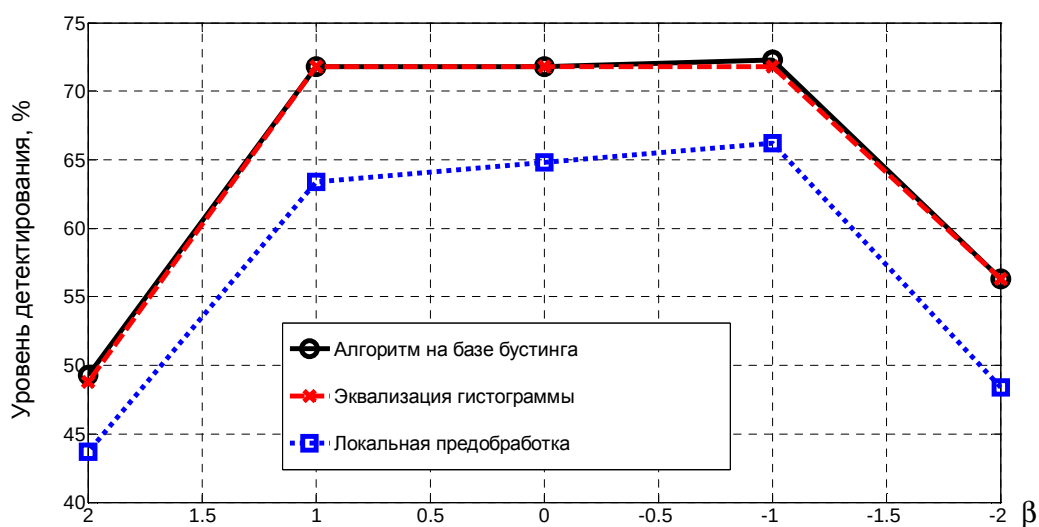


Рис. 2.20. Влияние алгоритмов предобработки на детектирование лиц алгоритмом на базе бустинга в условиях неравномерной освещенности изображений

Таблица 2.12.

Влияние алгоритмов предобработки на число ложно детектированных объектов алгоритмом на базе бустинга в условиях неравномерной освещенности

Алгоритм \ Амплитуда неравномерной освещенности β					
	-2	-1	0	1	2
Алгоритм на базе бустинга	2	10	12	6	1
Эквализация гистограммы	1	12	12	6	2
Локальная предобработка	3	9	9	12	5

Также как и в случае изменения контраста не удается достичь улучшения уровня верного детектирования алгоритма на базе бустинга посредством применения алгоритмов предобработки. Результат не изменяется при использовании эквализации гистограммы и уменьшается в среднем на 7% после локальной предобработки.

2.6 Краткие выводы

Результаты проведенной работы позволяют сделать следующие основные выводы.

- Исследована устойчивость трех современных алгоритмов детектирования лиц на телевизионных изображениях к изменению условий освещенности. При изменении яркости уровень верного детектирования алгоритмов на базе МОВ, РСПЭ и бустинга уменьшается в среднем на 17,0%, 10,6% и 2,2% соответственно. При изменении контраста уровень верного детектирования алгоритма на базе МОВ уменьшается в среднем на 25,7%. Неравномерная освещенность приводит к уменьшению средних показателей алгоритма на базе МОВ на 16,9%; алгоритма на базе РСПЭ – на 17,9%; алгоритма на базе бустинга – на 7,5%.

- Исследована возможность использования алгоритма эквализации гистограммы и алгоритма локальной предобработки для повышения устойчивости тестируемых алгоритмов к изменению условий освещенности.

- Результаты показывают целесообразность использования эквализации гистограммы для улучшения работы алгоритма на базе МОВ. Выигрыш по среднему уровню верного детектирования составляет в этом случае 8,1% при изменении яркости изображений и 19,3% при изменении контраста.

- Для алгоритма на базе РСПЭ лучшие результаты достигаются при использовании локальной предобработки. Выигрыш по среднему уровню верного детектирования составляет в этом случае 2% при изменении яркости изображений и 2,1% в условиях неравномерной освещенности.

– Алгоритм на базе бустинга наиболее устойчив к изменению условий освещенности. Удастся улучшить его уровень верного детектирования в среднем на 0,7% при изменении яркости изображений за счет использования алгоритма эквализации гистограммы.

3. РАСПОЗНАВАНИЕ ПОЛА ЛЮДЕЙ ПО ИЗОБРАЖЕНИЮ ЛИЦА ПРИ ИЗМЕНЕНИИ УСЛОВИЙ ОСВЕЩЕННОСТИ

3.1. Вводные замечания

Данный раздел посвящен разработке системы, различающей людей на изображениях по гендерному признаку (разбиение на два класса: «мужчины» и «женщины»). Подобные системы находят широкое применение в различных областях: робототехнике, интерфейсах взаимодействия человек-компьютер, системах сбора и оценки демографических показателей, видеонаблюдении, системах интеллектуального представления информации с цифровых носителей, и многих других [19-25]. К примеру, распознавание пола может использоваться как первый этап при решении задачи идентификации личности. Методы машинного обучения [6-10], применяемые при построении гендерных классификаторов, носят универсальный характер, что позволяет использовать разработки в области распознавания пола в системах понимания изображений [11] для решения задач классификации многих других объектов [1-5].

Существуют всеобъемлющие сравнительные исследования методов распознавания лиц на изображениях [14], методов обнаружения лиц [19, 22], а также методов распознавания выражения лица [20]. Задача распознавания пола на сегодняшний день менее изучена. Сравнительный анализ гендерных классификаторов, предложенных за последнее время, приведен в работе [65]. Существуют подходы, основанные на нейронных сетях [91], на комбинации вейвлетов Габора и метода главных компонент [92], на анализе независимых компонент и линейном дискриминантном анализе [93, 94]. В работе [95] используется генетический алгоритм для генерации признаков и классификатор на базе метода опорных векторов (МОВ). В работе [96] были предложены локальные бинарные признаки в комбинации с обучающей процедурой AdaBoost [67]. Алгоритм на основе сетей радиальных базисных функций и индуктивных решающих деревьев описан в работе [97].

Для организации полностью автоматической системы распознавание пола используется совместно с алгоритмом выделения лиц, который отбирает кандидатов для анализа [66]. От качества работы алгоритма выделения лиц во многом зависит итоговый результат работы всей системы, поскольку неточности при определении местоположения лица на изображении могут привести к принятию ошибочного решения на этапе распознавания. Для решения задачи выделения лиц на изображении мною использовался алгоритм на базе бустинга, описанный в работе [66]. Визуальный пример работы автоматической системы распознавания пола приведен на рис. 3.1.



Рис. 3.1. Визуальный пример результата работы автоматической системы распознавания пола (класс «мужчины» условно обозначен синим цветом, класс «женщины» – красным)

Результаты сравнительного анализа современных методов обучения классификатора применительно к задаче распознавания пола, приведенные в первой главе, обуславливают выбор метода опорных векторов в качестве базового. Работа алгоритма на базе МОВ рассматривается в контексте его функционирования в радиотехническом телевизионном устройстве, где наличие внешних факторов, таких как изменение условий освещенности, может оказать значительное влияние на точность классификации. И, поскольку вопрос устойчивости алгоритма гендерной классификации на базе

МОВ к различным условиям яркости, контраста, неравномерному освещению ранее не исследовался, этой задаче уделяется внимание в данной работе. Предлагаются три модификации алгоритма на базе МОВ, позволяющие достичь преимущества в уровне верного распознавания пола по сравнению с базовым алгоритмом в рассмотренных условиях. Эти модификации основаны на применении различных алгоритмов предобработки изображений. Первая базируется на общепринятой процедуре эквализации гистограммы, другие две – на алгоритмах предобработки, предложенных автором данной работы.

3.2. Исследование работы алгоритма на базе метода опорных векторов при изменении условий освещенности

Рассмотрим результаты исследования алгоритма на базе МОВ на устойчивость к изменению условий освещенности. Для тестирования использовалась выборка из собственной базы лиц, содержащая 500 изображений мужчин и женщин – по 250 на каждый из классов. Изображения из нее были недоступны в процессе обучения, что гарантирует корректность тестирования. Изображения из тестовой выборки искажались согласно описанным выше моделям для того, чтобы симитировать различные условия освещенности. При этом искажения яркости и контраста комбинировались, применяясь последовательно к одному и тому же изображению при всех возможных dY и α . В процессе тестирования вычислялся процент верной классификации (D). В общем случае при равной цене ошибки для обоих классов процент верной классификации показывает среднее значение между процентом верно классифицированных образцов первого и второго класса. Однако, в связи с равным количеством тестовых изображений, используемых на каждый класс, D вычислялся, как отношение числа верно классифицированных образцов обоих классов к общему числу тестируемых образцов. Полученная в ходе тестирования зависимость приведена на рис. 3.2.

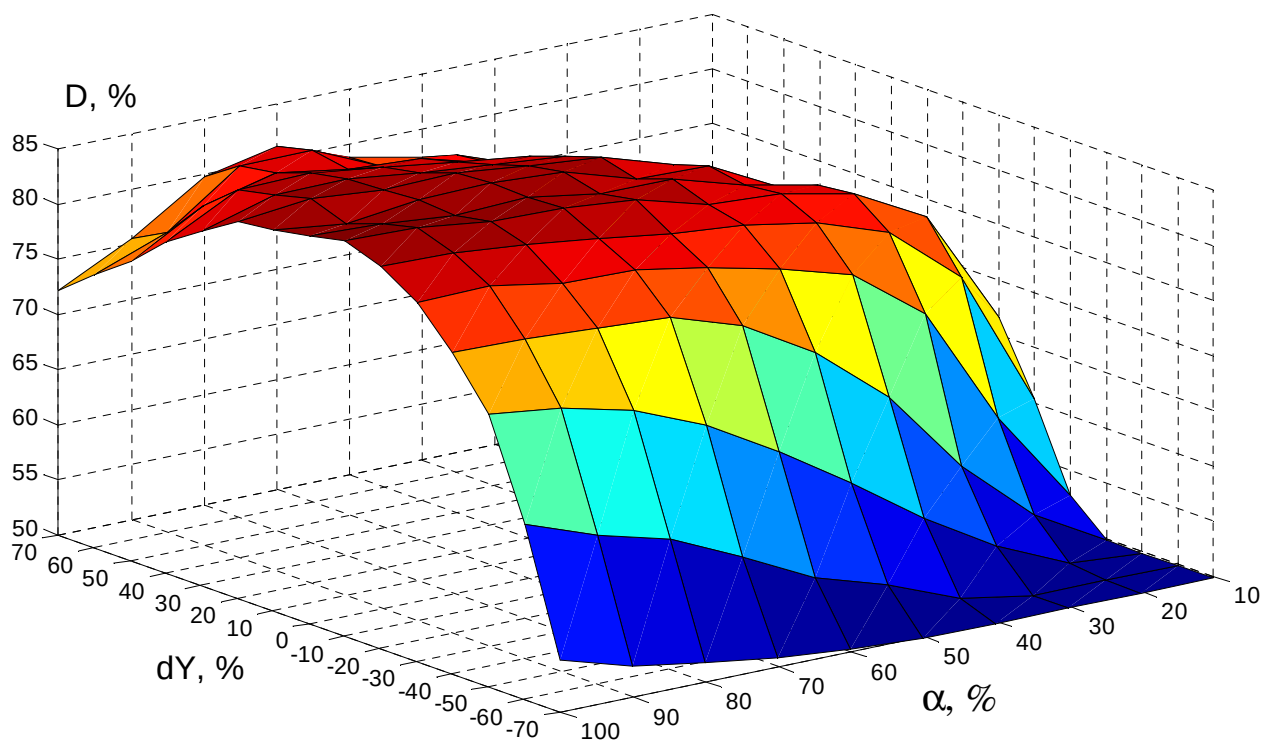


Рис. 3.2. Зависимость процента верной классификации алгоритма на базе МОВ от изменения яркости и контраста

Результаты показывают, что алгоритм на базе МОВ не обладает устойчивостью к изменению яркости и контраста. Уровень верного распознавания пола быстро падает при увеличении амплитуды искажений, а при комбинации максимального искажения по яркости и контрасту достигает уровня в 50%, что свидетельствует о практически случайном угадывании и полном отсутствии какой бы то ни было информации на выходе классификатора.

Рассмотрим более детально приведенную на рис. 3.2 поверхность в различных ее срезках. На рис. 3.3 приведена зависимость процента верной классификации от изменения яркости dY . Наибольшая потеря в вероятности верной классификации возникает при уменьшении яркости изображения. D снижается с 85% до 55% при затемнении изображения на 70%. Увеличение яркости также приводит к уменьшению эффективности классификатора, однако, с меньшей динамикой (с 85% до 72%).

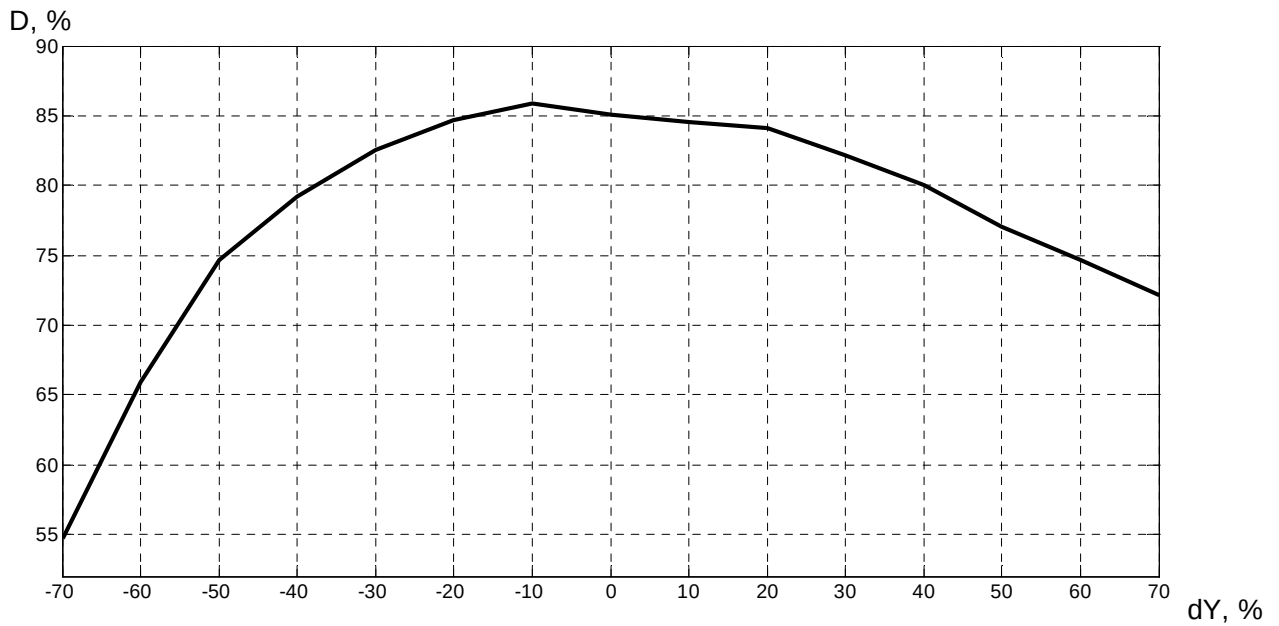


Рис. 3.3. Зависимость процента верной классификации алгоритма на базе МОВ от изменения яркости

Можно сделать вывод, что алгоритм на базе МОВ не обладает устойчивостью к изменению яркости изображения.

Рассмотрим ситуацию, когда средняя яркость изображения остается неизменной, но имеет место ухудшение контраста (рис. 3.4).

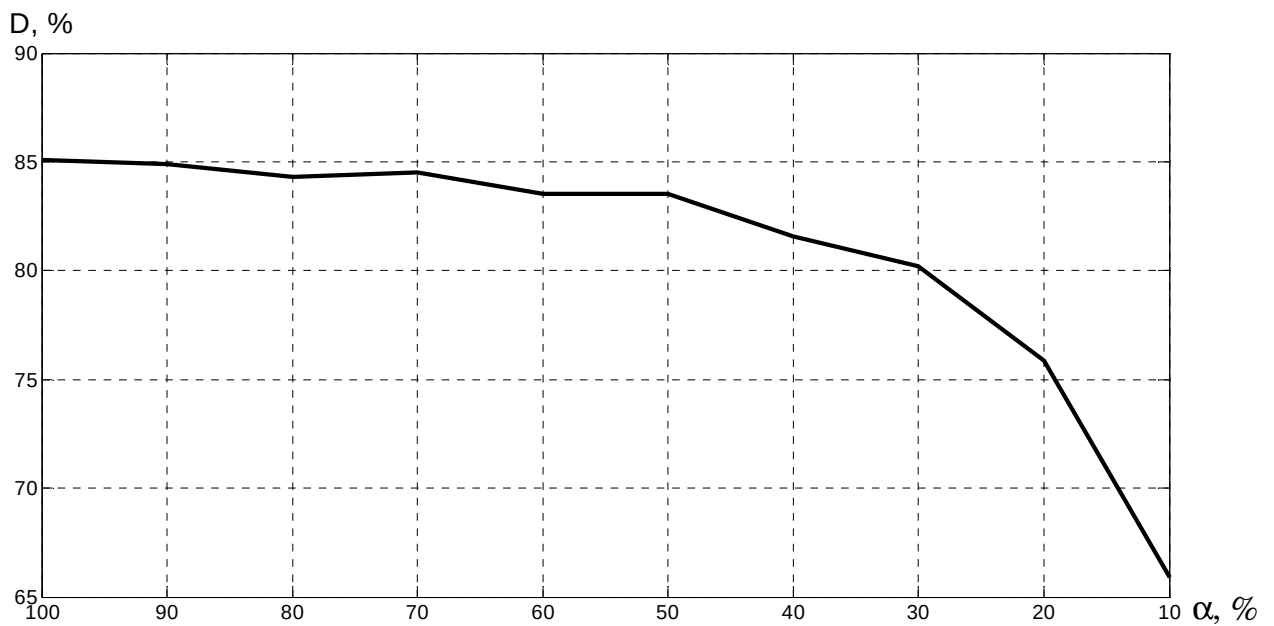


Рис. 3.4. Зависимость процента верной классификации алгоритма на базе МОВ от изменения контраста

Снижение контраста не более чем на 50% незначительно влияет на результат классификации. Но при дальнейшем ухудшении наблюдается резкий спад вероятности верной классификации. Так, при значении контраста искаженного изображения 10% от исходного D составляет всего 66%.

Рассмотрим результаты тестирования алгоритма на базе МОВ с помощью последней модели искажений – неравномерной освещенности (рис. 3.5).

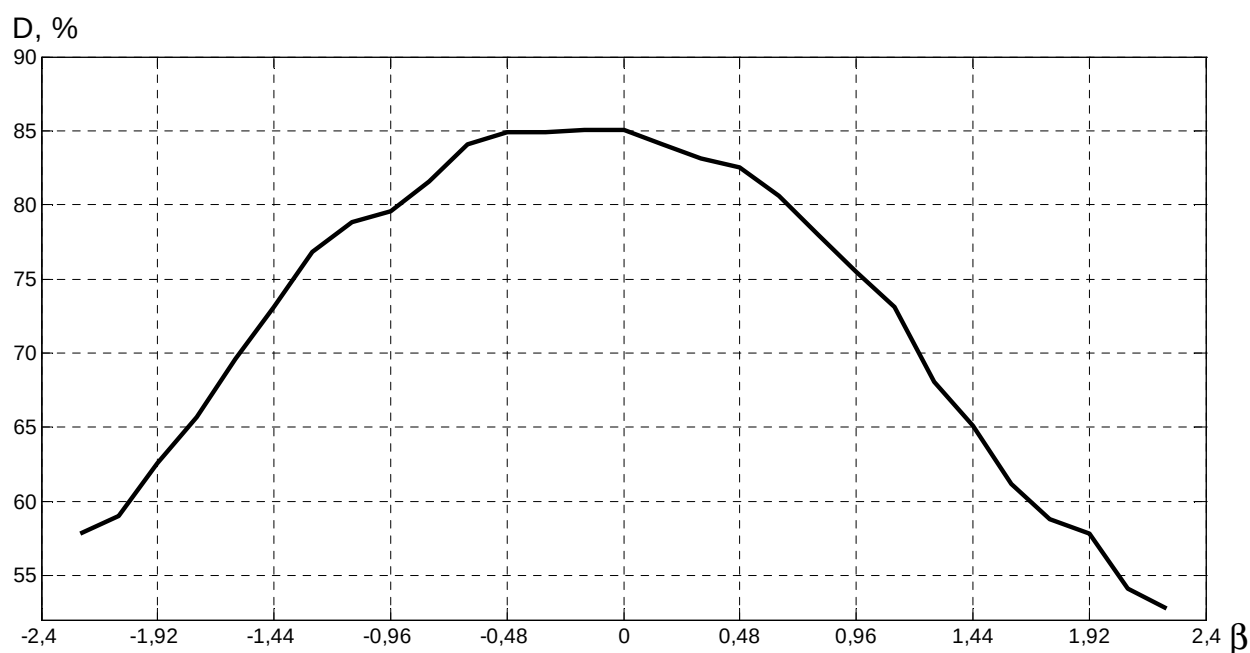


Рис. 3.5. Процент верной классификации алгоритма на базе МОВ в условиях неравномерной освещенности

Зависимость, представленная на рис. 3.5 симметрична, поскольку нет принципиальной разницы в том, с какой стороны находится моделируемый источник света – слева или справа. Процент верной классификации быстро падает при увеличении амплитуды неравномерного освещения. Даже при незначительном искажении ($\beta = 1,44$) D составляет порядка 65-70% от изначальных 85%.

Таким образом, проведенное исследование показывает, что алгоритм классификации пола, основанный на методе опорных векторов, не устойчив к

изменению условий освещенности, что ограничивает его применение в практической ситуации и подтверждает необходимость его модификации.

3.3. Предлагаемые модификации базового алгоритма на основе метода опорных векторов

Существует два основных подхода к решению проблемы устойчивости любого алгоритма на базе машинного обучения к тому или иному типу искажений. Первый подход - расширение обучающей выборки. Добавив в обучающую выборку искаженные изображения, можно заложить информацию о характере искажения в итоговую структуру классификатора. Другой подход заключается в выравнивании характеристик изображения таким образом, чтобы устранить или значительно снизить влияние искажений на результат подобной предобработки, который и подается на вход классификатора.

В случае изменения условий освещенности второй подход является более предпочтительным, поскольку позволяет учесть все возможные виды искажений и не приводит к усложнению процедуры обучения и структуры классификатора. Поэтому далее предлагаются три модификации алгоритма на базе МОВ, заключающиеся в добавлении этапа предобработки. Эти модификации отличаются используемыми алгоритмами предобработки. Первая из них основана на процедуре эквализации гистограммы изображения, являющейся общепринятым известным алгоритмом. Вторая и третья модификации основаны на алгоритмах, предлагаемых в данной работе. Это алгоритм локальной предобработки изображения и алгоритм локальной предобработки с предварительной сегментацией по цвету кожи.

Перейдем к детальному рассмотрению трех различных процедур предобработки, лежащих в основе предлагаемых модификаций:

- эквализация гистограммы изображения;
- локальное выравнивание яркости и контраста;
- локальная предобработка с сегментацией.

3.3.1. Эквиализация гистограммы изображения [50]

Эквиализация гистограммы (ЭГ) является общепринятым и наиболее часто используемым инструментом для приведения яркостных характеристик изображения к стандартному виду. Под гистограммой изображения $h(k)$ понимают вектор размерности 255, каждая составляющая которого равна числу пикселей изображения, интенсивность которых равна k . Целью процедуры эквиализации является минимизация модуля разности между гистограммой изображения и эталонной равномерной гистограммой $h_{gram}(k)$:

$$F(k) = |c_1(T(k)) - c_0(k)|, \quad (3.1)$$

где $F(k)$ – минимизируемая функция; $T(k)$ – оператор преобразования значений яркости; $c_0(k)$ – сумма отсчетов гистограммы обрабатываемого изображения меньших или равных k ; $c_1(k)$ – сумма отсчетов эталонной гистограммы h_{gram} меньших или равных k .

Полученная в результате минимизации функция используется, чтобы преобразовать пиксели исходного изображения с яркостью a в пиксели обработанного изображения с яркостью b согласно формуле:

$$b = T(a), \quad (3.2)$$

При минимизации на функцию T накладывают следующие ограничения: T должна быть монотонной и $c_1(T(k))$ не должно превышать $c_0(a)$ более, чем на половину количества точек с яркостью a .

Результат применения алгоритма эквиализации гистограммы к неискаженному изображению, затемненному изображению ($dY = -60\%$), засвеченному изображению ($dY = 60\%$), изображению с уменьшенным контрастом ($\alpha = 20\%$) и изображению с неравномерной освещенностью ($\beta = -2,4$) приведен на рис. 3.6.

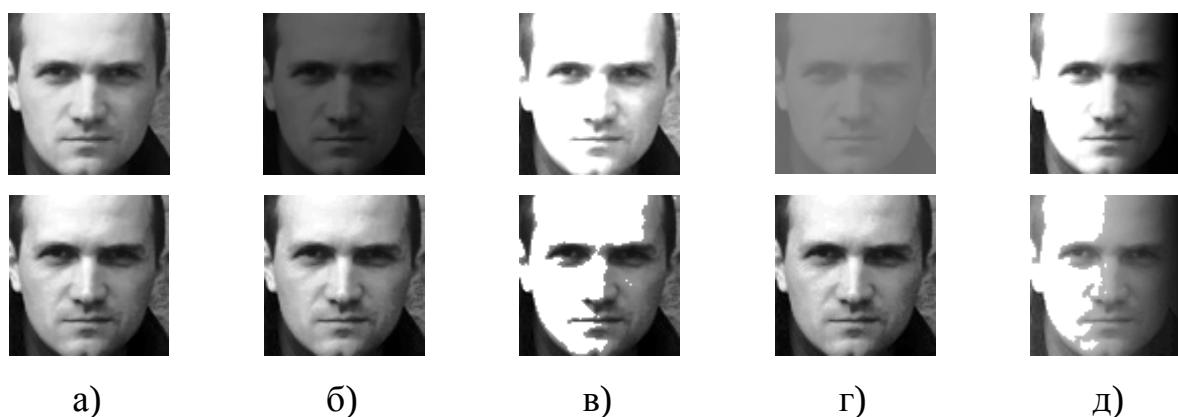


Рис. 3.6. Пары исходное изображение (сверху) – изображение, обработанное алгоритмом ЭГ (снизу), для следующих условий: а) неискаженное изображение; б) $dY = -60\%$; в) $dY = 60\%$; г) $\alpha = 20\%$; д) $\beta = -2,4$

Анализ визуальных примеров показывает, что при затемнении и уменьшении контраста обработанные алгоритмом ЭГ изображения схожи с неискаженным обработанным изображением. Таким образом, эквализация гистограммы позволяет снизить вариативность входных изображений при обучении классификатора. Однако, в случае засвечивания, а также неравномерной освещенности появляются артефакты, не устранимые алгоритмом ЭГ, что может негативно сказаться на результате классификации.

3.3.2. Алгоритм локального выравнивания яркости и контраста (локальная предобработка)

Для того чтобы учесть тот факт, что средняя яркость и контраст отличаются на различных участках изображения, предлагается разделить его на перекрывающиеся между собой блоки. Для каждого заданного блока осуществляется выравнивание яркости и контраста путем применения соответствующих преобразований на основе статистики этого блока. Чтобы получить итоговое значение яркости для пикселей из областей перекрытия применяется взвешенное суммирование значений пикселей перекрывающихся блоков в зависимости от близости пикселя к границе того или иного блока.

В работе использовались изображения разрешением 80×80 пикселей. Изображение разбивалось на 4 блока разрешением 48×48 пикселей (рис. 3.7).

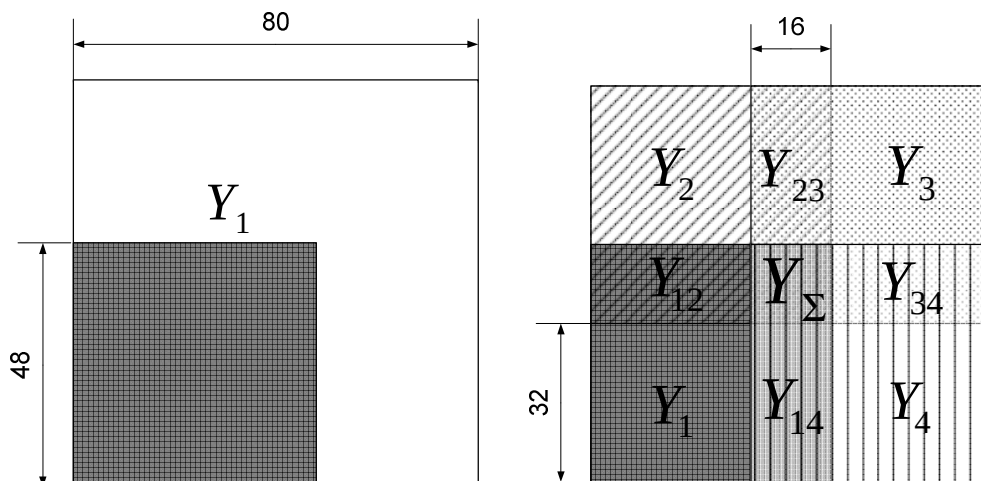


Рис. 3.7. Разбивка изображение на перекрывающиеся блоки

В результате наложения блоков возникало 5 областей перекрытия – 4 области разрешением 16×32 пикселя на границе двух блоков и 1 общая область разрешением 16×16 пикселей.

Для преобразования яркости блока использовалось выражение:

$$Y_k = Y - \langle Y_k \rangle + 0,5, \quad (3.3)$$

где Y – интенсивность пикселей исходного изображения, представленная в формате double (по шкале от 0 до 1); $\langle Y_k \rangle$ – среднее значение интенсивности пикселей исходного изображения, входящих в блок k . В результате такого преобразования средняя яркость блока оказывается точно на середине шкалы яркости и равна 0,5. Для пикселей, интенсивность которых в результате преобразования оказалась за пределами шкалы, применяется усечение, после чего значение остальных пикселей вновь корректируется, чтобы средняя яркость блока оставалась равной 0,5.

Для преобразования контраста значения всех пикселей внутри блока сортируются по возрастанию. Определенный небольшой процент самых

ярких и самых темных пикселей отбрасывается, после чего яркость оставшихся пикселей рассчитывается по формуле:

$$Y_k = \frac{Y - 0,5}{K} + 0,5, \quad (3.4)$$

где $K = \max\{(1 - 2 \cdot Y_{min}), (2 \cdot Y_{max} - 1)\}$ – коэффициент, показывающий контраст входного изображения и определяющий степень увеличения контраста обработанного изображения. Таким образом, контраст каждого блока максимально расширяется при сохранении информации. Пиксели, интенсивность которых в результате преобразования оказалась за пределами шкалы, подвергаются усечению.

Для того чтобы избежать эффектов блочности на границах перекрытия блоков, производится взвешенное суммирование значений пикселей соседних блоков:

$$Y_{lm} = \frac{a_{lj} \cdot Y_l + a_{mj} \cdot Y_m}{h}, \quad (3.5)$$

где a_{lj} , a_{mj} – расстояние от j -ого пикселя до границы соответствующего блока; Y_l , Y_m – интенсивность пикселя, посчитанная для блока l и m соответственно; h – ширина зоны перекрытия.

Рассмотрим результат применения предлагаемого алгоритма локальной предобработки (ЛП) к неискаженному изображению, затемненному изображению ($dY = -60\%$), засвеченному изображению ($dY = 60\%$), изображению с уменьшенным контрастом ($\alpha = 20\%$) и изображению с неравномерной освещенностью ($\beta = -2,4$) (рис. 3.8).

Анализ визуальных примеров показывает, что несмотря на небольшую потерю высокочастотной информации, вызванную усечением, алгоритм ЛП позволяет привести яркость и контраст изображения к единым характеристикам при любом виде искажений, связанных с изменением условий освещенности. Таким образом, алгоритм ЛП снижает вариативность входных изображений при обучении классификатора, его использование может дать преимущество в проценте верного распознавания по сравнению с использованием процедуры эквализации гистограммы.

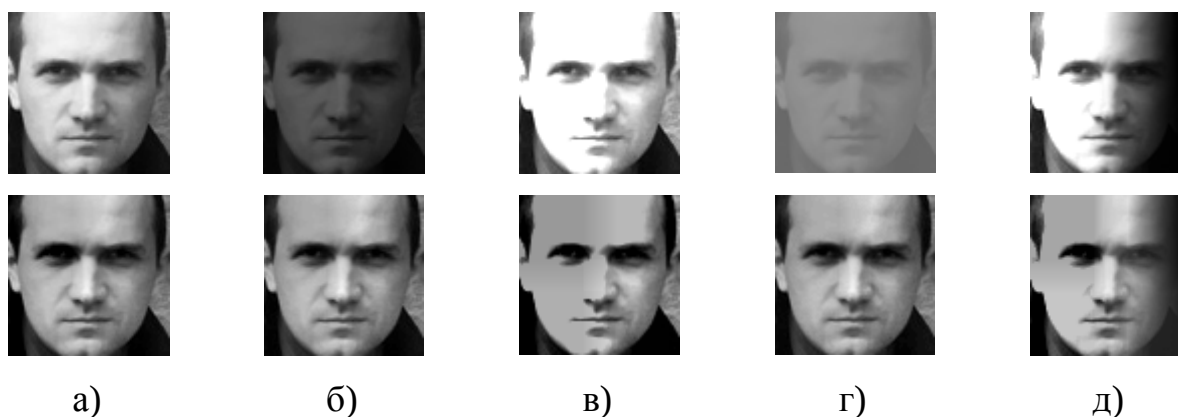


Рис. 3.8. Пары исходное изображение (сверху) – изображение, обработанное алгоритмом ЛП (снизу), для следующих условий: а) неискаженное изображение; б) $dY = -60\%$; в) $dY = 60\%$; г) $\alpha = 20\%$; д) $\beta = -2,4$

3.3.3. Алгоритм локальной предобработки с сегментацией

Предложенный алгоритм локальной предобработки хорошо справляется с различными условиями освещенности, однако, существует возможность еще больше снизить вариативность изображений за счет применения сегментации. В связи с тем, что область детекции лица имеет прямоугольную форму, а положение самого лица внутри этой области различается в зависимости от наклона и угла поворота головы, выделенный фрагмент содержит области, не относящиеся к лицу. Эти области могут иметь какую угодно яркость и структуру, что может помешать как работе алгоритма предобработки, так и непосредственно работе классификатора. Поэтому предлагается алгоритм локальной предобработки с сегментацией (СЛП), отличающийся от алгоритма ЛП предварительным разделением всех пикселей изображения на пиксели лица и пиксели фона.

Для этого сначала производится сегментация пикселей кожи. Следует отметить, что алгоритмы, рассмотренные ранее, отбрасывали цветовые компоненты изображения, как несущие малое количество информации, здесь же для сегментации кожи используется именно информация о цвете. В работе применялся известный классификатор, который задает явные численные правила, ограничивающие кластер цвета кожи в пространстве

RGB [99]. Пиксель (R, G, B) относится к коже, если для него выполнены все следующие условия:

- $R > 95$;
- $G > 40$;
- $B > 20$;
- $\max\{R, G, B\} - \min\{R, G, B\} > 15$;
- $|R - G| > 15$;
- $R > G$;
- $R > B$.

Данный алгоритм не требует перехода в цветовую систему, отличную от цветовой системы исходного изображения, не требует больших вычислительных ресурсов, показывает хорошие результаты в ряде задач смежных с распознаванием пола, например, при детектировании области лица на изображении [100].

Далее для каждого пикселя из полученной маски определяем, является ли он пикселем фона, либо пикселем лица, не являющимся кожей. Для этого применяется морфологическая фильтрация. Если число пикселей кожи, окружающих данный пиксель, меньше порога, то этот пиксель принадлежит фону. Пример полученной маски в результате процедуры сегментации изображения для некоторых тестовых изображений приведен на рис. 3.9.



Рис. 3.9. Пары исходное изображение (сверху) – маска, полученная в результате сегментации (снизу), для некоторых тестовых изображений

Здесь пиксели, классифицированные как кожа, обозначены белым, пиксели, отнесенные к лицу по результатам морфологической фильтрации обозначены синим, пиксели фона – черным, пиксели, отнесенные к границе лица – желтым. Визуальные примеры показывают, что предложенный алгоритм сегментации успешно решает поставленную задачу, позволяя в конечном счете разделить все пиксели на пиксели лица и пиксели фона.

Возвращаясь к алгоритму локальной предобработки с сегментацией, предлагается пиксели, относящиеся к фону, приравнять к 0,5 (средней яркости), остальные пиксели обрабатывать с помощью процедуры, аналогичной работе алгоритма ЛП.

Результат применения алгоритма СЛП к неискаженному изображению, затемненному изображению ($dY = -60\%$), засвеченному изображению ($dY = 60\%$), изображению с уменьшенным контрастом ($\alpha = 20\%$) и изображению с неравномерной освещенностью ($\beta = -2,4$) приведен на рис. 3.10.

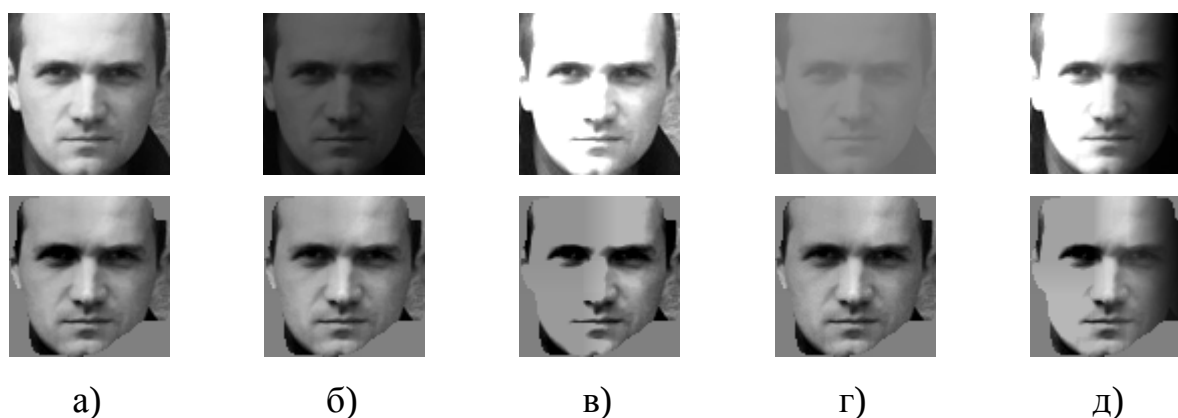


Рис. 3.10. Пары исходное изображение (сверху) – изображение, обработанное алгоритмом СЛП (снизу), для следующих условий: а) неискаженное изображение; б) $dY = -60\%$; в) $dY = 60\%$; г) $\alpha = 20\%$; д) $\beta = -2,4$

Анализ визуальных примеров показывает, что в дополнение к выравниванию яркости и контраста алгоритм СЛП позволяет отсеять области, не относящиеся к лицу, что снижает вариативность и может положительно сказаться на результате классификации.

3.4. Результаты тестирования предложенных модификаций

Для тестирования эффективности предложенных алгоритмов предобработки в задаче повышения устойчивости алгоритмов классификации пола к изменению условий освещенности было обучено три различных модификации базового классификатора на базе МОВ:

- с предварительной эквализацией гистограммы (ЭГ-МОВ);
- с локальной предобработкой яркости и контраста (ЛП-МОВ);
- с сегментацией и последующей предобработкой (СЛП-МОВ).

Для обучения и тестирования использовалась та же база лиц, что и для базового алгоритма, что обеспечивает равные условия при их сравнении. ROC кривые тестируемых алгоритмов приведены на рис. 3.11.

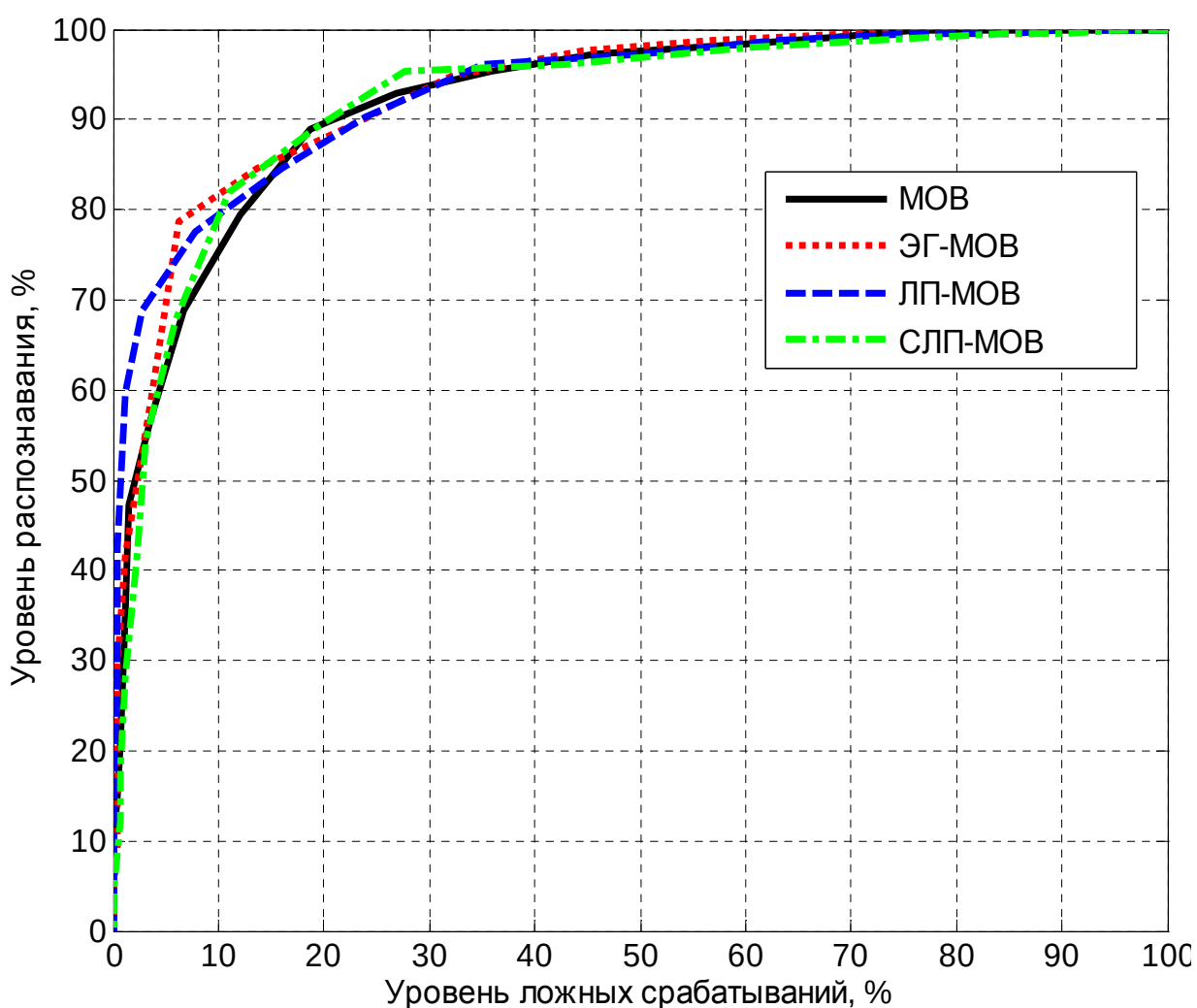


Рис. 3.11. ROC кривые классификатора на базе МОВ и трех его модификаций

Вероятность верной классификации, полученная на наборе неискаженных изображений, приведена в табл. 3.1.

Таблица 3.1

Уровень верного распознавания пола тестируемых алгоритмов

Алгоритм	МОВ		ЭГ-МОВ		ЛП-МОВ		СЛП-МОВ	
Параметр	Верно	Неверно	Верно	Неверно	Верно	Неверно	Верно	Неверно
Уровень распознавания мужчин, %	89,0	11,0	87,4	12,6	83,9	16,1	85,8	14,2
Уровень распознавания женщин, %	81,3	19,7	81,6	19,4	85,5	14,5	86,3	13,7
Общий уровень распознавания, %	85,1	14,9	84,5	15,5	84,7	15,3	86,1	13,9

Результаты показывают, что введение блока предобработки в состав алгоритма на базе МОВ незначительно сказывается на результате классификации на наборе неискаженных изображений. В случае эквализации гистограммы и локальной обработки наблюдается снижение вероятности верного распознавания на 0,5%. Предобработка с сегментацией позволяет добиться улучшения вероятности верного распознавания на 1%. В связи с такой малой разницей по уровню верного распознавания ROC кривые обученных классификаторов практически идентичны. Также следует отметить, что предобработка позволяет снизить несимметричность в вероятности распознавания пола мужчин и женщин. Так для МОВ эта несимметричность составляет 7,7%; для ЭГ-МОВ – 5,8%; для ЛП-МОВ – 1,6%; а для СЛП-МОВ – 0,5%.

Перейдем к рассмотрению результатов исследования предлагаемых модификаций классификатора на базе МОВ в различных условиях освещения. Зависимость процента верной классификации алгоритмов ЭГ-МОВ, ЛП-МОВ и СЛП-МОВ от одновременного изменения яркости и контраста приведена на рис. 3.12, 3.13 и 3.14 соответственно.

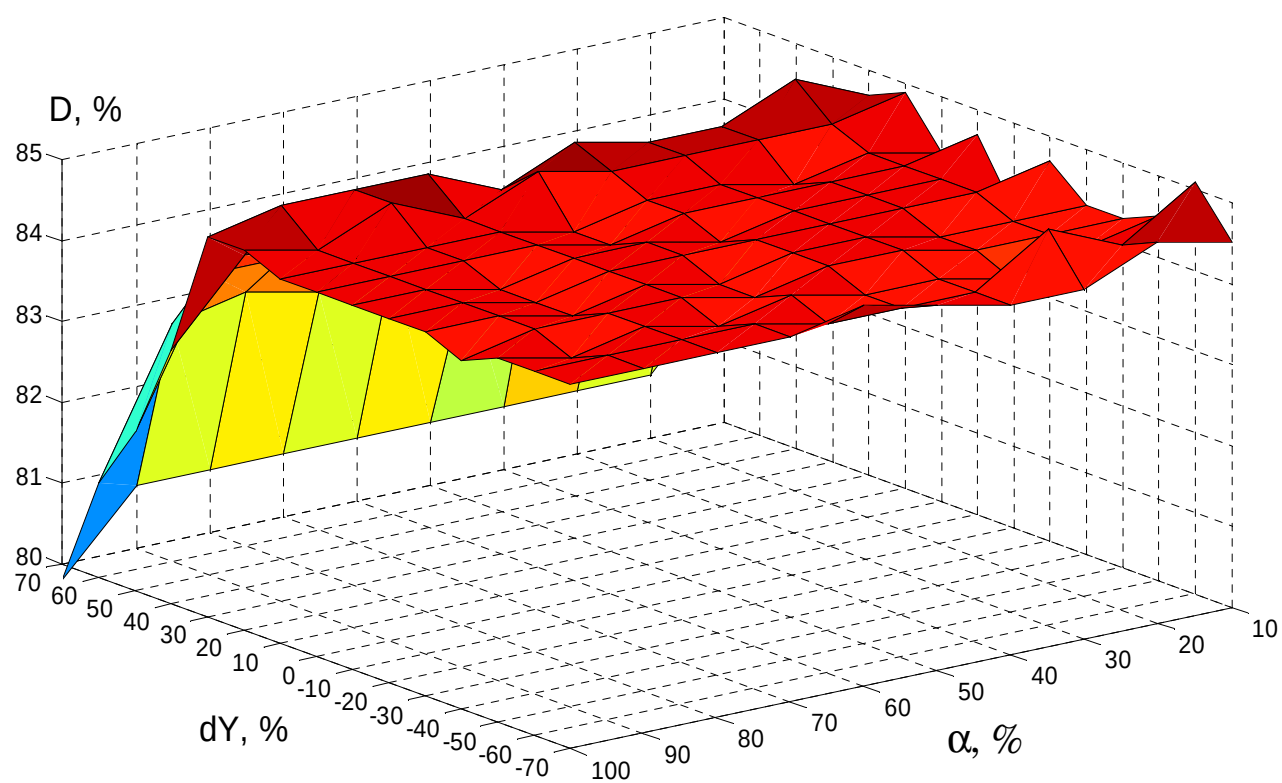


Рис. 3.12. Зависимость процента верной классификации алгоритма ЭГ-МОВ от изменения яркости и контраста

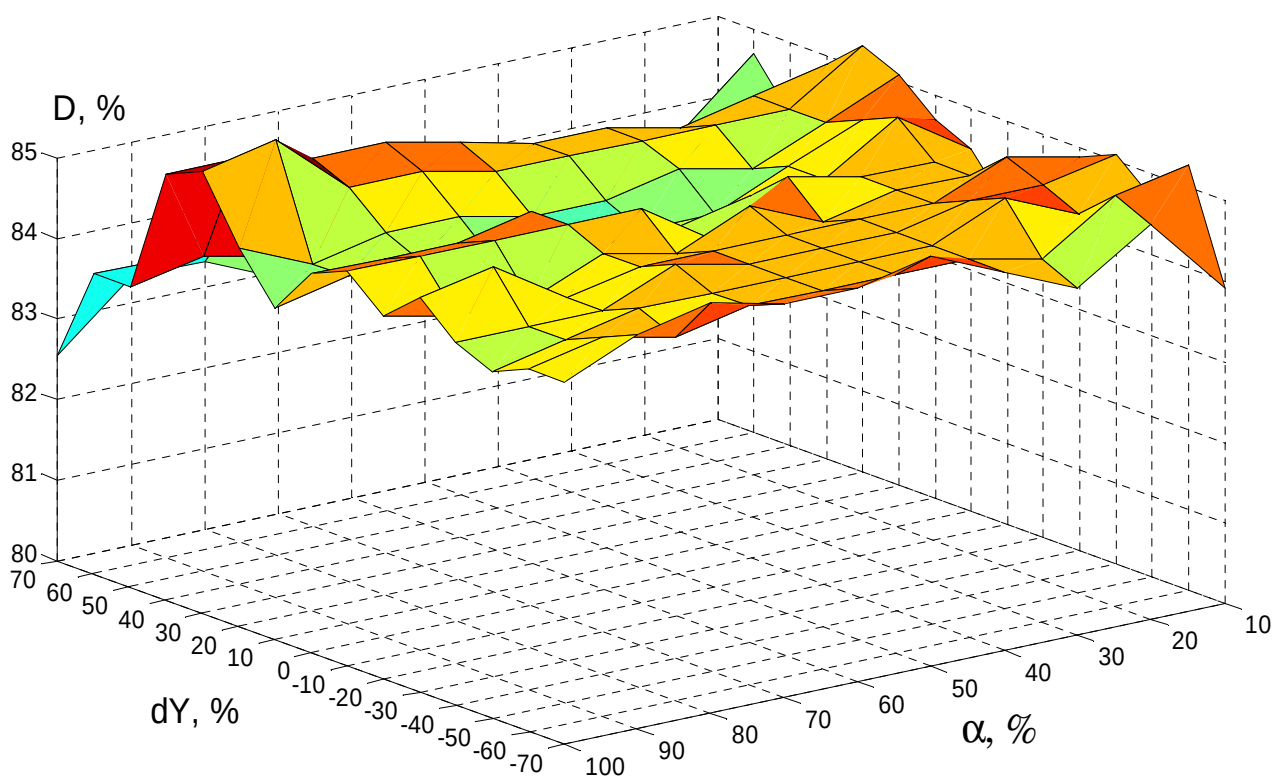


Рис. 3.13. Зависимость процента верной классификации алгоритма ЛП-МОВ от изменения яркости и контраста

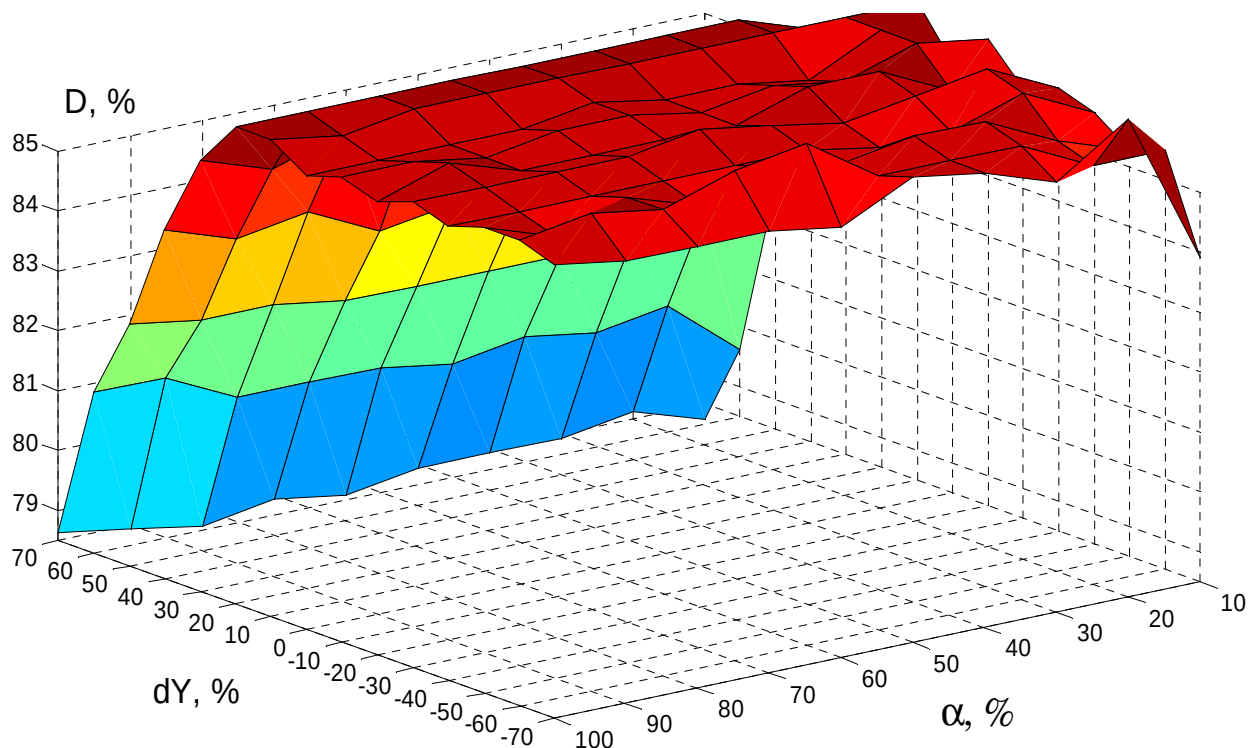


Рис. 3.14. Зависимость процента верной классификации алгоритма СЛП-МОВ от изменения яркости и контраста

Результаты показывают, что все три алгоритма предобработки позволяют добиться существенного улучшения уровня верного распознавания на наборе искаженных изображений по сравнению с изначальным классификатором на базе МОВ. Так, алгоритмы ЭГ-МОВ и ЛП-МОВ имеют вероятность верной классификации порядка 84% в широком диапазоне изменения яркости и контраста, и только при существенном засвечивании изображений их эффективность снижается до 80% и 82,5% соответственно. Алгоритм СЛП-МОВ имеет небольшое преимущество в среднем по уровню распознавания (порядка 1%), однако, в условиях засвечивания показывает только 79%.

Для того чтобы более детально сравнить поведение тестируемых алгоритмов, рассмотрим зависимости их уровня классификации от отдельных условий освещенности: изменения яркости, изменения контраста, неравномерной освещенности. Результаты исследований приведены на рис.

3.15, 3.16 и 3.17 для каждого из рассматриваемых условий освещенности соответственно.

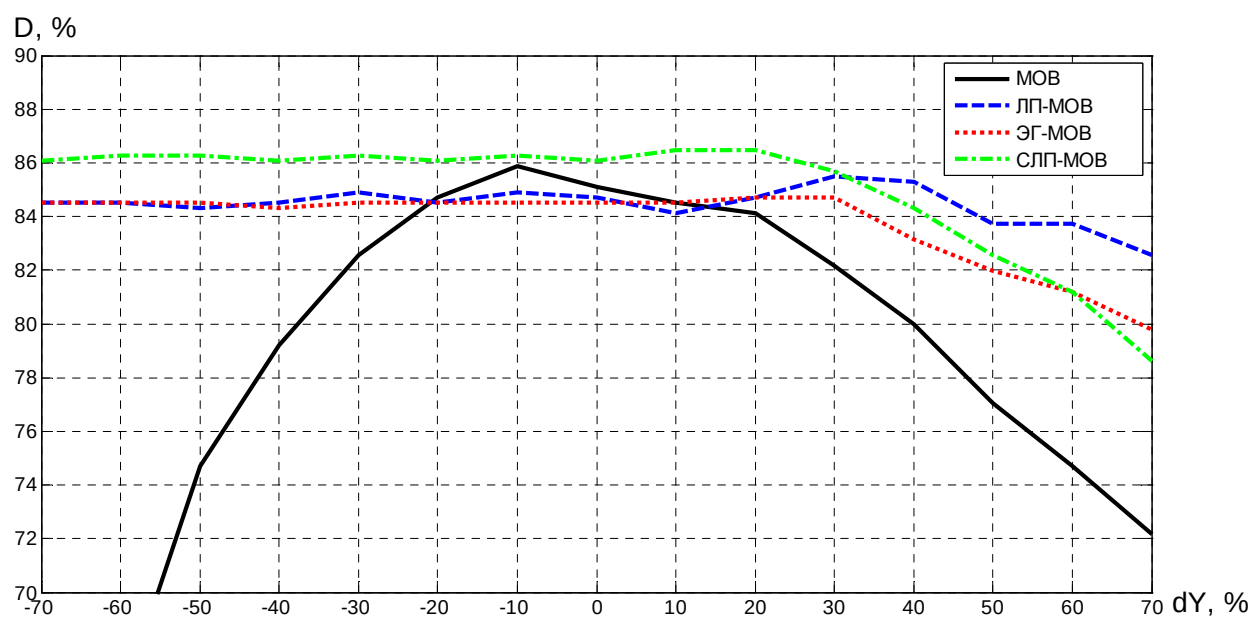


Рис. 3.15. Зависимость процента верной классификации алгоритма на базе МОВ и трех его модификаций от изменения яркости

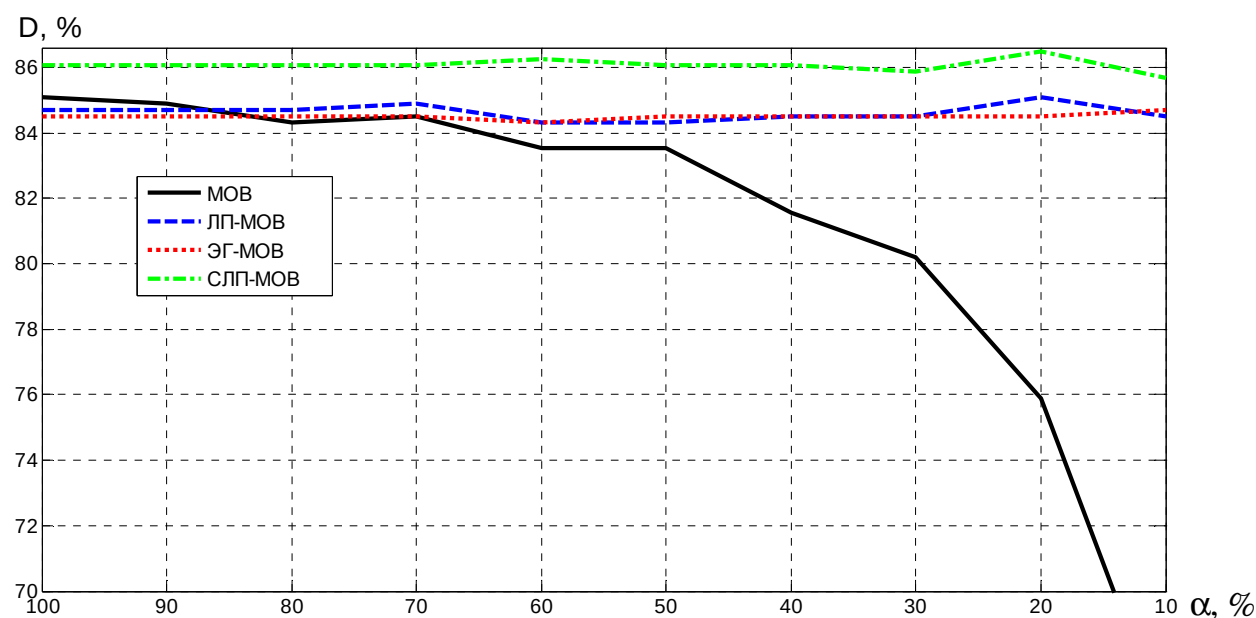


Рис. 3.16. Зависимость процента верной классификации алгоритма на базе МОВ и трех его модификаций от изменения контраста

Из рис. 3.15 следует, что на большем участке шкалы изменения яркости лидирует алгоритм СЛП-МОВ (86%), на втором месте ЛП-МОВ (84,5%), на

третьем – ЭГ-МОВ (84%). При сильном засвечивании лучший результат имеет классификатор ЛП-МОВ (82,5%), на втором и третьем местах алгоритмы ЭГ-МОВ (80%) и СЛП-МОВ (79%) соответственно. Выигрыш по уровню верного распознавания по сравнению с классификатором МОВ без этапа предобработки составляет в среднем более 10% для любой из предложенных модификаций.

Рассматривая зависимость процента верной классификации от изменения контраста, очевидно большое преимущество алгоритмов ЭГ-МОВ (84,5%), ЛП-МОВ (84,5%) и СЛП-МОВ (86%) над алгоритмом МОВ. Так, при ухудшении контраста в 5 раз разница между МОВ и СЛП-МОВ достигает 10% в пользу алгоритма СЛП-МОВ.

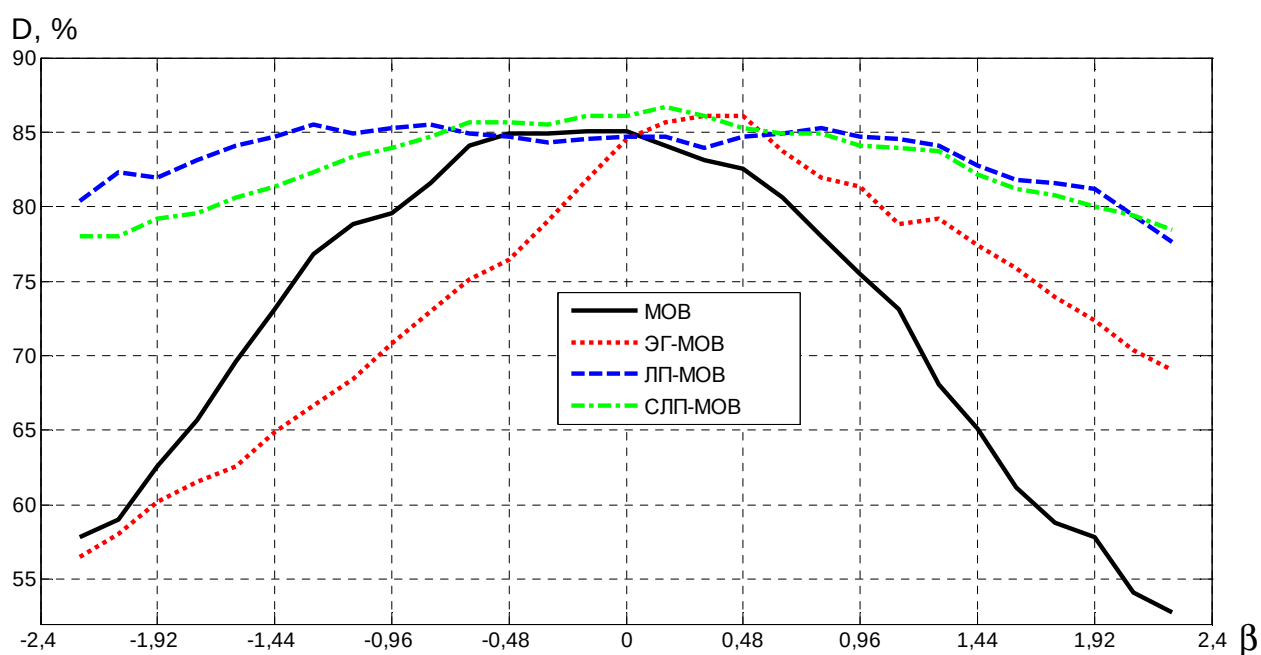


Рис. 3.17. Процент верной классификации алгоритма на базе МОВ и трех его модификаций в условиях неравномерной освещенности

В условиях неравномерной освещенности рассматриваемые алгоритмы ведут себя по-разному. Алгоритм ЭГ-МОВ оказывается неспособным привести характеристики изображения к единому виду при увеличении амплитуды бокового освещения. Его уровень верного распознавания в среднем аналогичен уровню, показанному классификатором МОВ без

предобработки. Алгоритмы ЛП-МОВ и СЛП-МОВ успешно справляются с выравниванием яркостных характеристик. Их уровень верной классификации составляет в среднем порядка 82,5%, что более чем на 10% больше того же показателя для алгоритмов МОВ и ЭГ-МОВ.

Таким образом, можно сделать вывод, что предложенные модификации алгоритма классификации пола на базе МОВ позволяют достичь преимущества по среднему уровню верного распознавания при изменении условий освещенности. Лучшие результаты демонстрирует алгоритм СЛП-МОВ. Он опережает алгоритм МОВ по уровню верной классификации в среднем более чем на 10% как при изменении яркости, так и при изменении контраста, а также в условиях неравномерной освещенности. Алгоритм СЛП-МОВ показывает уровень верного распознавания на 1% выше по сравнению с базовым алгоритмом МОВ на наборе неискаженных изображений.

3.5 Краткие выводы

Результаты проведенной работы позволяют сделать следующие основные выводы.

- Исследована устойчивость базового классификатора на основе метода опорных векторов к изменению условий освещенности при распознавании пола человека по изображению лица. Вероятность верного распознавания уменьшается с 85% при нормальных условиях до 55% и 72% при уменьшении и увеличении яркости соответственно, до 66% при уменьшении контраста и до 55% при неравномерном освещении.

- Разработано три модификации базового алгоритма на базе МОВ с применением одного стандартного и двух оригинальных алгоритмов предобработки изображений.

- Предлагаемые модификации имеют преимущество по среднему уровню верного распознавания порядка 10-15% по сравнению с базовым алгоритмом МОВ при различных условиях освещенности. Лучшие результаты демонстрирует алгоритм СЛП-МОВ.

4. МОДИФИКАЦИЯ ПРИЗНАКОВОГО ПРОСТРАНСТВА АЛГОРИТМА РАСПОЗНАВАНИЯ ПОЛА НА БАЗЕ МЕТОДА ОПОРНЫХ ВЕКТОРОВ

4.1. Вводные замечания

Данный раздел посвящен усовершенствованию базового алгоритма на базе МОВ с целью повышения вероятности верного распознавания на наборе неискаженных изображений. В предыдущем разделе рассматривались алгоритмы предобработки изображений, влияющие на вероятность верного распознавания классификатора в различных условиях освещенности. Однако, применение предложенных модификаций не позволяет добиться преимущества в нормальных условиях. Для решения этой задачи предлагается модифицировать другой этап работы классификатора, а именно – переход в пространство признаков.

Результаты экспериментов, приведенных в первой главе данной работы, показывают преимущество использования адаптивных признаков по сравнению с базовым алгоритмом МОВ. Адаптивные признаки были предложены в работе [85]. Они являются интегральными, т.е. значение конкретного признака зависит от интенсивности каждого пикселя анализируемого изображения. В работе [101] были предложены признаки, называемые локальными бинарными шаблонами (ЛБШ) (local binary patterns). Они хорошо зарекомендовали себя в задачах выделения объектов на изображениях [65], однако, их применение в задаче классификации пола по изображению лица оставалось неисследованным. Признаки ЛБШ являются локальными, поскольку для вычисления конкретного признака используются интенсивности пикселей из заданного региона анализируемого изображения, а не всего изображения целиком.

В данной работе рассматривается возможность применения локальных бинарных шаблонов в задаче классификации пола по изображению лица. Для этого осуществляется переход от значений пикселей в пространство ЛБШ

признаков и классификация на базе метода опорных векторов. Назовем полученную таким образом модификацию базового алгоритма на базе МОВ – ЛБШ-МОВ.

Другой возможностью повышения уровня верного распознавания является объединение нескольких признаков пространств в единый вектор признаков. В случае использования адаптивных признаков и локальных бинарных шаблонов этот выигрыш может оказаться значительным, поскольку процедуры вычисления этих признаков значительно отличаются друг от друга, и они несут в себе различную информацию, необходимую для разделения классов. Для того, чтобы проверить способность адаптивных признаков и локальных бинарных шаблонов дополнять друг друга был обучен классификатор на базе МОВ, входом которому служил объединенный вектор признаков. Эта модификация далее носит название АП-ЛБШ-МОВ.

Рассмотрим процедуру вычисления локальных бинарных шаблонов, а также результаты тестирования предложенных на их базе модификаций: ЛБШ-МОВ и АП-ЛБШ-МОВ.

4.2. Процедура вычисления локальных бинарных шаблонов

Процедуру вычисления локальных бинарных шаблонов удобно представить в виде двух этапов. На первом этапе к каждому пикселю изображения применяется определенный оператор. Этот оператор работает так, как показано на рис. 4.1.

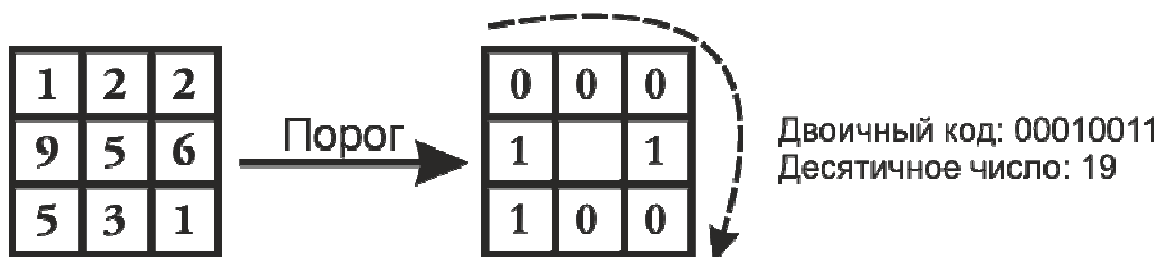


Рис. 4.1. Первый этап вычисления локальных бинарных шаблонов

Вначале значение интенсивности пикселя сравнивается со значениями всех пикселей из некоторой окрестности, например, размером 3×3 пикселя. Результат сравнения записывается как 0, если значение рассматриваемого пикселя меньше центрального, и как 1 в противном случае. Для рассматриваемой окрестности 3×3 пикселя получается 8 цифр, из которых составляется двоичный вектор. Этот вектор далее интерпретируется как двоичная запись целого числа, которое и является результатом применения оператора к пикселю. В случае выбора окрестности размером 3×3 пикселя диапазон значений, получаемых в результате применения оператора, составляет от 0 до 255. Этот диапазон совпадает с диапазоном интенсивностей пикселей в стандартном представлении яркостной компоненты изображения. Поэтому результат применения первого этапа вычисления ЛБШ признаков удобно представить в виде изображения, интенсивность каждого пикселя которого равна результату рассмотренного выше преобразования.

На втором этапе полученное в ходе первого этапа изображение разбивается на прямоугольные области. В данной работе использовалось несколько последовательно примененных разбиений: на области 10×10 пикселей; 20×20 пикселей и 40×40 пикселей. Такие значения были выбраны, исходя из разрешения анализируемых объектов 80×80 пикселей и условия целого числа блоков при разбиении. В результате общее число областей, на которые разбивалось изображение, составило 84.

Далее для каждой из 84 используемых областей вычислялась её гистограмма – отношение количества чисел заданного значения (от 0 до 255) к общему количеству чисел в рассматриваемой области. Таким образом, каждая гистограмма содержала 256 значений и являлась вектором размерности 256.

Итоговый вектор признаков получался путем конкатенации гистограмм по всем областям в один вектор. Размерность этого вектора в нашем случае

составила 21504. Второй этап вычисления локальных бинарных шаблонов показан на рис. 4.2.

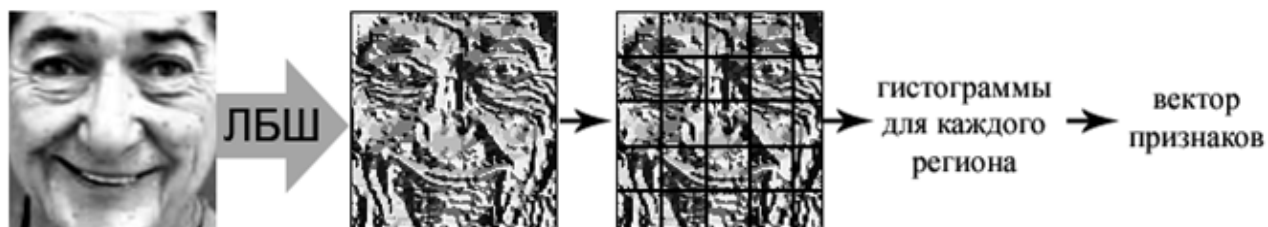


Рис. 4.2. Второй этап вычисления локальных бинарных шаблонов

ЛБШ признаки имеют ряд преимуществ, обусловленных особенностями их вычисления на каждом из этапов. Благодаря квантованию, применяемому на первом этапе при сравнении интенсивностей пикселей, ЛБШ признаки устойчивы к небольшим изменениям освещенности. Применение гистограмм на втором этапе обеспечивает устойчивость к небольшим сдвигам в положении лица, которая достигается за счет того, что подсчет ведется не индивидуально для каждого пикселя, а используются области достаточно большого размера.

4.3. Результаты тестирования предложенных модификаций признакового пространства

Для того, чтобы оценить эффективность рассмотренного пространства признаков в задаче распознавания пола по изображению лица был обучен классификатор на базе МОВ, входом которому служили вычисленные вектора ЛБШ признаков. Обучение и тестирование классификатора проводилось с использованием тех же обучающих и тестовых баз, что и в случае обучения и тестирования базового алгоритма, что гарантирует корректность сравнения.

Рассмотрим ROC кривую алгоритма ЛБШ-МОВ по сравнению с ROC кривой базового классификатора МОВ (рис. 4.3).

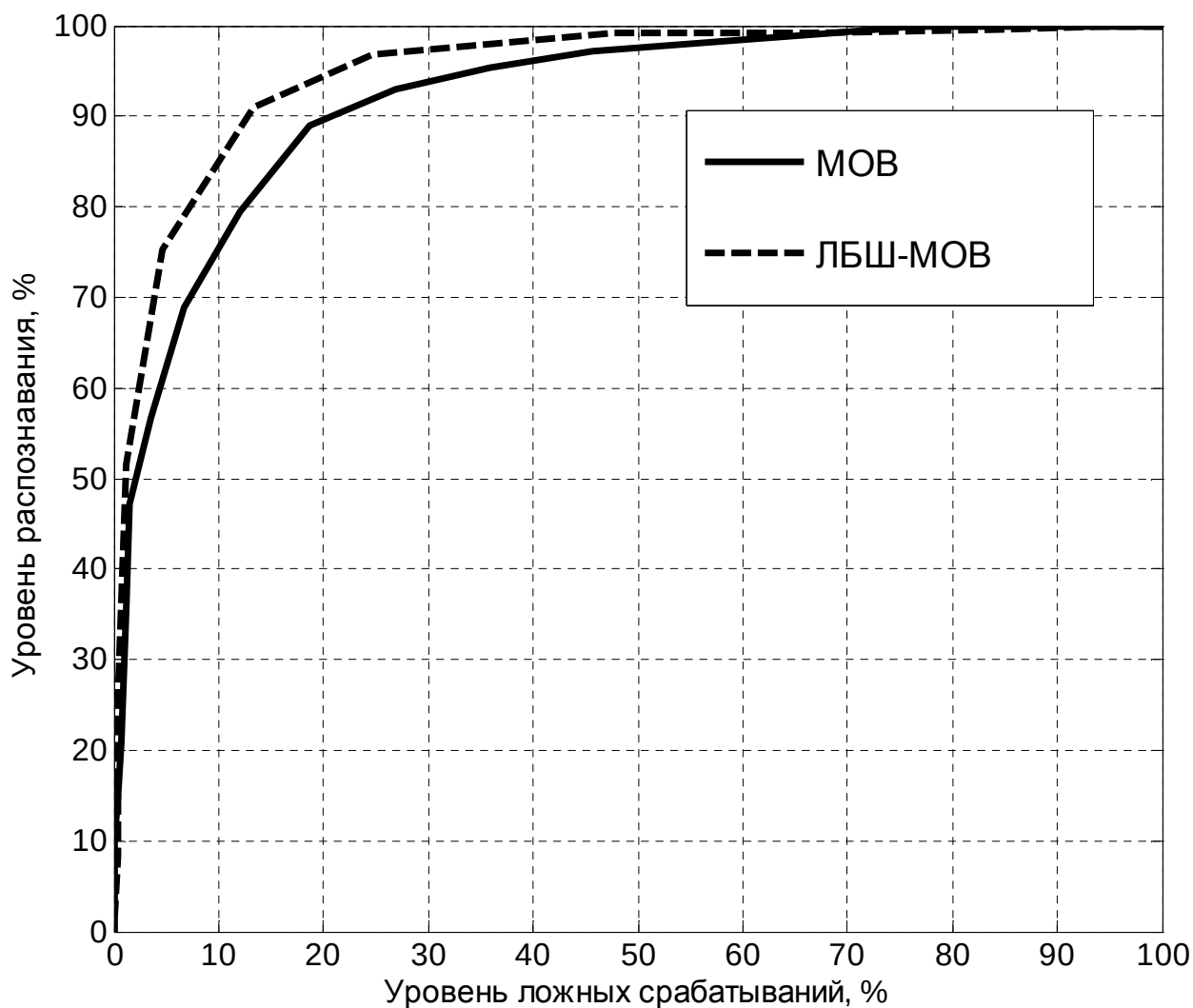


Рис. 4.3. ROC кривая алгоритма ЛБШ-МОБ по сравнению с базовым алгоритмом

Использование ЛБШ признаков позволяет добиться преимущества в уровне распознавания по сравнению с классификатором МОБ. Так, при уровне ложных срабатываний 10% вероятность верного распознавания пола алгоритмом на базе МОБ составляет 75%, а та же величина для алгоритма ЛБШ-МОБ – 85%. Из рис. 4.3 можно сделать вывод, что алгоритм ЛБШ-МОБ имеет преимущество над базовым классификатором МОБ на всех участках ROC кривой.

Сравним основные показатели известного алгоритма АП-МОБ, предложенной модификации ЛБШ-МОБ и базового алгоритма на базе МОБ. Результаты классификации алгоритмами МОБ, АП-МОБ и ЛБШ-МОБ приведены в табл. 4.1.

Таблица 4.1

Уровень распознавания алгоритма ЛБШ-МОВ по сравнению с существующими алгоритмами

Алгоритм	ЛБШ-МОВ		АП-МОВ		МОВ	
Параметр	Верно	Неверно	Верно	Неверно	Верно	Неверно
Уровень распознавания мужчин, %	92,9	7,1	90,6	9,4	89,0	11,0
Уровень распознавания женщин, %	85,9	14,1	91,0	9,0	81,3	19,7
Общий уровень распознавания, %	89,4	11,6	90,8	9,2	85,1	14,9

ЛБШ-МОВ показывает уровень верного распознавания равный 89,4% и опережает алгоритм МОВ на 4,3%, однако, уступает классификатору АП-МОВ на 1,4%. В случае алгоритма ЛБШ-МОВ сохраняется несимметричность в уровне распознавания мужчин и женщин, наблюдаемая у базового классификатора МОВ. Лучшие результаты в этом отношении демонстрирует алгоритм АП-МОВ.

Рассмотрим вторую предложенную модификацию – АП-ЛБШ-МОВ, заключающуюся в комбинировании адаптивных и ЛБШ признаков. Адаптивные признаки обучались только для яркостной компоненты изображения в соответствии с методикой, описанной в работе [85]. Было обучено 100 адаптивных признаков, и, таким образом, размерность итогового пространства признаков составила 21604. Каждый признак был нормирован относительно его максимального и минимального значения в обучающей выборке, чтобы его значения находились в диапазоне от 0 до 1. Далее в полученном признаковом пространстве, как и в предыдущем случае, был обучен классификатор на базе метода опорных векторов. Для обучения и тестирования использовались те же наборы изображений, что и для базового классификатора МОВ.

Сравнение ROC кривых предлагаемой модификации АП-ЛБШ-МОВ, алгоритма АП-МОВ и базового алгоритма МОВ представлено на рис. 4.4.

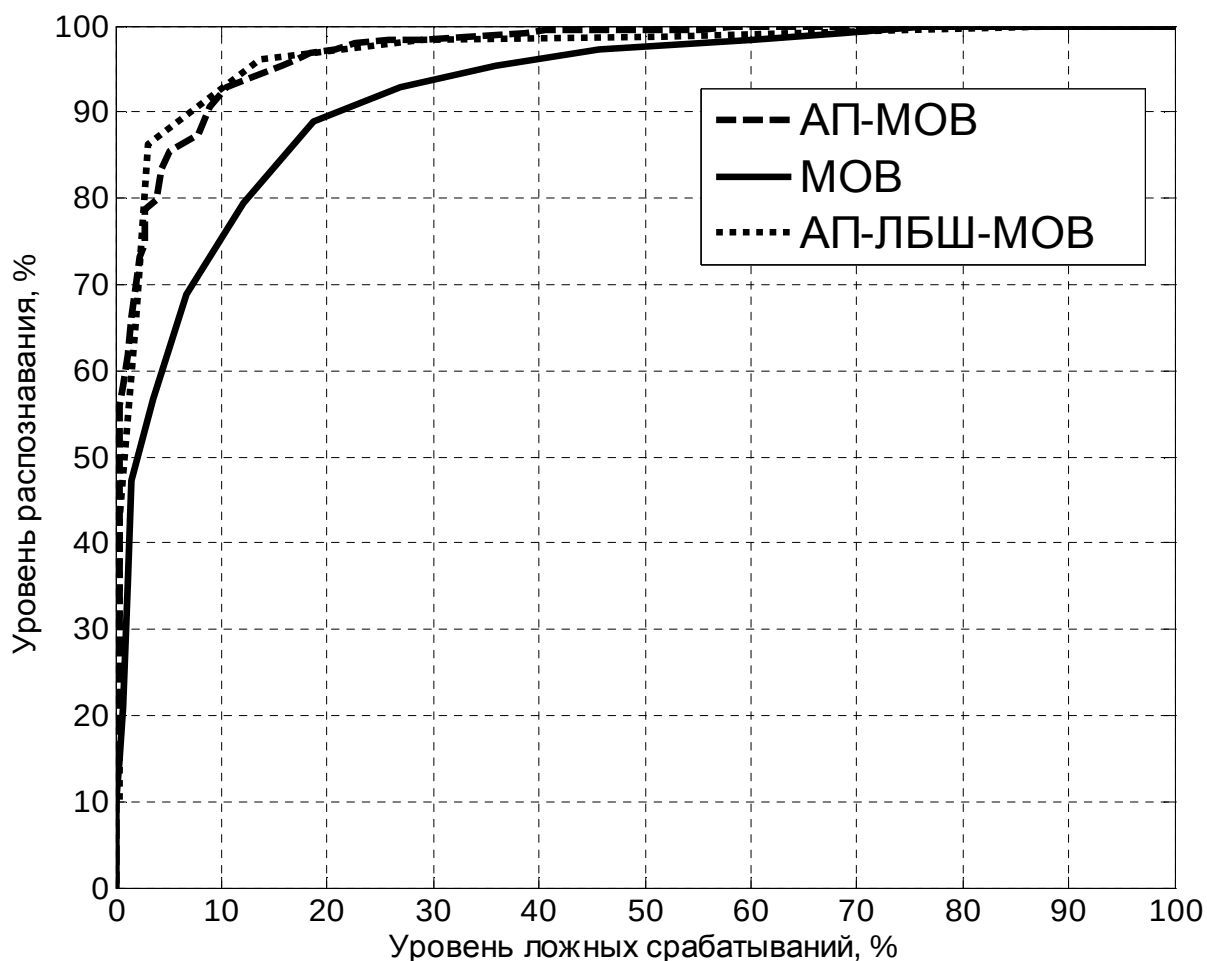


Рис. 4.4. ROC кривая алгоритма АП-ЛБШ-МОВ по сравнению с базовым алгоритмом и алгоритмом АП-МОВ

Алгоритм АП-ЛБШ-МОВ опережает алгоритм АП-МОВ по уровню распознавания при одинаковом уровне ложных срабатываний. Выигрыш составляет 3-5% в диапазоне уровня ложных срабатываний 4-8%. Необходимо отметить тот факт, что алгоритм АП-МОВ содержит адаптивные признаки, посчитанные для цветовых компонент изображения в пространстве HSV, а алгоритм АП-ЛБШ-МОВ использует только яркостную компоненту, т.е. пригоден для распознавания, как цветных, так и монохромных изображений. Оба алгоритма существенно опережают базовый алгоритм MOB.

Результаты классификации алгоритмами MOB, АП-МОВ и АП-ЛБШ-МОВ приведены в табл. 4.2.

Таблица 4.2

Уровень распознавания алгоритма АП-ЛБШ-МОВ по сравнению с существующими алгоритмами

Алгоритм	АП-ЛБШ-МОВ		АП-МОВ		МОВ	
Параметр	Верно	Неверно	Верно	Неверно	Верно	Неверно
Уровень распознавания мужчин, %	93,9	6,1	90,6	9,4	89,0	11,0
Уровень распознавания женщин, %	94,3	5,7	91,0	9,0	81,3	19,7
Общий уровень распознавания, %	94,1	5,9	90,8	9,2	85,1	14,9

Алгоритм АП-ЛБШ-МОВ превосходит алгоритм АП-МОВ по общему уровню распознавания на 3,3% и достигает уровня в 94,1%. При этом уровни верного распознавания мужчин и женщин у алгоритма АП-ЛБШ-МОВ приблизительно равны. Выигрыш по общему уровню распознавания по сравнению с базовым классификатором МОВ составляет 9% и 5,7% для алгоритма АП-ЛБШ-МОВ и алгоритма АП-МОВ соответственно.

4.4 Краткие выводы

Результаты проведенной работы позволяют сделать следующие основные выводы.

- Разработана модификация базового классификатора МОВ на основе ЛБШ признаков. ЛБШ-МОВ показывает уровень верного распознавания равный 89,4% и опережает алгоритм МОВ на 4,3%, однако, уступает классификатору АП-МОВ на 1,4%.

- Разработана модификация базового классификатора МОВ на основе комбинированного пространства адаптивных и ЛБШ признаков. АП-ЛБШ-МОВ показывает уровень верного распознавания равный 94,1% и опережает алгоритмы МОВ и АП-МОВ на 9% и 3,3% соответственно.

ЗАКЛЮЧЕНИЕ

Основные результаты диссертации можно сформулировать в следующем виде.

1. Проведен анализ известных алгоритмов решения задач детектирования лиц людей на телевизионных изображениях и распознавания пола по детектированной области лица, позволяющий оценить эффективность существующих методов.
2. Разработана собственная база, содержащая более 10 тысяч изображений лиц людей, для проведения обучения и тестирования алгоритмов распознавания лиц.
3. Исследована устойчивость трех современных алгоритмов детектирования лиц на телевизионных изображениях к изменению условий освещенности. При изменении яркости уровень верного детектирования алгоритмов на базе МОВ, РСПЭ и бустинга уменьшается в среднем на 17,0%, 10,6% и 2,2% соответственно. При изменении контраста уровень верного детектирования алгоритма на базе МОВ уменьшается в среднем на 25,7%. Неравномерная освещенность приводит к уменьшению средних показателей алгоритма на базе МОВ на 16,9%; алгоритма на базе РСПЭ – на 17,9%; алгоритма на базе бустинга – на 7,5%.
4. Исследована возможность использования алгоритма эквализации гистограммы и алгоритма локальной предобработки для повышения устойчивости тестируемых алгоритмов детектирования лиц к изменению условий освещенности:
 - Результаты показывают целесообразность использования эквализации гистограммы для улучшения работы алгоритма на базе МОВ. Выигрыш по среднему уровню верного детектирования составляет в этом случае 8,1% при изменении яркости изображений и 19,3% при изменении контраста.

– Для алгоритма на базе РСПЭ лучшие результаты достигаются при использовании локальной предобработки. Выигрыш по среднему уровню верного детектирования составляет в этом случае 2% при изменении яркости изображений и 2,1% в условиях неравномерной освещенности.

– Алгоритм на базе бустинга наиболее устойчив к изменению условий освещенности. Удастся улучшить его уровень верного детектирования в среднем на 0,7% при изменении яркости изображений за счет использования алгоритма эквализации гистограммы.

5. Проведено тестирование трех современных алгоритмов классификации пола по изображению лица с использованием предлагаемой базы лиц. Лучший результат демонстрируют алгоритмы, в основе которых лежит классификатор на базе метода опорных векторов. Уровень верного распознавания алгоритма АП-МОВ составил 90,8%; алгоритма МОВ – 85,1%; алгоритма KDDA – 69,7%.
6. Исследована устойчивость базового классификатора на базе МОВ к изменению условий освещенности при распознавании пола человека по изображению лица. Вероятность верного распознавания уменьшается с 85% при нормальных условиях до 55% и 72% при уменьшении и увеличении яркости соответственно, до 66% при уменьшении контраста и до 55% при неравномерном освещении.
7. Разработано три модификации базового алгоритма на базе МОВ с применением одного стандартного и двух оригинальных алгоритмов предобработки изображений. Предлагаемые модификации имеют преимущество по среднему уровню верного распознавания порядка 10-15% по сравнению с базовым алгоритмом МОВ при различных условиях освещенности. Лучшие результаты демонстрирует алгоритм СЛП-МОВ.
8. Разработана модификация базового классификатора МОВ на основе ЛБШ признаков. ЛБШ-МОВ показывает уровень верного распознавания

равный 89,4% и опережает алгоритм МОВ на 4,3%, однако, уступает классификатору АП-МОВ на 1,4%.

9. Разработана модификация базового классификатора МОВ на основе комбинированного пространства адаптивных и ЛБШ признаков. АП-ЛБШ-МОВ показывает уровень верного распознавания равный 94,1% и опережает алгоритмы МОВ и АП-МОВ на 9% и 3,3% соответственно.

СПИСОК ЛИТЕРАТУРЫ

1. Форсайт Д.А., Понс Д. Компьютерное зрение. Современный подход // М.: «Вильямс», 2004. 928 с.
2. Люгер Дж. Ф. Искусственный интеллект: стратегии и методы решения сложных проблем // М.: «Вильямс», 2003. 864 с.
3. Мерков А.Б. Распознавание образов. Введение в методы статистического обучения // М.: Едиториал УРСС, 2011. 256 с.
4. Fukunaga K. Introduction to Statistical Pattern Recognition // Academic Press, Boston, 1990. 592 p.
5. Фу К. Структурные методы распознавания образов // М.: Мир, 1977.
6. Alpaydin E. Introduction to Machine Learning // The MIT Press, 2010.
7. Sammut C., Webb G.I. Encyclopedia of Machine Learning // Springer, 2011.
8. Николенко С. И., Тулупьев А.Л. Самообучающиеся системы // М.: МЦНМО, 2009. 288 с.
9. Quinlan J.R. C4. 5: Programs for Machine Learning // Kluwer Academic, 1993. 302 p.
10. Mitchell T. Machine Learning // McGraw Hill, 1997. 432 p.
11. Уоссермен Ф. Нейрокомпьютерная техника // М.: Мир, 1992. 118 с.
12. Дюк В., Самойленко А. Data Mining: учебный курс // СПб: Питер, 2001. 368 с.
13. Bradski G.R. Computer Vision Face Tracking For Use in a Perceptual User Interface // Proc. IEEE Workshop on Applications of Comp. Vision, Princeton, 1998. P. 214-219.
14. Zhao W., Chellappa R., Phillips P., Rosenfeld A. Face recognition: A literature survey // ACM Computing Surveys (CSUR). 2003. V. 35. № 4. P. 399-458.

15. Конушин А.С. Слежение за точечными особенностями сцены // Компьютерная графика и мультимедиа. Выпуск №1(5). 2003.
16. Lucas B.D., Kanade T. An iterative image registration technique with an application to stereo vision // Proceedings of Imaging Understanding Workshop, 1981. P. 121-130.
17. Борисов Ю.И., Дворкович В.П., Зубарев Ю.Б. Технические проблемы и перспективы внедрения цифрового телевидения и радиовещания в России. Часть 1 // Broadcasting. Телевидение и радиовещание, 2007. №1. С. 22-27.
18. Борисов Ю.И., Дворкович В.П., Зубарев Ю.Б. Технические проблемы и перспективы внедрения цифрового телевидения и радиовещания в России. Часть 2 // Broadcasting. Телевидение и радиовещание, 2007. №2. С. 14-18.
19. Hjelmas E. Face detection: A Survey // Computer vision and image understanding. 2001. V. 83. № 3. P. 236-274.
20. Fasel B., Luettin J. Automatic facial expression analysis: A survey // Pattern Recognition Letters. 2003. V. 36. № 1. P. 259-275.
21. Stan Z. Li, Anil K. Jain. Handbook of Face Recognition // Springer Science+Business Media, 2005. 408 p.
22. Kriegman D., Yang M.H., Ahuja N. Detecting faces in images: A survey. // IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, No 1, pp. 34-58, 2002.
23. Szeliski R. Computer Vision: Algorithms and Applications // Springer, 2010.
24. Yilmaz A., Javed O., Shah M. Object Tracking: A Survey // ACM Comput. Surv, 2006. V. 38 P. 13.
25. Chellappa R., Wilson C., Sirohey S. Human and Machine Recognition of Faces: A Survey // Proc. IEEE, 1995. V. 83 P.705-740.

26. Froba B., Kublbeck. C. Real-time Face Detection Using Edge-orientation Matching // Audio- and Video-based Biometric Person Authentication (AVBPA'2001), Halmstad, Sweden, 2001, P 78-83.
27. De Silva L.C., Aizawa K., Hatori M. Detection and tracking of facial features by using a facial feature model and deformable circular template // IEICE Trans. Inform. Systems EZ8-D(9), 1995. P. 1195-1207.
28. Kienzle W., Bakir G., Franz M., Schölkopf B. Face Detection - Efficient and Rank Deficient // Adv. in Neural Inf. Proc. Systems, V.17, P. 673-680, 2005.
29. Roth D., Yang M.-H., Ahuja N. A SNoW-based face detector // In Advances in Neural Information Processing Systems 12 (NIPS 12), MIT Press, Cambridge, MA, 2000. P. 855-861.
30. Juell P., Marsh R. A hierarchical neural network for human face detection // Pattern Recog. 29, 1996. P. 781-787.
31. Bar-Shalom Y., Foreman T. Tracking and Data Association // Academic Press Inc., 1988. 353 p.
32. Comaniciu D., Ramesh V., Andmeer P. Kernel-based object tracking // IEEE Trans. Patt. Analy. Mach. Intell. 25, 2003. P. 564-575.
33. Yang M.-H., Ahuja N., Kriegman D. Mixtures of Linear Subspaces for Face Detection // Proc. Fourth Int'l Conf. Automatic Face and Gesture Recognition, 2000, P. 70-76.
34. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // Proc. Int. Conf. on Computer Vision and Pattern Recognition. 2001. № 1. P. 511-518.
35. Зубарев Ю.Б., Дворкович В.П., Дворкович А.В. Проблемы и перспективы внедрения информационных мультимедийных систем в России // Электросвязь, 2004. №10. С. 11-16.
36. Ярославский Л.П. Введение в цифровую обработку изображений. – М.: Советское радио, 1979. 312 с.

37. Yaroslavsky L. Digital Picture Processing - An Introduction // Springer Verlag, 1985. 276 p.
38. Yaroslavsky L. Eden M. Fundamentals of Digital Optics // Birkhauser, Boston, 1996. 368 p.
39. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика) // Техносфера, 2012. 1008 с.
40. Цифровая обработка телевизионных и компьютерных изображений / Под ред. Зубарева Ю.Б. и Дворковича В.П. // М.: 1997. 252 с.
41. Чобану М.К. Многомерные многоскоростные системы обработки сигналов // М.: Техносфера, 2009. 480 с.
42. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов // М.: Мир, 1988. 488 с.
43. Прэтт У. Цифровая обработка изображений // М.: Мир, 1982. 311 + 479 с.
44. Ту Дж., Гонсалес Р. Принципы распознавания образов // Под редакцией Ю.И. Журавлева М.: Мир. 1978.
45. Гонсалес Р., Вудс Р. Цифровая обработка изображений // М.: Техносфера, 2005. 621 с.
46. Журавлев Ю.И., Рязанов В.В., Сенько О.В. Распознавание. Математические методы. Программная система. Практические применения. // М.: Фазис, 2005. 159 с.
47. Журавлев Ю.И. Избранные научные труды // М.: Издательство Магистр, 1998. 420 с.
48. Журавлев Ю.И., Никифоров В.В. Алгоритмы распознавания, основанные на вычислении оценок // Кибернетика. 1971. №3. С. 1-11.
49. Гренандер У. Лекции по теории образов // Под ред. Ю. Журавлева М.: Мир, 1979.

50. Методы компьютерной обработки изображений // Под ред. В.А. Сойфера М.: Физматлит, 2001. 784 с.
51. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов (статистические проблемы обучения) // М.: Наука, 1974. 416 с.
52. Загоруйко Н.Г. Прикладные методы анализа данных и знаний // Новосибирск: Изд-во Института математики, 1999. 270 с.
53. Загоруйко Н.Г. Методы распознавания и их применение // М.: Сов.радио, 1972.
54. Потапов А.С. Распознавание образов и машинное восприятие: общий подход на основе принципа минимальной длины описания. // СПб.: Политехника, 2007. 548 с.
55. Лукьяница А.А., Шишкин А.Г. Цифровая обработка видеоизображений // М.: «Ай-Эс-Эс Пресс», 2009. 518 с.
56. Ко, Т. A Survey on behavior analysis in video surveillance for homeland security application // 37th IEEE Applied Imagery Pattern Recognition Workshop, AIPR. P. 1-8, 2008.
57. Онлайн-измерение аудитории Digital Signage (<http://27faces.com/>).
58. Донской В.И. Алгоритмы обучения, основанные на построении решающих деревьев// Журнал выч. мат. и матем. физики. 1982, т. 22, №4, с. 963-974.
59. Sato K., Aggarwal J. Temporal spatio-velocity transform and its application to tracking and interaction // Comput. Vision Image Understand. 96, 2, 2004. P. 100-128.
60. Brodia, T., Chellappa, R. Estimation of object motion parameters from noisy images // IEEE Trans. Patt. Analy. Mach. Intell. 8, 1, 1986. P. 90-99.

61. Rowley H.A., Baluja S., Kanade T. Neural network-based face detection // IEEE Trans. Pattern Anal. Mach. Intell. 20, 1998. P. 23-38.
62. Lin S.-H., Kung S.-Y., Lin L.-J. Face recognition/detection by probabilistic decision-based neural network // IEEE Trans. Neural Networks 8, 1997. P. 114-132.
63. Belhumeur P.N., Hespanha J.P., Kriegman D.J. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection // IEEE Transactions on pattern analysis and machine intelligence, 1997, V. 19, № 7.
64. Шмаглит Л.А., Голубев М.Н. Использование информации о цвете в алгоритме выделения лиц на изображениях // Материалы 20-й Международной Конференции по Компьютерной Графике и Зрению «ГрафиКон'2010», 2010. С. 331-332.
65. Makinen E., Raisamo R. An experimental comparison of gender classification methods // Pattern Recognition Letters 29, 2008. № 10, P. 1544-1556.
66. Viola P., Jones M. Robust real time face detection // International Journal of Computer Vision, Vol. 57, No 2, pp. 137-154, 2004.
67. Freund Y., Schapire R. A decision-theoretic generalization of on-line learning and an application to boosting // Computational Learning Theory. 1995. P. 23-37.
68. Samarth Brahmhatt. Practical OpenCV // Apress, 2013. 223p.
69. Шмаглит Л.А., Голубев М.Н., Игнатов И.С., Хрящев В.В. Выделение лиц на изображениях в условиях искажений // Докл. 11-й междунар. конф. «Цифровая обработка сигналов и ее применение» (DSPA-2009), Москва, 2009. Т.2. С. 511-514.
70. Шмаглит Л.А., Ганин А.Н. Система слежения и гендерной классификации лиц на видеоизображениях // Тез. докл. 18-й междунар. науч.-тех. конф. «Радиоэлектроника, электротехника и энергетика». Москва, 2012. Т. 1. С. 128-129.

71. Streit R.L., Luginbuhl T.E. Maximum likelihood method for probabilistic multi-hypothesis tracking // In Proceedings of the International Society for Optical Engineering (SPIE.), 1994. V. 2235. P. 394-405.
72. Tao H., Sawhney H., Kumar R. Object tracking with bayesian estimation of dynamic layer representations // IEEE Trans. Patt. Analy. Mach. Intell. 24, 1, 2002. P. 75-89.
73. Black M., Jepson A. Eigenttracking: Robust matching and tracking of articulated objects using a view-based representation // Int. J. Comput. Vision, 26, 1, 1998. P. 63-84.
74. Isard M., Blake A. Condensation – conditional density propagation for visual tracking // Int. J. Comput. Vision 29, 1, 1998. P. 5-28.
75. Avidan S. Support vector tracking // In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2001. P. 184-191.
76. Huttenlocher D., Noh J., Rucklidge W. Tracking non rigid objects in complex scenes // In IEEE International Conference on Computer Vision (ICCV), 1993. P. 93-101.
77. Ганин А.Н., Шмаглит Л.А., Храбров Д.Е., Манов И.А. Программно-аппаратный комплекс для детектирования, слежения и гендерной классификации объектов по изображению лиц // Тр. LXVII науч. сессии, посвященной Дню Радио. Москва, 2012. С.163-166.
78. Shi, J., Tomasi, C.: Good features to track. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 593-600, 1994.
79. Tomasi C., Kanade T. Detection and Tracking of Point Features // Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.
80. Khryashchev V., Ganin A., Golubev M., Shmaglit L. Audience analysis system on the basis of face detection, tracking and classification techniques // Proc. International MultiConference of Engineers and Computer Scientists 2013. Hong Kong, 2013. V. 1. P. 446-450.

81. Khryashchev V., Shmaglit L., Golubev M., Shemyakov A. The Development of Object Tracking and Recognition Algorithms for Audience Analysis System // IAENG International Journal of Computer Science, 40:2, pp. 94-103, 2013.
82. Шмаглит Л.А., Хрящев В.В., Ганин А.Н., Матвеев Д.В. Распознавание пола человека по выделенной области лица на изображениях// Проектирование и технологии электронных средств, 2011. №4. Владимир. С. 36-43.
83. Burges C. A Tutorial on Support Vector Machines for Pattern Recognition // Data Mining and Knowledge Discovery, 1998. V. 2, P. 121-167.
84. Beveridge J.R., She K., Draper B.A., Givens G.H. A nonparametric statistical comparison of principal component and linear discriminant subspaces for face recognition // In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. P. 535-542.
85. Голубев М.Н., Ганин А.Н., Шмаглит Л.А. Алгоритм распознавания пола на основе адаптивных признаков и метода опорных векторов // X всероссийская научно-техническая конференция «Нейрокомпьютеры и их применение», Москва, 2012. С. 33.
86. Gao H., Davis J. Why direct LDA is not equivalent to LDA // Pattern Recognition Letters 39, 2006. № 5, P. 1002-1006.
87. Burges C. Simplified support vector decision rules // In International Conference on Machine Learning. 1996. P. 71-77.
88. L. Shmaglit, V. Khryashchev. Gender classification of human face images based on adaptive features and support vector machines. Optical Memory and Neural Networks. Vol. 22. Issue 4. October 2013. pp. 228-235.
89. Phillips P.J., Moon H., Rizvi S.A., Rauss P.J. The FERET evaluation methodology for face recognition algorithms // IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 2000. P. 1090-1104.

90. Fawcett T. ROC Graphs: Notes and Practical Considerations for Researchers // Pattern Recognition Letters. 2004. V. 27. № 8. P. 882-891.
91. Tamura S., Kawai H., Mitsumoto H. Male/female identification from 8 to 6 very low resolution face images by neural network // Pattern Recognition Letters. 1996. V. 29. № 2. P. 331-335.
92. Lyons M., Budynek J., Plante A., Akamatsu S. Classifying facial attributes using a 2-d Gabor wavelet representation and discriminant analysis // Proc. Int. Conf. on Automatic Face and Gesture Recognition. 2000. P. 202-207.
93. Jain A., Huang J. Integrating independent components and linear discriminant analysis for gender classification // Proc. Int. Conf. on Automatic Face and Gesture Recognition. 2004. P. 159-163.
94. Saatci Y., Town C. Cascaded classification of gender and facial expression using active appearance models // Proc. Int. Conf. on Automatic Face and Gesture Recognition. 2006. P. 393-400.
95. Sun Z., Bebis G., Yuan X., Louis S.J. Genetic feature subset selection for gender classification: A comparison study // Proc. IEEE Workshop on Applications of Computer Vision. 2002. P. 165-170.
96. Sun N. et al. Gender classification based on boosting local binary pattern // Proc. Int. Symposium on Neural Networks. 2006. V. 2. P. 194-201.
97. Gutta S., Wechsler H., Phillips P.J. Gender and ethnic classification of face images // Proc. Int. Conf. on Automatic Face and Gesture Recognition. 1998. P. 194-199.
98. Рекомендация МСЭ-R ВТ. 601-7. Студийные параметры кодирования цифрового телевидения для стандартного 4:3 и широкоэкрannого 16:9 форматов. 2011. 17 с.
99. Gomez G., Morales E. Automatic feature construction and a simple rule induction algorithm for skin detection // In Proc. of the ICML Workshop on Machine Learning in Computer Vision, 2002, P. 31-38.

100. Khryashchev V., Priorov A., Shmaglit L., Golubev M. Gender Recognition via Face Area Analysis // Proc. World Congress on Engineering and Computer Science 2012, Berkeley, USA. P. 645-649.
101. Ojala T., Pietikäinen M., Harwood D. A Comparative Study of Texture Measures with Classification Based on Feature Distributions // Pattern Recognition, vol. 29, pp. 51-59, 1996.
102. Шмаглит Л.А., Голубев М.Н., Ганин А.Н., Хрящев В.В. Гендерная классификация по изображению лица // Докл. 14-й междунар. конф. «Цифровая обработка сигналов и ее применение» (DSPА-2012), Москва, 2012. Т. 2. С. 425-428.
103. Голубев М.Н., Шмаглит Л.А., Ганин А.Н. Разработка и анализ системы автоматического распознавания пола людей по изображению лица // XIV Всероссийская научно-техническая конференция «Нейроинформатика-2012», Москва, 2012. С. 151-159.
104. Ганин А.Н., Хрящев В.В., Шемяков А.М., Шмаглит Л.А. Гендерная и возрастная классификация по видеоданным // Докл. 15-й междунар. конф. «Цифровая обработка сигналов и ее применение» (DSPА-2013), Москва, 2013. Т. 2. С. 292-294.
105. Khryashchev V., Shmaglit L., Priorov A., Shemyakov A. Adaptive Feature Extraction for Gender Classification of Human Faces // Proc. Of the 23rd International Conference on Computer Graphics and Vision, GraphiCon'2013, Vladivostok, 2013. P. 71-74.
106. Nilsson M., Dahl M., Claesson I. The successive mean quantization transform // Proceedings of IEEE Int. Conf. ICASSP, V. 4, P. 429-432, 2005.
107. D. Roth. The SNoW Learning Architecture // Technical Report UIUCDCS-R-99-2102, UIUC Computer Science Department, 1999.

108. D. Roth, M.-H. Yang, and N. Ahuja. A SNoW-based face detector // in Advances in Neural Information Processing Systems 12 (NIPS 12), MIT Press, Cambridge, MA, pp. 855-861, 2000.
109. E. Osuna, R. Freund, and F. Girosi. Training support vector machines: an application to face detection // In Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 1997.
110. E. Osuna and F. Girosi. Reducing the run-time complexity in support vector machines // Advances in Kernel Methods—Support Vector Learning, pp. 271–284, Cambridge, MA, 1999. MIT Press.
111. C. J. C. Burges and B. Schölkopf. Improving the accuracy and speed of support vector machines // Advances in Neural Information Processing Systems, volume 9, page 375. MIT Press, 1997.
112. C. J. C. Burges. Simplified support vector decision rules // In International Conference on Machine Learning, pages 71–77, 1996.
113. Шмаглит Л.А., Приоров А.Л., Хрящев В.В., Матвеев Д.В. Детектирование лиц на изображениях в условиях аддитивного белого гауссовского шума // Электромагнитные волны и электронные системы. 2014. Т. 19, № 5. С. 62–70.
114. Аминова Е.А., Никитин А.Е., А.Н., Хрящев В.В., Шмаглит Л.А. Faces.Video.Lab – программа для слежения за лицами на видеопоследовательностях // Свидетельство о государственной регистрации программы для ЭВМ №2011619048 от 21 ноября 2011 г.
115. Pavlov V., Khryashchev V., Pavlov E., Shmaglit L. Application for Video Analysis Based on Machine Learning and Computer Vision Algorithms// Proceedings of the 14th conference of open innovation association FRUCT, Helsinki, Finland. Publisher: SUAI, ISSN 2305-7254, 90-100 pp., 2013.

ПРИЛОЖЕНИЕ 1. АКТЫ ВНЕДРЕНИЯ РЕЗУЛЬТАТОВ РАБОТЫ

ООО «ДиММ»

150030, г. Ярославль, Московский пр-кт, д.123 кв.3

ИНН 7603019903 КПП 760401001,

р/с. №40702810800000000915,

в ИКБР «Яринтербанк» (ООО) г. Ярославль,

кор/сч.30101810300000000728,

БИК 047888728

УТВЕРЖДАЮ

Генеральный директор

ООО «ДиММ»



Личак Д.А.

«14» 2014 г.

АКТ

внедрения результатов диссертационной работы Шмаглита Льва Александровича, выполненной в Ярославском государственном университете имени П.Г. Демидова (ЯрГУ), на тему «Разработка и анализ алгоритмов распознавания лиц на телевизионных изображениях для биометрической идентификации».

Результаты диссертационной работы Шмаглита Л.А. «Разработка и анализ алгоритмов распознавания лиц на телевизионных изображениях для биометрической идентификации» нашли применение в работе ООО «ДиММ». Особый практический интерес представляют следующие результаты диссертации:

1. Результаты сравнительного анализа различных современных алгоритмов детектирования объектов при изменении условий освещенности.
2. Модифицированный алгоритм автоматического распознавания пола людей по изображению лица на основе локальных бинарных шаблонов и метода опорных векторов.

Разработанные Шмаглитом Л.А. алгоритмы классификации пола человека применяются в составе комплекса для анализа аудитории по видеоданным.

Ведущий инженер

A handwritten signature in blue ink, likely belonging to the Lead Engineer mentioned in the text.

Архангельский А.И.

**Общество с ограниченной
ответственностью «АйДата»**
ул. Володарского, д. 103, оф. 214,
г. Ярославль, 150040
тел./факс: (4852)446815
e-mail: idata@ystu.ru
ОКПО 69934420, ОГРН 1117606000632
ИНН 7606080222, КПП 760601001

УТВЕРЖДАЮ

Директор
ООО «АйДата»

Наумов Д.В.

2014 г.



АКТ

внедрения результатов диссертационной работы Шмаглита Льва Александровича, выполненной в Ярославском государственном университете имени П.Г. Демидова (ЯрГУ), на тему «Разработка и анализ алгоритмов распознавания лиц на телевизионных изображениях для биометрической идентификации».

Результаты диссертационной работы Шмаглита Л.А. «Разработка и анализ алгоритмов распознавания лиц на телевизионных изображениях для биометрической идентификации» нашли применение в новых разработках ООО «АйДата».

Разработанные в диссертации алгоритмы локальной предобработки и предобработки с предварительной сегментацией позволяют улучшить характеристики радиотехнических систем видеонаблюдения, в частности систем мониторинга транспортных потоков, работающих в условиях наличия искажений и изменения освещенности.

Научный руководитель

A handwritten signature in blue ink, likely belonging to S.Yu. Boykov, positioned above a horizontal line.

к.т.н. Бойков С.Ю.

"УТВЕРЖДАЮ"

Первый проректор Ярославского
государственного университета
им. П.Г. Демидова

С.А. Кащенко

"09" 

2014 г.



АКТ

внедрения результатов диссертационной работы Шмаглита Льва Александровича на тему «Разработка и анализ алгоритмов распознавания лиц на телевизионных изображениях для биометрической идентификации» в научно-исследовательские работы

Результаты диссертационной работы Шмаглита Л.А., представленной на соискание ученой степени кандидата технических наук по специальности 05.12.04 «Радиотехника, в том числе системы и устройства телевидения», были использованы в научно-исследовательских работах:

- при выполнении НИР «Развитие нелинейной теории цифровой обработки сигналов и изображений в технических системах» в рамках программы «Развитие научного потенциала высшей школы (2010-2012 годы)» (грант РФФИ №10-08-01186) внедрен алгоритм распознавания объектов на основе локальных бинарных шаблонов и метода опорных векторов.

Заведующий кафедрой динамики
электронных систем, профессор, д.т.н.



Ю.А. Брюханов

"УТВЕРЖДАЮ"

Первый проректор Ярославского
государственного университета
им. П.Г. Демидова

С.А. Кащенко

" 10 "

2014 г.

АКТ

внедрения результатов диссертационной работы

Шмаглита Л.А. на тему

**«Разработка и анализ алгоритмов распознавания лиц на телевизионных
изображениях для биометрической идентификации»**

в учебный процесс

Мы, нижеподписавшиеся, заведующий кафедрой динамики электронных систем, профессор, д.т.н. Брюханов Ю.А. и старший преподаватель кафедры динамики электронных систем, к.т.н. Волохов В.А. составили настоящий акт о том, что результаты диссертационной работы Шмаглита Л.А. внедрены в учебный процесс на кафедре динамики электронных систем физического факультета ЯрГУ (специальность «Радиотехника»):

- в курсе «Цифровая обработка изображений» – результаты применения алгоритмов предобработки изображений для повышения уровня детектирования и распознавания объектов в радиотехнических устройствах;
- в курсе «Компьютерное зрение» – результаты разработки алгоритмов классификации объектов на телевизионных изображениях;
- при выполнении студентами курсовых и дипломных проектов используется разработанная Шмаглитом Л.А. программа Faces.Video.Lab для слежения за лицами на видеопоследовательностях.

Заведующий кафедрой динамики
электронных систем, профессор, д.т.н.



Ю.А. Брюханов

Старший преподаватель кафедры динамики
электронных систем, к.т.н.



В.А. Волохов