

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

«Рязанский государственный радиотехнический университет имени В.Ф. Уткина»

*На правах рукописи*



**Черненко Анна Дмитриевна**

**«Алгоритмы построения пространственно-временных дескрипторов  
признаков для обнаружения микролицевых движений»**

Специальность: 2.2.12. Приборы, системы и изделия медицинского назначения

Диссертация на соискание ученой степени  
кандидата технических наук

Научный руководитель:  
доктор технических наук, доцент  
Мельник Ольга Владимировна

Рязань 2025

## ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

3DHOГ	–	Histograms of oriented 3D spatio-temporal gradients, трехмерная гистограмма ориентированного градиента
AAM	–	Active Appearance Model, модель активного внешнего вида
ASM	–	Active Shape Model, модель активной формы
CLM	–	Constraint Local Model, ограниченная локальная модель
DRMF	–	Discriminative Response Map Fitting, карта дискриминационного ответа
FHOFO	–	Fuzzy Histogram of Optical Flow Orientations, нечеткой гистограммы ориентации оптического потока
FHOOF	–	Fuzzy Histogram of Oriented Optical Flow, нечеткой гистограммы направленного оптического потока
HOG	–	Histogram of gradients, гистограмма градиентов
HOOF	–	Histogram of oriented optical flow, гистограмма ориентированного оптического потока
KLT	–	Kanade-Lucas-Tomasi, алгоритм Канада-Лукас-Томас
LBP	–	Local Binary Patterns, локальный бинарный шаблон
LBP-ТОР	–	Local Binary Patterns from Three Orthogonal Planes, локальные бинарные шаблоны по трем ортогональным плоскостям
MFE	–	Micro Facial Expressions, микромимика
MFM	–	Micro Facial Movement, микродвижения лица
MLP	–	Multilayer Perceptron, многослойный перцептрон
РСА	–	principal component analysis, анализ главных компонент
SAMM	–	Spontaneous Micro-Facial Movement Dataset, набор данных спонтанных действий и микродвижений
STIP	–	Spatio-Temporal Interest Point, детектор пространственно-временных точек интереса
SVM	–	Support Vector Machine, машина опорных векторов
ДТ	–	динамические текстуры
ЕД	–	единица действия
СКЛиД	–	система кодирования лицевых движений
МРТ	–	магнитно-резонансная томография

## Содержание

ВВЕДЕНИЕ .....	5
1 Современные области применения и алгоритмы вычисления пространственно-временных дескрипторов признаков .....	12
1.1 Пространственно-временные дескрипторы в задачах распознавания движений по видеопоследовательности изображений .....	12
1.2 Обнаружение микролицевых движений с помощью пространственно-временных дескрипторов признаков .....	17
1.3 Основные этапы обнаружения микролицевых движений .....	25
1.4 Существующие алгоритмы вычисления пространственно-временных дескрипторов признаков .....	30
1.5 Наборы данных спонтанных микровыражений лица.....	41
1.6 Выводы по первой главе.....	49
2 Разработка алгоритмов выбора областей возможного проявления микролицевых движений.....	51
2.1 Алгоритмы поиска лица и глаз на изображении .....	51
2.2 Алгоритм выбора антропометрических точек вокруг глаз .....	56
2.3 Классификация современных методов обнаружения антропометрических точек.....	69
2.4 Подготовка набора векторов признаков для обнаружения микролицевых движений .....	76
2.5 Выводы по второй главе.....	79
3 Модификация алгоритмов пространственно-временных дескрипторов признаков.....	79
3.1 Подходы к выбору алгоритмов для обнаружения микролицевых движений .....	81
3.2 Исследование влияния значений параметров алгоритма LBP-TOP на диагностическую эффективность обнаружения микролицевых движений..	83
3.3 Сравнительный анализ алгоритмов дескрипторов FHOOF и FHOFO ....	98

3.4 Общее сравнение алгоритмов LBP-TOP, FHOOF и FHOFO .....	108
3.5 Выводы по третьей главе .....	112
4 Разработка программного комплекса для экспериментальных исследований пространственно-временных дескрипторов .....	114
4.1 Перспективы диагностики психоэмоционального состояния человека на основе микровыражений лица .....	114
4.2 Основные функции разработанного программного комплекса .....	116
4.3 Результаты экспериментов с применением разработанного программного комплекса .....	117
4.4 Экспериментальные исследования на собственном наборе изображений .....	122
4.5 Выводы по четвертой главе .....	127
ЗАКЛЮЧЕНИЕ .....	128
СПИСОК ЛИТЕРАТУРЫ.....	130
ПРИЛОЖЕНИЕ А .....	146
ПРИЛОЖЕНИЕ Б.....	148

## ВВЕДЕНИЕ

**Актуальность темы исследования.** Объективная оценка текущего психоэмоционального состояния человека является актуальной задачей во многих областях – от медицины, психологии и педагогики до проблем подбора персонала, безопасности и криминалистики.

Один из наиболее популярных подходов к оценке эмоционального состояния – анализ выражений лица человека на основе автоматизированной интерпретации его фото- и видеоизображений. Крупнейший специалист в области психологии эмоций Пол Экман показал, что с точки зрения мимики люди любой культуры одинаково выражают свои эмоции, которые представлены семью базовыми выражениями лица: счастье, грусть, гнев, страх, удивление, отвращение и презрение. На сегодняшний день, с учетом активного прогресса в сферах машинного зрения и искусственного интеллекта, существует множество работ, направленных на обнаружение и идентификацию эмоциональных проявлений в мимике. Современные системы распознавания выражений лица достаточно успешно решают задачи предварительной обработки изображения, извлечения визуальных признаков и классификации эмоций, в том числе в режиме реального времени.

Ключевой проблемой большинства существующих систем распознавания выражений лица является неоднозначная связь между мимикой и реальным эмоциональным состоянием человека. Выражение лица во многих ситуациях социального взаимодействия может оказаться как наигранным, так и демонстративно нейтральным. Путем решения данной проблемы является переход от решения задачи анализа макровыражений лица, которые в зависимости от ситуации могут быть как спонтанными, так и фальшивыми, к задаче обнаружения микролицевых движений – кратковременных и низкоамплитудных движений лицевой мускулатуры, участвующих в формировании так называемых микровыражений.

Микровыражения – это кратковременные выражения лица низкой интенсивности, которые обычно появляются, когда люди намеренно или неосознанно пытаются скрыть свои истинные эмоции. Множество исследований, проведенных в данной области, показали, что человек не может контролировать проявление микровыражений и не способен их подавить. Когда запускается эмоциональный эпизод, возникает импульс, который вызывает сокращение соответствующих лицевых мышц. Как только выражение начинает проявляться, человек пытается остановить его, что вызывает кратковременное изменение лица. Оценка продолжительности микролицевого движения различается у разных исследователей, но общепринято, что микролицевое движение должно длиться не более 200 мс.

Процесс обнаружения микролицевых движений состоит из нескольких этапов, а именно: обнаружение ключевых точек лица, извлечение пространственно-временных признаков, классификация признаков. При этом, с учетом слабой выраженности микролицевых движений, требуются исследование и разработка специализированных алгоритмов для реализации каждого из этапов. В диссертационной работе основное внимание уделяется этапу извлечения пространственно-временных дескрипторов признаков, а также частично предшествующему и последующему этапам. Рассмотрена задача обнаружения микролицевых движений. Используется подход, который заключается в вычислении признаков для антропометрических точек на основе областей, в которых возможно возникновение микролицевых движений.

Исследования в направлении анализа микровыражений представляют существенный интерес из-за возможностей выявления скрытых эмоций человека. Актуальной задачей является создание программно-алгоритмических средств, выявляющих признаки подавляемых эмоций на видеопоследовательности изображений, что может быть полезно как при диагностике, так и в терапии различных нарушений эмоциональной сферы, в частности, фобий, депрессии и тревожного расстройства.

**Степень разработанности темы исследования.** Задача обнаружения характерных точек лица решалась в работах Cootes T.F., Taylor C.J., Cooper D.H., Graham J., Sullivan J., Raveendran K., Amos B. Решением задачи формирования дескрипторов признаков на видеопоследовательности для обнаружения макровыражений лица занимались Gong S., Pietikainen M., Zhao G., Hung X., Fei Z., Valstar M., Batista J., Li X. Вопросы классификации эмоций рассматривались Busso C., Jain J., Movellan J.R., D’Mello S., Graesser A., Mavromoustakos-Blom P., Tan C.T., Bakkes S.

Непосредственно исследованию научно-технических вопросов в области регистрации, распознавания и классификации микровыражений посвящено небольшое число работ отечественных и зарубежных исследователей, что связано с отсутствием подходящих алгоритмов и методов, а также наборов данных для проведения исследований в данной области. Исследованиями алгоритмов для решения задачи автоматизированного анализа занимаются зарубежные ученые Yarp M.H., Davison A., Pietikainen M., Zhao G., Chaudhry R., Klazer A., Pfister T. В отечественной науке исследования в области анализа микровыражений представлены немногочисленными работами Московского государственного университета имени М.В. Ломоносова (Гусев А.Н.) и Рязанского государственного радиотехнического университета имени В.Ф. Уткина (Никифоров М.Б., Мельник О.В., Саблина В.А.).

Следует отметить, что в научной литературе рассматриваются различные способы применения анализа микролицевых движений, при этом вопросы обеспечения приемлемой вычислительной сложности остаются без должного внимания. Кроме того, недостаточно проработана проблема обнаружения микродвижений лица в естественной обстановке.

Таким образом, существует научная проблема, связанная с необходимостью поиска новых подходов и алгоритмов для повышения точности и степени автоматизации систем анализа микровыражений лица. Требуется разработать

улучшенные алгоритмы анализа микролицевых движений с точки зрения качества и скорости решения поставленных задач.

**Цель диссертационной работы** заключается в повышении диагностической эффективности обнаружения микролицевых движений в областях их возможного проявления за счёт использования пространственно-временных дескрипторов признаков, вычисляемых с помощью улучшенных алгоритмов.

Для достижения поставленной цели требуется решить следующие **основные задачи:**

- выбор наборов данных для проведения исследований пространственно-временных дескрипторов признаков для анализа микролицевых движений;
- разработка алгоритмов выбора областей возможного проявления микролицевых движений для последующего построения пространственно-временных дескрипторов признаков;
- теоретические и экспериментальные исследования алгоритмов вычисления пространственно-временных дескрипторов признаков;
- разработка улучшенных алгоритмов пространственно-временных дескрипторов признаков для анализа микровыражений лица;
- обоснование повышения диагностической эффективности обнаружения микролицевых движений за счёт использования пространственно-временных дескрипторов признаков, вычисляемых с помощью разработанных алгоритмов.

**Объектом исследований** являются последовательности изображений с микродвижениями лица человека.

**Предметом исследований** является совершенствование пространственно-временных дескрипторов признаков для анализа микродвижений лица.

**Методы исследования.** В данной диссертационной работе теоретические исследования выполнялись с помощью теории методов машинного обучения, теории классических математических алгоритмов вычисления дескрипторов антропометрических точек. Моделирование и экспериментальные исследования



предлагаемых алгоритмов выполнялись с использованием набора данных спонтанных действий и микродвижений Spontaneous Micro-Facial Movement Dataset (SAMM), соответствующих системе кодирования лицевых движений, фреймворка MediaPipe Face Mesh на основе библиотеки TensorFlow, языков программирования MATLAB и Python.

**Научная новизна** работы определяется разработкой нового алгоритма поиска антропометрических точек вокруг глаз и новых методик подбора параметров алгоритмов вычисления пространственно-временных дескрипторов признаков и модификацией алгоритма вычисления дескриптора локальных бинарных шаблонов по трем ортогональным плоскостям.

**Практическая значимость и научная ценность** результатов заключается в том, что:

– применение разработанного алгоритма обнаружения антропометрических точек вокруг глаз позволяет уменьшить среднеквадратичное отклонение координат обнаруженных точек от искомым по сравнению с алгоритмом модели активной формы;

– применение модифицированного алгоритма и разработанной методики позволяет повысить диагностическую эффективность обнаружения микролицевых движений.

**Соответствие паспорту специальности.** Диссертация соответствует паспорту специальности 2.2.12. Приборы, системы и изделия медицинского назначения (технические науки) по направлениям исследований 4, 14, 19, 22.

**Достоверность и обоснованность** результатов и выводов подтверждается корректным использованием математического аппарата, результатами моделирования на ЭВМ и экспериментальными исследованиями с использованием специализированных аннотированных наборов данных.

**Основные положения,** выносимые на защиту:

1. Алгоритм обнаружения антропометрических точек вокруг глаз, обеспечивающий в 2,2 раза меньшее среднеквадратичное отклонение координат

обнаруженных точек от искомым по сравнению с алгоритмом модели активной формы.

2. Модифицированный алгоритм вычисления дескриптора локальных бинарных шаблонов по трем ортогональным плоскостям и методика подбора его параметров, которые позволили получить диагностическую эффективность обнаружения микролицевых движений, равную 98%, в то время как диагностическая эффективность при стандартной реализации алгоритма составляет 88%.

3. Методика подбора параметров для алгоритмов вычисления дескриптора нечеткой гистограммы направленного оптического потока и дескриптора нечеткой гистограммы ориентаций оптического потока, которая позволила получить диагностическую эффективность обнаружения микролицевых движений, равную 98% и 97% соответственно, в то время как диагностическая эффективность при стандартной реализации алгоритмов составляет 95% и 96% соответственно.

**Внедрение результатов диссертации.** Разработанные программно-алгоритмические средства были использованы в центре практической психологии ФГБОУ ВО РязГМУ Минздрава России при проведении исследований по влиянию звуковой стимуляции на эмоциональную сферу человека. Результаты диссертации внедрены в разработки ООО «Биотехпродуктс» и в учебный процесс ФГБОУ ВО «Рязанский государственный радиотехнический университет имени В.Ф. Уткина» при подготовке студентов, обучающихся по направлению 12.03.04 «Биотехнические системы и технологии».

**Апробация работы.** Основные результаты докладывались и обсуждались на конференциях: II Международный научно-технический форум «Современные технологии в науке и образовании» СТНО-2019, РГРТУ (Рязань, 2019 г.); V НТК магистрантов Рязанского радиотехнического университета, РГРТУ (Рязань, 2019 г.); 8-ая международная Средиземноморская конференция по встроенной вычислительной обработке МЕСО-2019 (Будва, Черногория, 2019 г.); XXXIV

Всероссийской научно-технической конференции студентов, молодых ученых, специалистов БИОМЕДСИСТЕМЫ-2021 (Рязань, 2021 г.); V Международный научно-технический форум «Современные технологии в науке и образовании» (СТНО-2022) (Рязань, 2022 г.); 32-я Международная конференция по компьютерной графике и машинному зрению ГРАФИКОН-2022 (Рязань, 2022 г.); XXXV Всероссийской научно-технической конференции студентов, молодых ученых, специалистов БИОМЕДСИСТЕМЫ-2022 (Рязань, 2022 г.); 78-я Научно-техническая конференция Санкт-Петербургского НТО РЭС им. А.С. Попова, посвященная Дню радио, 26 апреля 2023 года, СПбГЭТУ «ЛЭТИ»; Международная научная конференция Системы и технологии цифровой медицины STDH-23 (Ташкент, Узбекистан, 2023 г.); Десятая международная конференция по когнитивной науке (Пятигорск, 2024 г.); Физика и Радиоэлектроника в Медицине и Экологии ФРЭМЭ-2024 (Суздаль, 2024 г.); XXXVII Всероссийской научно-технической конференции студентов, молодых ученых, специалистов БИОМЕДСИСТЕМЫ-2024 (Рязань, 2024 г.).

**Публикации.** По теме диссертации опубликованы 22 печатные работы, в том числе из них 4 статьи в научных рецензируемых изданиях, рекомендованных ВАК для публикации результатов диссертаций по специальности 2.2.12; 3 публикации, индексируемые в международных реферативных базах Web of Science и (или) Scopus; 13 статей и тезисов докладов внутривузовских, всероссийских и международных конференций, 2 свидетельства о государственной регистрации программы для ЭВМ.

**Структура и объем работы.** Диссертация состоит из введения, 4-х глав, заключения, списка используемой литературы и приложений. Общий объем диссертации составляет 150 страниц, работа содержит 53 рисунка и 19 таблиц. Список использованных источников включает 136 наименований.

# **1 Современные области применения и алгоритмы вычисления пространственно-временных дескрипторов признаков**

## **1.1 Пространственно-временные дескрипторы в задачах распознавания движений по видеопоследовательности изображений**

Коммуникативное поведение человека включает мимику, язык жестов, действия, походку. Анализ коммуникативного поведения с помощью компьютеров является одной из самых активно развивающихся областей исследования вычислительных технологий, связанных с людьми. Интерес вызван широким спектром потенциальных приложений во многих сферах деятельности, таких как перцептивные интерфейсы, визуальное наблюдение, интеллектуальная среда, идентификация людей [1], индексирование и поиск видео [2], взаимодействие человека с компьютером, цифровые развлечения, видео наблюдения [3], уход за пожилыми людьми [4]. Движение может быть легко записано в формате видео с помощью камер, которые являются одним из основных элементов оборудования для получения данных. За счет невысокой стоимости и компактности их удобно использовать повсеместно, что увеличивает сферы применимости.

Анализ коммуникативного поведения человека включает в себя обнаружение, отслеживание и распознавание, способствующие пониманию человеческого поведения на основе последовательностей изображений с участием людей. Благодаря развитию исследований в данной области компьютеры могут распознавать людей, анализировать события, касающиеся людей, предоставлять интеллектуальные, адаптивные услуги [5, 6].

В общем случае деятельность человека представляет собой динамическую текстуру (ДТ) – последовательность изображений движущихся сцен, обладающих определенными свойствами стационарности во времени. Примерами ДТ в природе являются морские волны, огонь, дым, листва, вихрь [7].

Описание и распознавание ДТ используется в видеопоисковых системах, которые привлекают все большее внимание исследователей. Из-за неизвестной

пространственной и временной протяженности ДТ их распознавание представляет собой более сложную задачу по сравнению со статическим случаем [8].

Характеристики алгоритма распознавания ДТ [9]:

- 1) сочетание признаков движения с чертами внешности;
- 2) локальная обработка для улавливания информации о переходе в пространстве и времени, например, изменение состояния огня, постепенно меняющегося от искры до большого пожара;
- 3) определение функций, устойчивых к преобразованию изображения, например вращение;
- 4) нечувствительность к перепадам освещенности;
- 5) вычислительная простота [10].

Движение – это совокупность временных объектов, которые характеризуются пространственно-временными признаками. Важная информация обозначается изменениями во внешнем виде и движении элементов во времени [11], которые представляют собой последовательности простых действий [12], где пространственные и временные характеристики связаны для описания всего события. Для описания действий в видеороликах извлекаются функции, позволяющие представлять их компактным способом [13]. Из-за большого количества пространственной информации в каждом кадре рассматриваемой области обычно применяются двумерные дескрипторы локальных признаков. Однако данные дескрипторы не учитывают вклад времени, что приводит к необходимости использовать пространственно-временные дескрипторы [14].

Во многих подходах [10, 15-18] глобальные пространственно-временные вариации ДТ моделировались как динамическая система. В [15] для распознавания ДТ использовался динамический фрактальный анализ. В [18] для инвариантного к внешнему виду распознавания ДТ использовалась линейная динамическая система. Все подходы подчеркивают временные вариации ДТ.

Многие дескрипторы признаков для 2D-изображений обобщаются на трехмерную пространственно-временную область [19], например, детектор

пространственно-временных точек интереса (Spatio-Temporal Interest Point, STIP) [20], который показал многообещающие результаты [21]. Концепция STIP основана на функции Харриса [22] для учета вариаций в пространственной и временной области с целью нахождения геометрических точек – ориентиров, которые представляют собой точки интереса с относительно высокой дисперсией. Отслеживая ориентиры в детекторах STIP можно записывать форму и характеристики движения на видео. STIP был протестирован на классах действий различной сложности [23] и доказал свою эффективность в отслеживании многих событий. Другими дескрипторами, которые были обобщены на трехмерную пространственно-временную область, являются: трехмерный дескриптор масштабно-инвариантного преобразования признаков (3-dimensional Scale-Invariant Feature Transform descriptor, 3DSIFT) [24], пространственно-временной дескриптор масштабно-инвариантного преобразования признаков (Spatio-Temporal Scale-Invariant Feature Transform descriptor, Spatio-temporal SIFT) [25] и трехмерная гистограмма ориентированных градиентов 3DHOG (Histograms of oriented 3D spatio-temporal gradients) [26].

Дескриптор локального бинарного шаблона (Local Binary Patterns, LBP) — это простой дескриптор локальных признаков, устойчивый к изменениям освещения и ошибкам выравнивания. В [27] для распознавания ДТ использованы дескрипторы объемных локальных бинарных шаблонов (Volume Local Binary Pattern, VLBP) и дескрипторы локальных бинарных шаблонов по трем ортогональным плоскостям (Local Binary Patterns from Three Orthogonal Planes, LBP-TOР). Недостаточное количество элементов для построения гистограмм, вероятность возникновения редких паттернов, зашумленность изображения являются проблемами, связанными с надежностью LBP. Для однородного локального бинарного шаблона (Uniform Local Binary Patterns, ULBP) [28] неоднородные шаблоны считаются редкими и зашумленными, поэтому группируются в один столбец гистограммы. Поскольку LBP чувствителен к шуму квантования, в [29] был предложен локальный тройной шаблон с тремя

состояниями (Local Ternary Patterns, LTP). В работе [30] предложена методика удаления недостоверной информации, содержащейся в дескрипторах LBP для повышения надежности гистограмм. Распознавание ДТ в значительной степени зависит от пространственной информации. Кроме того, все кадры ДТ в целом подобны друг другу. Гистограмма строится путем усреднения во временном направлении по всем кадрам. Временная информация переносится на гистограмму.

Расширением дескриптора LBP на временную область является дескриптор LBP-TOP, который рассматривает пиксель в трехмерном пространстве с пространственными и временными свойствами и вычисляет LBP каждого пикселя в трех ортогональных плоскостях, затем вычисляются три гистограммы, соответствующие трем ортогональным плоскостям. LBP-TOP не только наследует достоинства LBP, такие как нечувствительность к изменениям освещения, поворотам, но также расширяет свое применение для анализа характеристик видео высокой размерности [31,32].

В работе [33] предложен пространственно-временной дескриптор для обнаружения аномального поведения человека. Вместо того, чтобы обнаруживать несколько точек интереса, в STIP рассматриваются пространственно-временные вариации текстуры в видеорядах. Угловой детектор Харриса, используемый в STIP, применяется для формирования пространственно-временной текстурной карты (Spatio-Temporal Texture Map, STTM) представления видео.

В работе [26] представлен пространственно-временной дескриптор, 3DHOOG. Данный дескриптор основан на пространственно-временных трехмерных градиентах, которые имеют низкую вычислительную сложность. Квантование ориентации выполняется с использованием правильных многогранников. Применены интегральные гистограммы для уменьшения сложности вычисления признаков в произвольных пространственных и временных масштабах.

В работе [14] представлен подход к распознаванию человеческой деятельности на видео с помощью построения пространственно-временных

дескрипторов, основанных на моментах Цернике (Pyramidal Zernike Over Time, PZOT). Временная информация собирается из комбинации двумерных откликов Цернике. Для согласования дескрипторов применяются изображения в разных пространственных масштабах. Данные дескрипторы могут применяться в классификаторах, не требующих большой обучающей выборки, что приводит к уменьшению вычислительных затрат на обучение.

В работах [34, 35] представлен процесс распознавания речи по визуальной информации, для чего используются признаки внешнего вида на основе пространственно-временных локальных бинарных шаблонов с учетом движения области рта и временного ряда произношения. Детектор лица и глаз служит для автоматического определения области рта.

Другим подходом к решению задачи распознавания движения на видеопоследовательности является применение оптического потока, одной из реализаций которого является метод вычисления дескриптора гистограммы ориентированного оптического потока (Histogram of oriented optical flow, HOOF). Данный дескриптор не зависит от масштаба и направления движущегося объекта. В HOOF вместо самого оптического потока используется распределение оптического потока. Для каждого кадра вычисляется оптический поток. На его основе рассчитывается гистограмма для ориентаций векторов оптического потока, причем векторы, расположенные зеркально относительно вертикальной оси, рассматриваются одинаково, чтобы не создавать различий между движениями влево и вправо [36].

В [37] представлен дескриптор тензорного движения с использованием информации из дескриптора 3DНОG. В работе рассматривается метод, основанный на объединяющем два дескриптора тензоре агрегации, в котором один содержит полиномиальные коэффициенты для аппроксимации оптического потока, а другой – данные из дескриптора гистограммы градиентов (Histogram of gradients, HOG).



Таким образом, значительная часть прогресса в задачах распознавания движений была достигнута за счет разработки различающих пространственно-временных дескрипторов признаков. Процесс извлечения признаков очень важен, поскольку он позволяет представить изображение/видео в более компактном виде по сравнению с простым рассмотрением пикселей.

Базовый подход к распознаванию действий включает следующие основные этапы:

- 1) выделение и представление низкоуровневых признаков;
- 2) классификация действий высокого уровня.

Для низкоуровневого этапа представления действий путем выделения характерных признаков (краев, углов, информации об интенсивности и ориентации), выбор которых сильно влияет на эффективность высокоуровневой классификации действий [39] наиболее распространенными методами являются: 3DNOG [26], HOOF [38], LBP-TOP [27].

## **1.2 Обнаружение микролицевых движений с помощью пространственно-временных дескрипторов признаков**

Микродвижение – это бессознательное проявление эмоционального состояния, имеющее короткое время проявления и низкую интенсивность действий лицевых мышц [40].

Обнаружение микродвижений лица (Micro Facial Movement, MFM) — это новая и сложная область исследований в компьютерном зрении, основанная на работах психологов, изучающих микромимику (Micro Facial Expressions, MFEs) [41].

Несмотря на трудности, анализ микромимики становится все более популярным в последние годы из-за потенциала практического применения в сферах безопасности и при допросах [42-44], здравоохранении [45], переговорах [46], оценке эффективности обучения [47].

Выражения лица универсальны, а не культурно определены [48]. Экман и соавт. [49-51] описывают 7 выражениями лица: счастье, грусть, гнев, страх,

удивление, отвращение и презрение. При запуске эмоционального эпизода возникает импульс, который может вызвать одно или несколько выражений эмоций. Микровыражение возникает, когда человек пытается скрыть свои истинные эмоции. Как только выражение начинает проявляться, человек пытается остановить его, что вызывает временное изменение лица. Принято, что продолжительность микролицевого движения не более 500 мс [52]. Другие определения длительности показывают, что микровыражения длятся менее 250 мс [53] и менее половины секунды [42]. Экман и Фризен первыми определили микровыражение [54], продолжительность которого составила менее 200 мс. Длительность движения является основным признаком, отличающим микровыражение от макровыражения [55].

В системе кодирования лицевых движений (СКЛиД) выражение лица представлено комбинацией различных визуально наблюдаемых движений лица, также называемых единицами действий (ЕД) [56]. СКЛиД описывает критерии наблюдения и кодирования каждой ЕД, а также описание того, как ЕД проявляются в комбинациях. Несмотря на то, что есть конкретные описания выполнения движений, точные изменения лица меняются от одного человека к другому. На визуальное представление эмоции оказывают влияние структура кости, лицевые мышцы, жировые отложения и морщины. ЕД в руководстве СКЛиД представлены двумя основными группами – действиями верхней и нижней части лица. Затем каждая из групп делится на подгруппы. Подгруппы и ЕД, принадлежащие к каждой из групп, показаны в таблице 1.

Как и все мышцы человеческого тела, лицевая мышца может сокращаться с разной силой. В СКЛиД сила сокращения мышц лица классифицируется как показатель интенсивности и обозначается буквами А, В, С, D и Е. А — едва заметная или самая низкая интенсивность движения, Е — максимально возможная, следовательно, самая высокая интенсивность. Данная классификация относится к началу, вершине и завершению движения, причем вершина является самой высокой интенсивностью во временной последовательности движения [27].

В макровыражении активированные ЕД явно просматриваются из-за значительного изменения геометрии и внешнего вида лица. Микровыражение представляет собой едва уловимые движения, которые легко пропустить, если не следить за его эволюцией во времени. Таким образом, ключом к распознаванию микровыражений является анализ временной динамики. Имеется ряд методов извлечения пространственно-временных признаков, которые кодируют внешний вид, а также временные свойства в последовательности кадров.

Таблица 1 – Группы и подгруппы ЕД СКЛиД

Основные группы СКЛиД	Подгруппы	ЕД
Верхняя часть лица	Брови	4
	Лоб	1, 2
	Веки	5, 6, 7, 43, 45, 46
Нижняя часть лица	Вверх/Вниз	9, 10, 15, 16, 17
	Горизонтальное	14, 20
	Косое	11, 12, 13
	Орбитальное	18, 22, 23, 24, 28
	Смешанное	8+25, 19, 21, 29. 30, 31, 32, 38, 39

Одним из алгоритмов вычисления дескрипторов является алгоритм трехмерных гистограмм ориентированных градиентов как простейшее расширение «традиционных» функций НОГ, после появления которых были разработаны такие алгоритмы вычисления дескрипторов, как локальный бинарный шаблон по трем ортогональным плоскостям, гистограммы ориентированного оптического потока и их разновидности [57].

*Дескриптор трехмерных гистограмм ориентированных градиентов, 3D Histograms of Oriented Gradients (3DHOG)*

Polikovskiy S. и соавт. [59] предложили функцию трехмерного градиента для описания локальной пространственно-временной динамики лица. После сегментации лица на 12 областей в соответствии с СКЛиД [51], каждая область представляла собой самостоятельный мышечный комплекс, таким образом, было

получено 12 отдельных пространственно-временных блоков. Величины проекций градиента вдоль каждого из трех направлений использовались для построения гистограммы признаков по каждой из областей. Авторы предполагают, что в каждом кадре представления выражения лица задействована только одна ЕД, представляющая в СКЛиД один конкретный активированный комплекс лицевых мышц, и ЕД может быть использована в качестве аннотации к изображению. Основным недостатком данной работы заключался в предположении активности только одной ЕД в каждом кадре, что на практике создает серьезное ограничение.

В [61] представлен метод обнаружения микролицевых движений с использованием метода временной разности 3D-гистограммы ориентированных градиентов. На основе СКЛиД определялся шаблон лица, состоящий из 26 областей. Затем вычислялись дескрипторы 3DНОГ для каждой из областей. Чтобы найти тонкие движения лица в локальных областях использовалось расстояние хи-квадрат. Для обнаружения микродвижений выше порогового значения использовался пиковый детектор. Данный метод фокусировался на движении 26 областей лица.

*Дескриптор локальных бинарных шаблонов по трем ортогональным плоскостям, Local Binary Pattern-Three Orthogonal Planes (LBP-TOP)*

LBP — это дескриптор, первоначально предложенный для описания внешнего вида изображений. Ключевая идея заключается в том, что для описания внешнего вида используется относительная яркость соседних пикселей [62-64]. Базовая функция LBP опирается на два настраиваемых параметра:  $R$  - радиус и  $P$  - количество соседних пикселей вокруг центрального пикселя. Равномерная выборка точек  $P$  на окружности с радиусом  $R$  и получение их яркости относительно центрального пикселя позволяет охарактеризовать окрестность  $P$ -битным числом. При распознавании микровыражений для кодирования пространственно-временного шаблона используется дескриптор LBP-TOP [65].

LBP-TOP унаследовал вычислительную эффективность LBP и является расширением его низкоуровневого представления [28], описывающим локальные

изменения текстуры вдоль круговой области с помощью двоичных кодов, которые затем кодируются в гистограмму. LBP-TOP извлекает признаки из локальных пространственно-временных окрестностей в трех плоскостях: пространственной плоскости ( $XU$ ), вертикальной пространственно-временной плоскости ( $YT$ ) и горизонтальной пространственно-временной плоскости ( $XT$ ). Это позволяет LBP-TOP динамически кодировать временные изменения.

Впоследствии, для задачи распознавания микровыражений было предложено несколько вариантов LBP-TOP. Чтобы выделить тонкие изменения лица, Xu F. и соавт. [66] предложили метод описания микролицевого движения с разной степенью детализации, названный Карта динамики лица (Facial Dynamics Map, FDM), в котором для пиксельного выравнивания последовательностей микролицевых движений реализована оценка оптического потока. Также разработана итеративная стратегия для расчета основного направления оптического потока для каждого прямоугольного параллелепипеда, состоящая в делении последовательности выражений в соответствии с выбранной степенью детализации, что позволяет охарактеризовать движения лица, содержащие информацию во временной области.

Pfister T. и соавт. [58] предприняли одну из первых попыток автоматического распознавания микровыражений лица. Метод, в котором LBP-TOP используется для извлечения признаков, оказал большое влияние на многие последующие работы. Сначала для поиска ключевых точек лица была применена 68-точечная модель активной формы (Active Shape Model, ASM) [60]. На основе полученных ключевых точек соотношение деформации между первым кадром лица каждой последовательности и кадром нейтрального лица рассчитывалось с использованием метода локального средневзвешенного значения (Local Weighted Mean, LWM) [67]. Чтобы нормализовать небольшие изменения позы и грубые изменения выражения выполнялось геометрическое преобразование. При наличии различий в количестве кадров между входными последовательностями использовалась модель временной интерполяции (Time interpolation model, TIM),

которая позволяла нормализовать длину последовательности до определенного количества кадров. Затем из последовательности извлекались признаки LBP-TOP. Wang S-J. и соавт. [68] представили последовательность микровыражений и дескрипторы LBP с помощью тензора и выполнили разреженный тензорный канонический корреляционный анализ, чтобы выявить взаимосвязь между самой последовательностью микровыражений и ее дескрипторами LBP.

В [69] представлен локальный бинарный шаблон с шестью точками пересечения (LBP with Six Intersection Points, LBP-SIP), то есть из LBP-TOP рассматриваются только 6 уникальных точек, лежащих на трех пересекающихся линиях трех ортогональных плоскостей. LBP-SIP показал более высокую производительность благодаря уменьшению избыточной информации по сравнению с LBP-TOP.

Локальный бинарный шаблон с основными ортогональными плоскостями (LBP with Mean Orthogonal Planes, LBP-MOP), был построен путем объединения признаков LBP только из трех средних изображений, которые являются результатом временного объединения стеков изображений по трем ортогональным плоскостям. Производительность LBP-MOP была сравнима с LBP-SIP [70].

Между тем, из-за проблемы потери информации при кодировании только в трех ортогональных плоскостях, вычисление LBP было расширено до кубического пространства и разработан дескриптор локальных кубических бинарных шаблонов (Local Cube Binary Patterns, LCBP) для распознавания микровыражений лица [71]. LCBP объединил гистограммы признаков направления, амплитуды и центральной разности для того, чтобы получить пространственно-временные признаки, которые уменьшают размерность вектора признаков при сохранении пространственно-временной информации. В дополнение к вычислению знака различия пикселей, Huang X. и соавт. [72] предложили пространственно-временной локальный бинарный шаблон с интегральной проекцией (Spatiotemporal Local Binary Pattern with an Integral

Projection, STLBP-IP), который расширяет возможности LBP-TOP для решения задачи распознавания микролицевых движений. Он получает горизонтальную и вертикальную проекции, сохраняя при этом атрибуты формы изображения лица, что делает более информативным представление во временной области, но теряет некоторую пространственную информацию. Чтобы решить эту проблему, был предложен дискриминационный пространственно-временной локальный бинарный шаблон с пересмотренной интегральной проекцией (Discriminative Spatiotemporal Local Binary Pattern with Revisited Integral Projection, DISTLBP-RIP), включающий атрибут формы в пространственно-временном представлении [73]. Huang X. и Zhao G. [74] предложили вариант бинарного шаблона, названный пространственно-временным локальным бинарным шаблоном Радона (Spatio-Temporal Local Radon-based Binary Pattern, STRBP), который использует преобразование Радона для получения надежных признаков внешнего вида. Ven X. и соавт. [75] предложили альтернативный двоичный дескриптор под названием шаблоны Хот Вил (Hot Wheel Patterns, HWP) для кодирования отличительных признаков изображений макро- и микровыражений лица.

Шаблоны кодирования существующих методов основаны на захвате временной информации по трем ортогональным плоскостям. Сигналы движения в других направлениях и временная корреляция между соседними кадрами в некоторой степени игнорируются. Подход, основанный на LBP, обеспечивает уточненное описание локальной области. Большие размеры векторов признаков и избыточность представлений ограничивают эффективность распознавания.

*Дескриптор гистограммы ориентированного оптического потока, Histogram of Oriented Optical Flow (HOOF)*

В работе [76] используется дескриптор, основанный на оптическом потоке. Авторы выделяют основное направление движения в видеоряде и рассчитывают средние характеристики оптического потока в лицевых блоках. Вводится функция основного направленного среднего оптического потока (Main Directional Mean Optical Flow, MDMO). MDMO – это функция оптического потока на основе

областей интереса, которая учитывает как локальную статистику (то есть среднее значение векторов оптического потока в ячейке с максимальным количеством областей интереса), так и ее пространственное положение (то есть область интереса, к которой она принадлежит). Ключевая точка лица каждого кадра определяется с помощью модели подбора карты дискриминационного ответа (Discriminative Response Map Fitting, DRMF) [77]. Далее поле оптического потока каждого кадра относительно следующего кадра используется для нахождения матрицы аффинного преобразования, которая корректирует изменение позы. Матрица преобразования сводит к минимуму отличие ориентиров лица в каждом кадре по отношению к первому кадру. Затем авторы рассчитывают среднее значение наиболее похожих векторов движения оптического потока в каждой области как характеристику НООФ движения области [38] и квантуют все векторы направления оптического потока до восьми интервалов, чтобы получить гистограмму направлений.

Liu Y-J. и соавт. [78] предложили функцию разреженного оптического потока (Sparse Main Directional Mean Optical-flow, Sparse MDMO), которая использует метод вычисления метрики расстояния для эффективного выявления базовой структуры. Данная модификация позволяет использовать разреженность точек выборки в пространстве признаков MDMO, которая достигается за счет добавления новых метрик в классическую схему графического разреженного кодирования. Liong S. и соавт. [79] предложили двухвзвешенный ориентированный оптический поток (Bi-Weighted Oriented Optical Flow, Bi-WOOF), который применяет различные алгоритмы для локального и глобального взвешивания дескриптора НООФ. Локально компоненты магнитуды использовались для взвешивания столбцов гистограммы ориентации в каждой области интереса. Полученные локально взвешенные гистограммы затем снова взвешивались (глобально) путем умножения на величину средней оптической деформации каждой области интереса. Большее изменение в движении или деформации пикселя способствует более четкой гистограмме. Вместо того, чтобы



рассматривать целые последовательности изображений, авторы продемонстрировали многообещающую эффективность распознавания, используя только два кадра (начальный и пиковый кадр) вместо всей последовательности, что позволило значительно сократить время обработки данных. Нarry S. L. и Routray A. [57] предложили метод нечеткой гистограммы ориентации оптического потока (Fuzzy Histogram of Optical Flow Orientations, FHOFO) и метод нечеткой гистограммы ориентированного оптического потока (Fuzzy Histogram of Oriented Optical Flow, FHOOF), в которых при вычислении дескрипторов не учитывается роль величины движения. Извлеченные признаки устойчивы к изменению интенсивности микровыражений.

Allaert B. и соавт. [80] определили последовательное движение лица, которое может происходить в нескольких направлениях в одной области лица. Предположение было сделано на основании того факта, что движения лица распространяются постепенно из-за эластичности кожи, поэтому для построения согласованного представления карты оптического потока должны быть извлечены только те направления движения, которые являются когерентными в соседних областях лица.

Поле оптического потока может эффективно передавать временные микроизменения в выражении лица. Поскольку функции на основе оптического потока могут описывать местоположение и геометрию ориентиров лица, они требуют более точного процесса выравнивания лица, что вносит некоторую сложность в процесс распознавания микровыражений. Между тем, чувствительность к изменениям освещения делает функции на основе оптического потока очень восприимчивыми к фоновому шуму, что приводит к увеличению избыточности информации.

### **1.3 Основные этапы обнаружения микролицевых движений**

Анализ микролицевых движений включает в себя две задачи: обнаружение и распознавание. Обнаружение микролицевых движений – это задача автоматического определения временного интервала микролицевого движения в

последовательности видеок кадров; распознавание — задача классификации микровыражений для идентификации эмоции, выраженной микролицевым движением, представленным в видеопоследовательности. Системы автоматического определения эмоционального состояния субъекта разрабатываются для автоматического поиска в потоковом видео временной динамики микролицевых движений. Временная динамика относится к движениям лица, которые включают начало, пик, затухание и нейтральные фазы. На рисунке 1 показан пример видеопоследовательности, изображающей данные фазы.

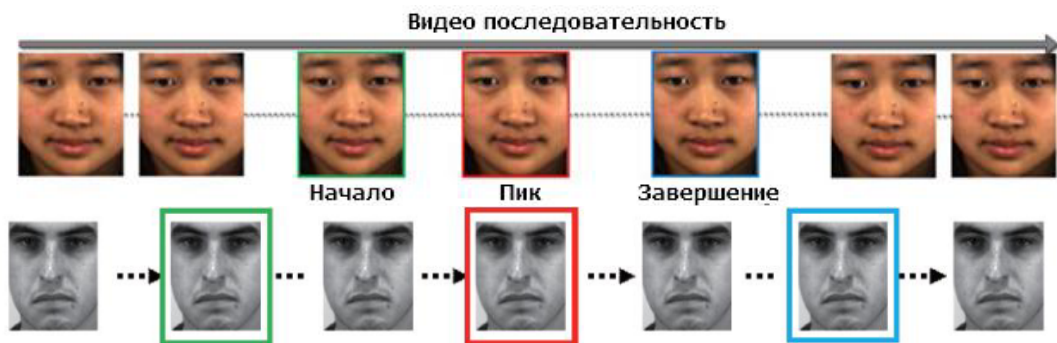


Рисунок 1 – Пример видеопоследовательности изображений, в которой представлена динамика микролицевого движения

Фаза начала — это момент, когда мышцы начинают сокращаться и «деформация» лица усиливается; пик — это момент максимума (вершины) экспрессии; фаза затухания — это момент, когда мышцы расслабляются и лицо возвращается к своей нейтральной форме (минимальная активация лицевых мышц или ее отсутствие) [81].

Процесс обнаружения микролицевых движений состоит из нескольких этапов, а именно: обнаружение антропометрических точек лица, извлечение признаков, классификация признаков [82,83] (рисунок 2).

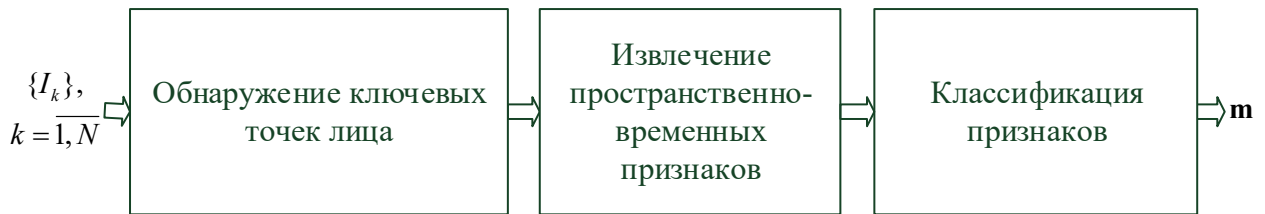


Рисунок 2 – Основные этапы обнаружения микролицевых движений

### *Обнаружение антропометрических точек лица*

Обнаружение антропометрических точек лица — это первый и самый важный этап в задаче определения микролицевых движений по видеопоследовательности. Локализация антропометрических точек на лице может выполняться вручную или методами автоматического определения лицевых ориентиров. В работе [59] антропометрические точки выбирались вручную только на первом кадре последовательности и фиксировались на последующих кадрах, поскольку предполагалось, что исследуемые изображения лиц расположены относительно одинаково на всех кадрах последовательности. В большинстве работ [84-87] применяются методы автоматического определения антропометрических точек лица. Вместо того, чтобы запускать обнаружение ориентиров для всей последовательности изображений лица, точки лица обнаруживаются только на первом кадре и автоматически фиксируются на последующих кадрах с предположением, что данные ориентиры будут изменяться минимально из-за тонкости микролицевых движений.

Методы обнаружения антропометрических точек лица, используемые для выявления микролицевых движений: модель активной формы [88], модель подбора карты дискриминационного ответа DRMF [77], средние сдвиги, ограниченные подпространством (Subspace Constraint Mean Shifts, SCMS) [89], автоматический детектор лицевых точек Face++ [90], ограниченная локальная модель (Constraint Local Model, CLM) [91], детектор лицевых точек на основе модели глубокого обучения MediaPipe Face Mesh [92]. ASM, DRMF и CLM являются примерами деформируемых моделей. Деформируемые модели можно

условно разделить на две основные категории: целостные (генеративные) модели и модели на основе частей (дискриминационные). Первые применяют целостное представление лица на основе текстуры, а последние используют локальные участки изображения вокруг ориентиров. Целостные модели позволяют достичь приемлемого результата, но если целевые изображения лиц не включены в обучающую выборку, то данные модели неверно определяют антропометрические точки лица на изображениях. Поэтому, для определения антропометрических точек лица чаще используют модели, основанные на локальных участках. ASM, DRMF и CLM основаны на локальных деформируемых моделях, однако их механизмы различаются. ASM применяет ограничения формы и выполняет локальный поиск наилучшего местоположения каждой ключевой точки; DRMF изучает изменение внешнего вида в наборе областей, окружающих отдельные точки и соответствующим образом обновляет модель формы; CLM изучает модель изменения формы и текстуры, но текстура выбирается фрагментами вокруг отдельных ключевых точек. Представленные методы устойчивы к «непредвиденным» сценариям, в которых могут существовать большие отклонения от плоскости и окклюзия [93].

Методы, основанные на регрессии, напрямую обучаются отображению внешнего вида изображения в местоположения антропометрических точек и не строят глобальной модели формы. Методы, основанные на регрессии, считаются более перспективными, по сравнению с представленными выше [94].

В работе [90] разработан автоматический точечный детектор лица Face++. Он использует конвейер с нейронной сетью и последовательной регрессией, устойчив к таким воздействиям, как частичные окклюзии и неправильное положение головы с углом наклона до  $90^\circ$ .

В работе [95] для поиска антропометрических точек лица представлен метод каскадной регрессии. Каждая регрессионная функция в каскаде оценивает форму по начальной оценке и интенсивности разреженного набора пикселей, проиндексированных относительно начальной оценки. Каждый регрессор

обучается через градиентный бустинг с функцией потерь. Разреженный набор пикселей, используемый на входе регрессора, выбирается с помощью комбинации алгоритма градиентного бустинга и априорной вероятности расстояния между парами входных пикселей. Каскад таких регрессоров может обнаруживать антропометрические точки при инициализации усредненным положением лица [94].

#### *Вычисление пространственно-временных дескрипторов признаков*

Обзор алгоритмов вычисления пространственно-временных дескрипторов признаков для анализа микровыражений лица представлен в параграфе 1.2.

#### *Классификация признаков*

Последний этап задачи обнаружения микродвижений лица – классификация признаков. На данном этапе преимущественно используются следующие типы классификаторов: k-ближайших соседей (k-nearest neighbour, k-NN), машина опорных векторов (support vector machine, SVM), случайный лес (random forest, RF), классификатор разреженного представления (Sparse Representations Classification, SRC), упрощенное k-сингулярное разложение (simplified K-singular value decomposition, K-SVD), обучение разреженной группы (Group Sparse Learning, GSL) и машина экстремального обучения (Extreme Learning Machine, ELM). Наиболее широко используемым классификатором является SVM. SVM представляет собой алгоритм, который строит гиперплоскость или набор гиперплоскостей в многомерном или бесконечном пространстве. При обучении SVM стремится к максимальному разбросу между границами разных классов. По сравнению с другими классификаторами в случаях, когда количество обучающих выборок невелико, SVM надежен, точен и эффективен. Два других классификатора — RF и k-NN редко используются в задаче обнаружения микродвижений лица. Несмотря на то, что RF быстрее, чем SVM, он склонен к переобучению при работе с зашумленными данными. Метод k-NN использует процесс обучения, который может не подходить для разреженных многомерных данных, таких как данные лиц [93].

Чтобы справиться с разреженностью микролицевых движений были разработаны упрощенные методы K-SVD, SRC и GSL. Однако каждый из методов по-разному решает проблему разреженности. K-SVD [96] путем изучения разреженного словаря с целью различения микролицевых движений минимизирует дисперсию разреженных коэффициентов. SRC [97], использованный в работе [98], представляет тестовую выборку как разреженную линейную комбинацию всех обучающих выборок, следовательно, разреженные ненулевые коэффициенты представления, будут сосредоточены на обучающих выборках того же класса, что и тестовая выборка. Kernelized GSL (kernelized group sparse learning) [99] облегчает процесс изучения набора весов важности иерархических пространственно-временных дескрипторов.

Нейронные сети способны выполнить за один этап извлечение признаков и классификацию, таким образом выявить сложные закономерности из данных. Однако для правильного обучения нейронной сети требуется значительный объем размеченных данных, в результате чего она, по сравнению с представленными выше методами, менее подходит для решения задачи обнаружения микролицевых движений, поскольку количество размеченных данных ограничено [93].

#### **1.4 Существующие алгоритмы вычисления пространственно-временных дескрипторов признаков**

Вычисление пространственно-временных дескрипторов признаков играет важную роль при решении задачи обнаружения микровыражений лица. Необходимо, чтобы дескриптор признаков был способен улавливать микро-различия, имеющиеся в пространственном и временном измерениях.

Наиболее распространенными пространственно-временными дескрипторами признаков, применяемыми в области анализа микровыражений лица, являются: дескриптор локальных бинарных шаблонов по трем ортогональным плоскостям [34], дескриптор гистограммы ориентированного оптического потока [38], дескриптор трехмерных гистограмм ориентированных градиентов [59].

Краткий обзор перечисленных выше пространственно-временных дескрипторов признаков представлен в параграфе 1.2. Далее указанные дескрипторы рассматриваются более подробно.

*Дескриптор локальных бинарных шаблонов по трем ортогональным плоскостям LBP-TOP*

Алгоритм LBP-TOP является расширением алгоритма LBP, то есть для статических изображений LBP расширяется до пространственно-временной области. В LBP-TOP радиусы в пространственных и временных осях  $X$ ,  $Y$ ,  $T$  и количество соседних точек в плоскостях  $XY$ ,  $XT$  и  $YT$  могут быть различными, что обозначается как:  $R_X$ ,  $R_Y$ ,  $R_T$ ,  $P_{XY}$ ,  $P_{XT}$ ,  $P_{YT}$ , дескриптор LBP-TOP представлен как  $LBP - TOP_{P_{XY}, P_{XT}, P_{YT}, R_X, R_Y, R_T}$ .

Дескрипторы LBP-TOP последовательности изображений вычисляются путем объединения локальных бинарных шаблонов по трем ортогональным плоскостям  $XY$ ,  $XT$  и  $YT$ , принимая во внимание только статистику совпадений в трех направлениях. На рисунке 3а представлена последовательность изображений движения рта; 3б - изображение в плоскости  $XY$ ; 3в - изображение в плоскости  $XT$ , дающее визуальное представление об изменении во времени одного ряда; 3г - движение одного столбца во временном пространстве. Дескриптор LBP, вычисленный для всей последовательности изображений, кодирует только вхождение микршаблонов без каких-либо указаний на их местонахождение.

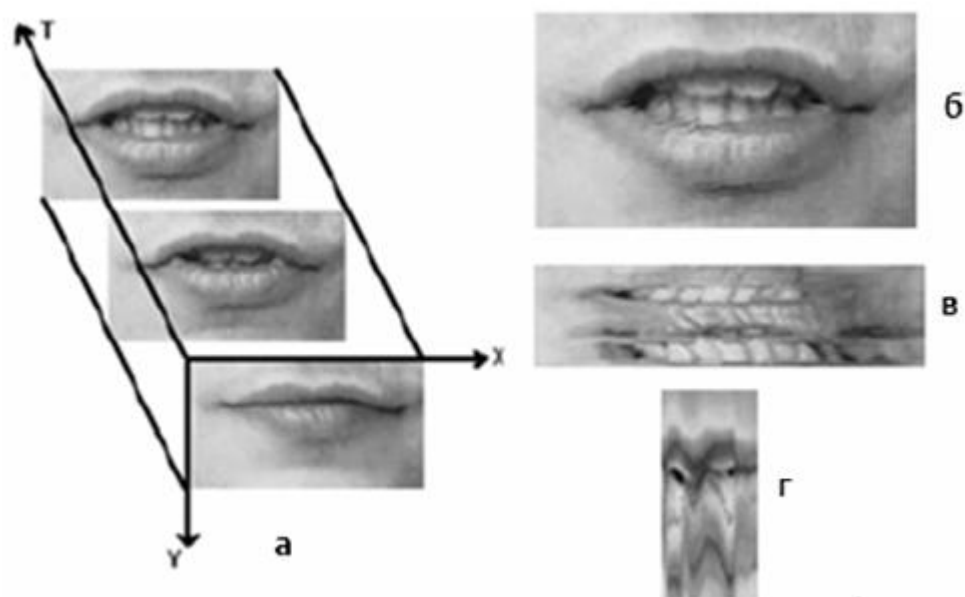


Рисунок 3 – Процесс вычисления дескриптора LBP-TOP:

- а – последовательность изображений локальной области лица (рта),  
 б – изображение в плоскости  $XY$ , в – изображение в плоскости  $XT$ ,  
 г – изображение в плоскости  $YT$

Чтобы решить данную задачу, вводится представление, состоящее из разделения изображения на несколько перекрывающихся блоков. На рисунке 4 приведены примеры изображений вычисленного LBP, на рисунке 4а представлено исходное изображение рта, на рисунках 4б, 4в, 4г показаны изображения LBP, которые получены путем вычисления LBP для каждого пикселя из плоскостей  $XY$  (рисунок 4б),  $XT$  (рисунок 4в) и  $YT$  (рисунок 4г), что соответствует изображениям области рта в первой строке. Таким образом, на рисунке 4 видно изменение внешнего вида и движение рта.

Однако помимо расположения микрошаблонов необходимо учесть временной порядок, иначе микродвижения приобретут почти одинаковые черты. Поэтому, вся последовательность изображений делится на блоки не только в соответствии с пространственными областями, но и в соответствии с временной областью, как показано на рисунке 5.



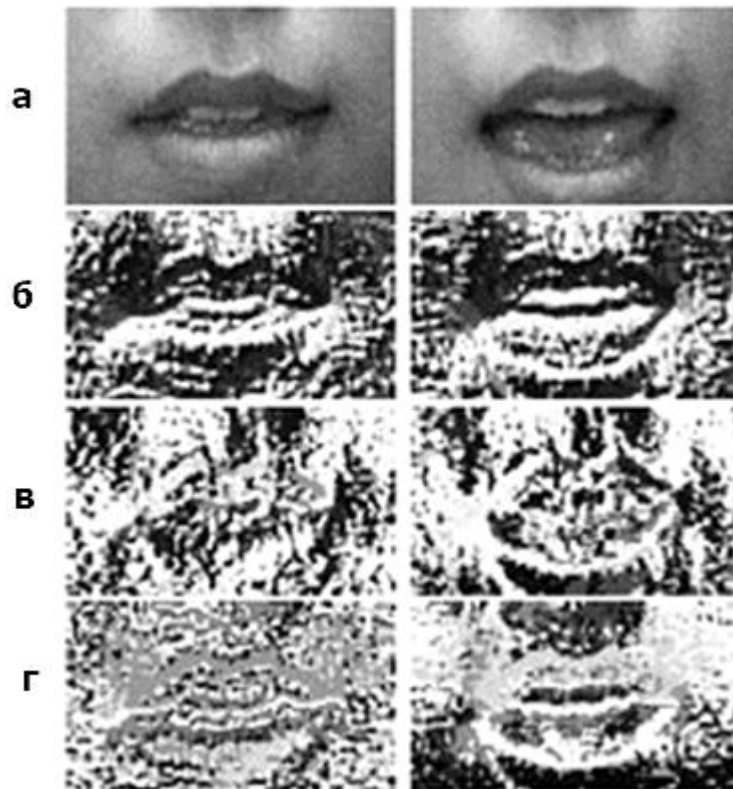


Рисунок 4 – Изображения LBP: а – изображения области рта,  
 б – LBP-XY изображения, в – LBP-XT изображения,  
 г – LBP-YT изображения

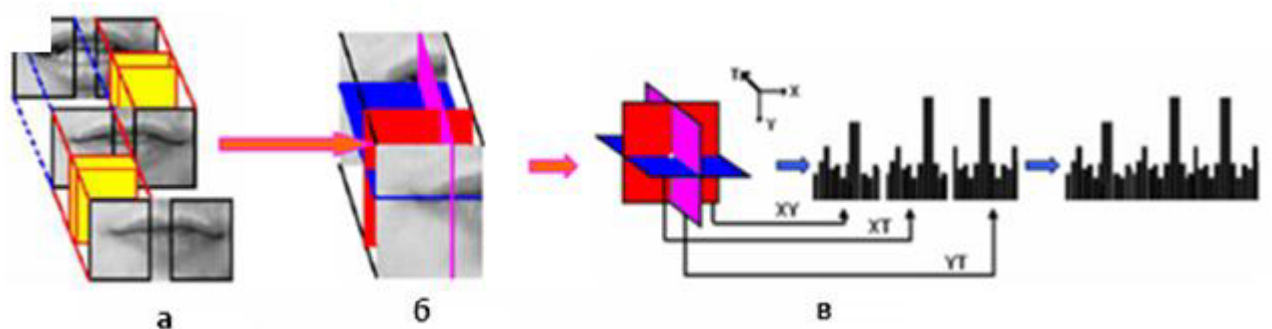


Рисунок 5 – Процесс вычисления гистограммы LBP-TOP: а – блоки внешнего вида и движения, б – LBP дескрипторы по трем ортогональным плоскостям, в – объединенные дескрипторы LBP одного из блоков

В каждом блоке вычисляются гистограммы LBP-TOP, которые затем объединяются в одну гистограмму. Все дескрипторы, извлеченные из каждого

блока, содержат информацию о внешнем виде и движении последовательности областей (рисунок 6).

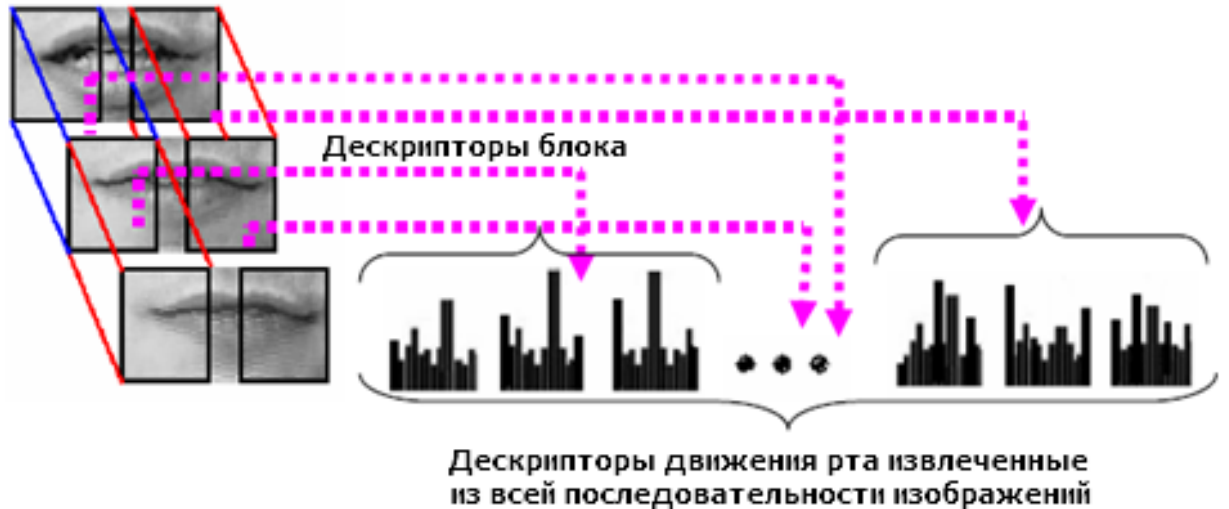


Рисунок 6 – Представление движения рта

Метки гистограммы содержат информацию, описывающую внешний вид и временную информацию на уровне пикселей, по трем ортогональным плоскостям. Метки суммируются в небольшие блоки для получения информации о внешнем виде и движении в определенных местах и временных сегментах на региональном уровне, затем вся информация с регионального уровня объединяется для создания глобального описания движения области [7].

Гистограмму движения области рта можно представить формулой (1):

$$H_{r,c,d,j,i} = \sum_{x,y,t} I\{f_j(x,y,t) = i\}, i = 0, \dots, n_j - 1; j = 0, 1, 2, \quad (1)$$

где  $r$  – индекс строк,  $c$  – индекс колонок,  $d$  – время,  $n_j$  – количество меток, созданных дескриптором LBP в  $j$ -й плоскости ( $j=0: XY, 1: XT, 2: YT$ ),  $f_j(x,y,t)$  – LBP код центрального пикселя с координатами  $(x,y,t)$  в  $j$ -ой плоскости

$$\text{и } I\{A\} = \begin{cases} 1, \text{ если } A \text{ – истина;} \\ 0, \text{ если } A \text{ – ложь.} \end{cases} \quad [100].$$

Нормализация гистограмм выполняется по формуле (2):

$$N_{r,c,d,j,i} = \frac{H_{r,c,d,j,i}}{\sum_{k=0}^{n_j-1} H_{r,c,d,j,k}} . \quad (2)$$

*Дескриптор гистограммы ориентированного оптического потока НООФ*

Дескрипторы НООФ [38] представляют собой гистограммы направлений движения, взвешенные по величине движения. Вектор оптического потока  $v = [v_x, v_y]^T$  вычисляется для каждого местоположения пикселя изображения. Величина движения выражается формулой (3):

$$\rho = \sqrt{v_x^2 + v_y^2} \quad (3)$$

и направление движения выражается формулой (4):

$$\theta = \tan^{-1}\left(\frac{y}{x}\right), \quad (4)$$

заданное в диапазоне:  $-\frac{\pi}{2} + \pi \frac{b-1}{B} \leq \theta \leq -\frac{\pi}{2} + \pi \frac{b}{B}$ , где  $b$  – количество интервалов от общего количества интервалов  $B$ ,  $1 \leq b \leq B$ , вычисляется на основе векторов оптического потока путем преобразования евклидовых координат в полярные координаты [100].

Таким образом, для каждого пикселя имеется пара  $(\rho_i, \theta_i)$ , состоящая из величины и направления движения. Затем все векторы потока группируются в соответствии с их углом поворота и взвешиваются по величине, после чего выполняется нормализация гистограммы, благодаря которой представление гистограммы инвариантно к масштабированию.

НООФ изначально был предложен для распознавания действий человека. Чтобы сделать дескрипторы НООФ независимыми от направления движения, номера интервалов для формирования столбцов гистограмм располагались симметрично вдоль вертикальной оси. Количество интервалов является настраиваемым параметром. Для распознавания микролицевых движений необходимо вычисление гистограмм по всем направлениям. Исследования, проведенные ранее, показывают, что для получения положительных результатов распознавания движений необходимо не менее 30 интервалов.

На рисунке 7 представлен процесс формирования гистограммы с 8 интервалами.

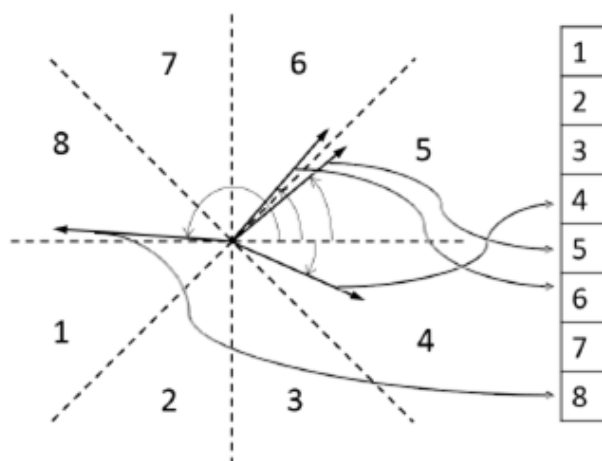


Рисунок 7 – Формирование гистограммы с 8 интервалами

Для области изображения  $I$  размером  $N \times M$  дескрипторы НООФ представлены формулой (5):

$$H(I) = [h_1, h_2, \dots, h_n],$$

$$h_i = \frac{h'_i}{\sum_{k=1}^B h'_k}, h'_i = \sum_{\forall \theta_j \in (i^{\text{th}} \text{ bin})} \rho_j, \quad (5)$$

где  $B$  – количество интервалов гистограммы,  $\rho_j$  – величина движения,  $\theta_j$  – направление движения в местоположении  $j$ -го пикселя [101].

Для различения действий необходимо анализировать динамику временных рядов дескрипторов НООФ, определенных для каждого изображения последовательности. Временной ряд гистограмм выражается формулой (6):

$$h_t = [h_{t,1}, h_{t,2}, \dots, h_{t,B}]^T. \quad (6)$$

Формирование временного ряда, описывающего модель движения, заключается в объединении дескрипторов НООФ для каждой области каждого изображения. При вычислении дескрипторов НООФ отсутствует необходимость в сегментации субъекта.

Метрикой в пространстве гистограмм, называемой минимальной разностью попарных значений, является сумма расстояний между точками, из которых построены гистограммы. Если гистограммы формируются по линейно изменяющимся данным, то расстояние вычисляется по формуле (7):

$$d_{MDPA}(h_1, h_2) = \sum_{i=1}^B \left| \sum_{j=1}^i (h_{1,i} - h_{2,j}) \right|. \quad (7)$$

Другим расстоянием между двумя гистограммами является хи-расстояние, которое определяется формулой (8):

$$d_{\chi^2}(h_1, h_2) = \frac{1}{2} \sum_{i=1}^B \frac{|h_{1,i} - h_{2,i}|}{h_{1,i} + h_{2,i}}. \quad (8)$$

Для оценки подобия расстояний используется ядро радиальной базисной функции. Так как дескрипторы НООФ не являются евклидовыми, для этапа классификации необходима предварительная обработка.

Описанный выше процесс вычисления дескрипторов HOOF представлен на рисунке 8.

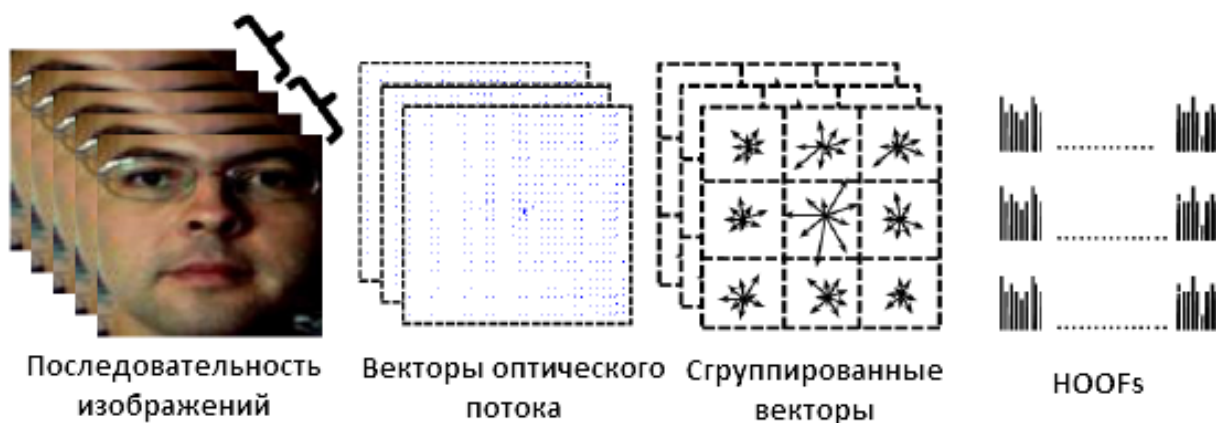


Рисунок 8 – Процесс вычисления дескрипторов HOOF

### *Дескриптор трехмерных гистограмм ориентированных градиентов 3DHOG*

Дескриптор 3DHOG применяется к выделенным областям, соответствующим системе кодирования лицевых движений СКЛид. Для каждой выделенной области лица создаются трехмерные блоки, представляющие собой выбранную последовательность изображений выделенной области. Процесс вычисления дескрипторов лицевых кубов представлен на рисунке 9 [100].

Первым шагом является вычисление первой производной по трем направлениям  $x, y, z$ , которые обозначаются как  $\delta v_x(x, y, t)$ ,  $\delta v_y(x, y, t)$  и  $\delta v_t(x, y, t)$ .

Пары производных содержат в себе информацию:

$\delta v_{xy} = (\delta v_x, \delta v_x)$  – форма поверхности,  $\delta v_{yt} = (\delta v_y, \delta v_t)$  – вертикальное движение,  $\delta v_{xt} = (\delta v_x, \delta v_t)$  – горизонтальное движение от изображения к изображению.

После чего выполняется процесс вычисления гистограмм ориентированных градиентов на основе вычисленных кубов  $\delta v_{xy}$ ,  $\delta v_{yt}$ ,  $\delta v_{xt}$  [59].

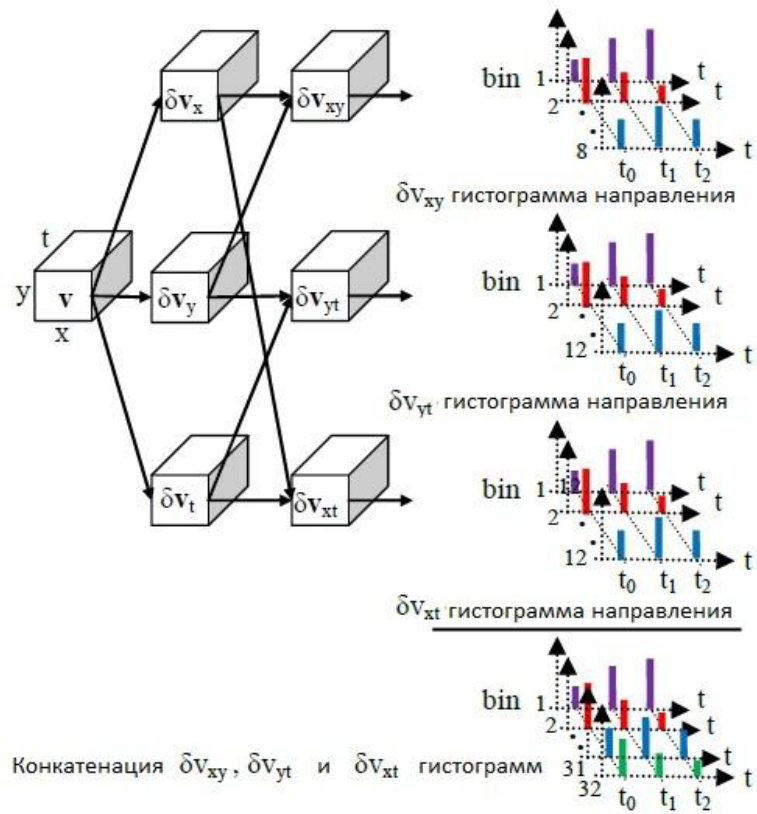


Рисунок 9 – Диаграмма вычисления дескриптора лицевого куба

3DHOOG представляет собой дескриптор, основанный на гистограмме, поэтому он идеально подходит для оценки изменений между кадрами. Концепция анализа различий признаков заключается в том, что для каждого текущего обрабатываемого кадра выполняется сравнение с признаками среднего кадра, который представлен усредненными признаками начального кадра,  $k$ -го кадра перед текущим кадром и конечного кадра,  $k$ -го кадра после текущего кадра.

$k$ -кадр описывается формулой (9):

$$k = \frac{1}{2}(N - 1), \quad (9)$$

где  $N$  — значение микроинтервала, которое обычно устанавливается равным нечетному числу.

Вычисляемую разницу также можно описать как различие между гистограммами в каждой области. Расстояние хи-квадрат  $\chi^2$  определяется формулой (10):

$$\chi^2(P, Q) = \sum_b^B \frac{(P_b - Q_b)^2}{(P_b + Q_b)}, \quad (10)$$

где  $b$  — это  $b$ -я ячейка на гистограммах  $P$  и  $Q$ , которые имеют одинаковое количество ячеек для общего количества ячеек  $B$ .

Уравнение (10) используется во всех временных плоскостях ( $XY$ ,  $XT$ ,  $YT$ ) для расчета различий между гистограммами. Однако, разница не должна зависеть от масштаба значений интервала, то есть каждый признак должен считаться одинаково важным.

Так как дескриптор 3DНОГ применяется к выделенным областям, поэтому более крупные области имеют больше пикселей в области и, следовательно, различий, которые не отражают того, какие области более значимы.

Нормализация определяется формулой (11):

$$F_i = \frac{F_i}{A_r}, \quad (11)$$

где  $A_r$  — количество пикселей в каждой отдельной области,  $i$  — индекс значения в признаке  $F$ .

Данный шаг выполняется для каждой области и необходим, чтобы убедиться, что площади областей, которые не были прямоугольными, верно рассчитаны для нормализации.

Каждое новое значение  $i$ -го признака области  $R$  вычисляется по формуле (12):

$$F'_{R,i} = F_{R,i} - \frac{1}{2}(F_{R,i+k} + F_{R,i-k}). \quad (12)$$



Уравнение (12) вычисляется для каждого кадра за исключением первого и последнего  $k$ -кадра видеопоследовательности из-за временных границ.

Любые отрицательные значения в  $F'_{R,i}$  обнуляются, как любые значения меньше нуля указывают на то, что значение в текущем кадре было меньше среднего значения между кадрами  $(i + k)$  и  $(i - k)$ . Следовательно, в текущем кадре нет быстрых изменений [61].

Дескриптор 3DHOГ показывает наилучшие результаты при анализе наличия движений в выделенных областях, что делает его одним из подходящих алгоритмов для решения задачи оценки наличия микролицевых движений по видеопоследовательности.

### **1.5 Наборы данных спонтанных микровыражений лица**

Разработка наборов данных спонтанных микровыражений является одной из самых сложных задач в области исследования микролицевых движений. Наиболее широко используемыми наборами данных, связанными с микровыражениями, являются *USF-HD* [102], набор данных Поликовского [59], набор данных обнаружения лжи (York Deception Detection Test, YorkDDT) [103], набор микровыражений Академии наук Китая (Chinese Academy of Sciences Micro-Expressions, CASME) [104], набор спонтанных микровыражений (Spontaneous Micro-expression Corpus, SMIC) [105], набор микровыражений II Академии наук Китая (Chinese Academy of Sciences Micro-Expressions II, CASME II) [106], набор спонтанных действий и микродвижений (Spontaneous Actions and Micro-Movement, SAMM) [107] и набор спонтанных макровыражений и микровыражений Академии наук Китая (A Dataset of Spontaneous Macro-Expressions and Micro-Expressions, CAS(ME)<sup>2</sup>) [108].

Набор *USF-HD* используется для оценки методов, направленных на различие макро- и микровыражений. Набор данных Поликовского создан для оценки обнаружения ключевых кадров в контексте микровыражения. YorkDDT предназначен для обнаружения лжи. Данные наборы содержат видеофрагменты постановочных микровыражений, которые не отражают естественное поведение

лицевых мышц, поэтому они не используются при проведении исследований распознавания микролицевых движений.

### *Набор данных SMIC*

Набор данных *SMIC* (рисунок 10) [106] содержит видеопоследовательности 20 участников (6 мужчин и 14 женщин), демонстрирующих 164 спонтанных микролицевых движения. Отличительной чертой *SMIC* от других наборов данных является наличие нескольких модальностей визуализации. Первая часть набора данных *SMIC-HS* содержит видеоролики, снятые в видимом спектре со скоростью 100 кадров в секунду. Вторая часть – *SMIC-VIS* содержит видеопоследовательности, снятые в видимом спектре, но с меньшей частотой кадров — 25 кадров в секунду. Наконец, видеопоследовательности *SMIC-NIR*, снятые в ближнем инфракрасном (*NIR*) спектре.

Принятая продолжительность микровыражения в *SMIC* составляет 1/2 секунды, разрешение кадров – 640x480 пикселей. Набор данных *SMIC-HS* содержит 164 видеопоследовательности с микровыражениями от 16 участников, в то время, как наборы данных *SMIC-VIS* и *SMIC-NIR* содержат 71 видеопоследовательность от 8 участников. Эмоции в *SMIC* разделены на 3 класса (положительные, отрицательные и удивление). Частота каждой эмоциональной категории набора данных *SMIC* представлена в таблице 2.

Таблица 2 – Частота появления каждой категории эмоции набора данных *SMIC*

Набор данных	Положительные	Отрицательные	Удивление	Общее количество видеопоследовательностей
<i>SMIC-HS</i>	51	70	43	164
<i>SMIC-VIS</i>	28	23	20	71
<i>SMIC-NIR</i>	28	23	20	71

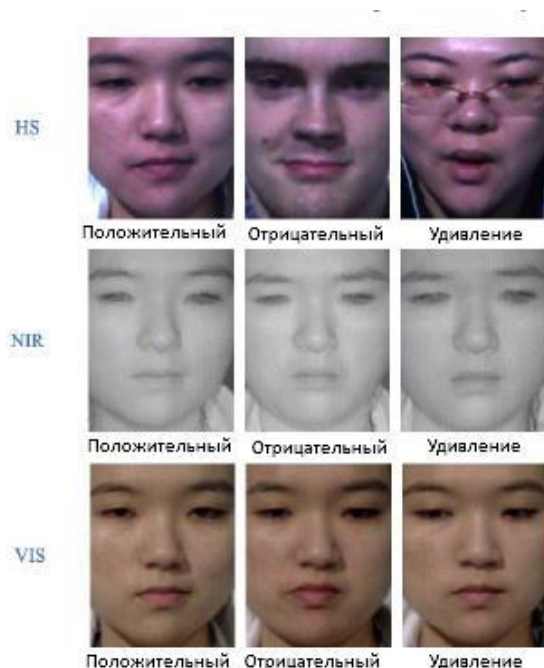


Рисунок 10 – Примеры изображений из набора данных SMIC, а именно: SMIC-HS, SMIC-NIR, SMIC-VIS

#### *Набор данных CASME*

Набор данных CASME (рисунок 11) [109] содержит 195 видеопоследовательностей спонтанно проявляющихся микровыражений, полученных при скорости съемки 60 кадров в секунду. При разработке набора данных принимали участие 35 участников (13 женщин, 22 мужчины). Видеопоследовательности были разделены на два класса в зависимости от условий окружающей среды и используемых камер, а именно: А и В. Изображения класса А были получены в помещении с двумя наклонно расположенными светодиодными лампами, используемыми для освещения лиц, разрешение изображений составляет 640x480 пикселей. Изображения класса В были получены при естественном освещении и имеют разрешение 1280x720 пикселей. CASME содержит микролицевые движения таких эмоций, как радость, грусть, отвращение, удивление, презрение, страх, напряжение. Количество видеопоследовательностей, полученное по каждой эмоциональной категории представлено в таблице 3. Учитывая, что некоторые эмоции труднее возбудить в

лабораторных условиях, чем другие, количество примеров в вышеупомянутых классах распределено неравномерно.

Таблица 3 – Частота появления каждой категории эмоции набора данных CASME

Класс эмоции	Кол-во видеопоследовательностей
Радость	5
Грусть	6
Отвращение	88
Удивление	20
Презрение	3
Страх	2
Подавление	40
Напряжение	28

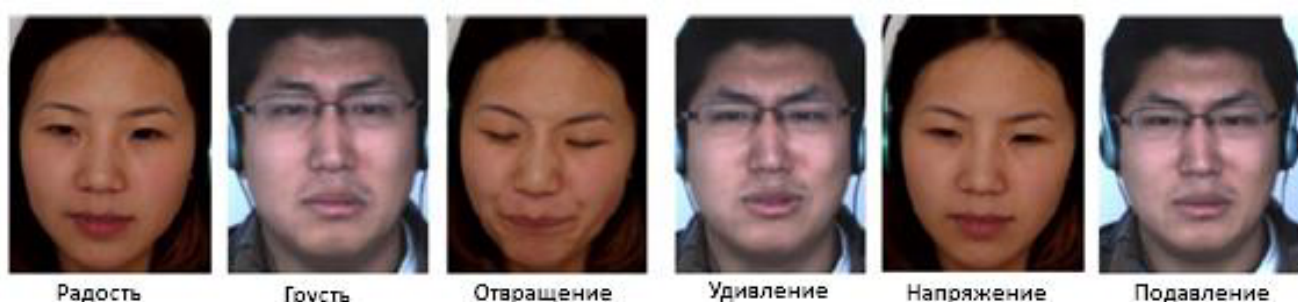


Рисунок 11 – Примеры изображений из набора данных CASME

### *Набор данных CASME II*

Набор данных CASME II (рисунок 12) [106] представляет собой коллекцию спонтанных микровыражений, содержащую 247 видеопоследовательностей 26 участников со средним возрастом около 22 лет. Видеопоследовательности были отобраны из почти 3000 видеороликов и помечены единицами действий на основе СКЛид. Данные были получены при равномерном освещении, без стробоскопа. Скорость съемки 200 кадров в секунду, разрешение – 280×340 пикселей. До и после каждого микровыражения сохраняется несколько кадров, но количество кадров может варьироваться в зависимости от клипа. CASME II содержит пять категорий эмоций: радость, отвращение, подавление,

удивление, печаль и «другие». Частота каждой эмоциональной категории представлена в таблице 4.

Таблица 4 – Частота появления каждой категории эмоции набора данных CASME II

Класс эмоции	Кол-во видеопоследовательностей
Радость	33
Отвращение	60
Удивление	25
Подавление	27
Другое	102



Рисунок 12 – Примеры изображений из набора данных CASME II

### *Набор данных CAS(ME)<sup>2</sup>*

Набор данных CAS(ME)<sup>2</sup> [108] состоит из двух классов: класс А содержит 87 видеороликов, в которых представлены макро- и микровыражения. Второй класс В состоит из 303 отдельных коротких видеороликов, каждый из которых длится до тех пор, пока демонстрируется выражение (макро- или микровыражение). Количество видео с макровыражениями – 250, микровыражениями - 53. Выражения классифицируются как положительные, отрицательные, удивленные и «другие». Скорость съемки 30 кадров в секунду, разрешение – 640x480 пикселей. Частота проявления каждой эмоциональной категории представлена в таблице 5.

Таблица 5 – Частота появления каждой категории эмоции набора данных CAS(МЕ)<sup>2</sup>

Класс эмоции	Количество видео-последовательностей с макровыражениями	Количество видео-последовательностей с микровыражениями
Положительный	87	6
Отрицательный	95	19
Удивление	13	9
Другое	55	19

### *Набор данных SAMM*

Набор данных SAMM [107] – это первый набор данных с высоким разрешением, содержащий 159 микродвижений, спонтанно воспроизводимых в ответ на зрительный стимул. SAMM является первым набором данных с разрешением 2040x1088 пикселей, площадь лица – 400x400 пикселей. В SAMM представлено 13 этнических групп со средним возрастом 33 года. С каждой видеопоследовательностью связаны индексы кадров, соответствующие началу и завершению микровыражения, и индекс так называемого пикового кадра (кадр с наибольшей интенсивностью микровыражения). В наборе данных представлено выражение семи эмоций: презрение, отвращение, страх, гнев, печаль, радость, удивление. Каждая видеопоследовательность в наборе данных содержит список единиц действий по СКЛид, задействованных в течение микровыражения (рисунок 13).



Рисунок 13 – Примеры изображений из набора данных SAMM

В таблице 6 представлен сводный обзор указанных выше наборов данных. В CASME и CASME II нет распределения по национальностям. SMIC задействует участников 3 национальностей, в то время как в SAMM представлены участники 13 национальностей. Скорость съемки в наборах данных CASME II и SAMM – 200 кадров в секунду. CAS(ME)<sup>2</sup> имеет ограниченное количество наборов микровыражений – 53. В CASME и SAMM 7 классов эмоциональных стимулов, CASME II - 5 классов, SMIC - 3 класса. CASME, CASME II и SAMM закодированы с использованием СКЛид. CASME II и SAMM являются основными наборами данных для исследования микролицевых движений, так как они удовлетворяют всем необходимым критериям для распознавания микровыражений, а именно: классы эмоций, высокая частота кадров, большое количество наборов микровыражений, различных по интенсивности движений лица [110].

В настоящее время количество наборов данных для проведения исследований по распознаванию микролицевых движений минимально. Основными препятствиями, ограничивающими исследование микровыражений, являются доступность и качество, а также недостаточная стандартизация данных, используемых исследователями.

В исследованиях участникам демонстрируют видеоролики, которые должны вызвать эмоции. В свою очередь, участникам следует попытаться скрыть возникшее эмоциональное возбуждение. Поскольку в некоторых случаях пробуждения эмоций не возникает, многие видеозаписи оказываются бесполезными, поскольку не содержат демонстрации микровыражений, что является одной из причин, по которой количество наборов микровыражений невелико, и почему каждый из наборов данных содержит относительно мало примеров классов эмоций.

Другая проблема связана с кодированием или маркированием данных, которые занимают очень много времени. Процесс требует наличия обученного и квалифицированного маркировщика, повторного изучения видеозаписей

участников (часто в замедленном темпе) и маркирования начала, пика и окончания микровыражения. Таким образом, помимо того, что процесс является трудоемким и медленным, он также является неточным, поскольку имеет место межметочная изменчивость. С этой проблемой тесно связан тот факт, что не существует единого и общепринятого стандарта классификации микровыражений.

Таблица 6 – Наборы данных спонтанных микролицевых движений

Набор данных	CASME	SMIC HS	SMIC VIS	SMIC NIR	CASME II	SAMM	CAS(ME) 2
Кол-во Микровыражений	195	164	71	71	247	159	57
Кол-во участников	35	20	10	10	35	32	22
Скорость съемки	60	100	25	25	200	200	30
Кол-во национальн.	1	Нет данных	3	Нет данных	1	13	1
Средний возраст	22,03	Нет данных	22,03	33,24	22,59	33	22,03
Разрешение	640x480	640x480	640x480	2040x1088	640x480	2040x1088	640x480
Разрешение лица	150x190	190x230	280x340	400x400	Нет данных	400x400	150x190
Классы эмоций	Радость, грусть, отвращение, удивление, презрение, страх, подавление, напряжение	Положительные, отрицательные, удивление	Радость, отвращение, удивление, подавление, другие	Презрение, отвращение, страх, злость, грусть, радость, удивление	Положительные, Отрицательные, удивление, другие	Радость, грусть, отвращение, удивление, презрение, страх, подавление, напряжение	Радость, грусть, отвращение, удивление, презрение, страх, подавление, напряжение
Соответствие кодированию СКЛиД	+	-	+	+	-	Соответствие кодированию СКЛиД	+



Подходы к маркированию, принятые для разных наборов данных, различаются (при этом аналогичные микровыражения рассматриваются как разные в зависимости от используемого набора данных), что создает проблемы для понимания относительных характеристик различных методов, их преимуществ и недостатков.

### **1.6 Выводы по первой главе**

В первой главе представлен обзор пространственно-временных дескрипторов признаков, применяемых в задачах распознавания движений, а также обнаружения микролицевых движений по видеопоследовательности изображений. Представлены основные этапы обнаружения микролицевых движений. Рассмотрены наборы данных, применяемые при решении задачи обнаружения микролицевых движений.

На основе анализа представленной информации можно сделать следующие выводы о возможных путях повышения эффективности распознавания микролицевых движений.

1. Известные подходы к вычислению пространственно-временных дескрипторов в задачах распознавания движений по видеопоследовательности изображений имеют ограничения и существенную специфику применения при решении задачи обнаружения микролицевых движений, что связано с быстротечностью и небольшой амплитудой произвольных микродвижений лицевых мышц.

2. Процесс обнаружения микродвижений лица состоит из трех последовательных этапов, корректность выполнения каждого из которых влияет на итоговую диагностическую эффективность обнаружения микролицевых движений. Первым этапом является локализация антропометрических точек лица, соответствующих областям возможного появления микродвижений, которая может выполняться различными методами автоматического определения лицевых ориентиров. Вторым этапом является вычисление пространственно-временных дескрипторов признаков микродвижений в локализованных областях. Последним

этапом является классификация вычисленных признаков с использованием методов машинного обучения. Таким образом, для повышения диагностической эффективности обнаружения микролицевых движений необходимо как повышать точность локализации областей интереса на видеоизображении, так и разрабатывать улучшенные алгоритмы вычисления пространственно-временных дескрипторов признаков для областей возможного проявления микродвижений.

3. Рассмотрение существующих аннотированных наборов данных спонтанных микровыражений лица показало, что подготовка таких наборов является одной из самых сложных проблем в области исследования микролицевых движений. Из-за триггерного механизма и непроизвольности микровыражений сложно построить большие наборы обучающих и тестовых данных для задач обнаружения микродвижений. Эта особенность ограничивает возможности применения активно развивающихся нейросетевых технологий глубокого обучения для реализации всего программного конвейера обнаружения микродвижений, но не исключает их использования на отдельных этапах, например, при обнаружении антропометрических точек лица.

В связи с вышеизложенным, в первой главе обосновано использование классических методов вычисления дескрипторов признаков, применяемых для обнаружения микродвижений лица, диагностическая эффективность которых может быть повышена за счет создания более точных алгоритмов обнаружения областей возможного проявления микролицевых движений на видеопоследовательности изображений и модификации последующих вычислительных процедур с учетом особенностей проблемы обнаружения микродвижений.

## **2 Разработка алгоритмов выбора областей возможного проявления микролицевых движений**

### **2.1 Алгоритмы поиска лица и глаз на изображении**

Обнаружение лица является важной задачей для многих приложений компьютерного зрения и биометрии, например, автоматическое распознавание лиц, видеонаблюдение, взаимодействие человека с компьютером и крупномасштабные системы поиска изображений лиц [111]. Первым этапом в любой из перечисленных систем является определение наличия лица, а затем положения его на изображении или видео.

Основная задача — справиться с широким спектром вариаций человеческого лица, а именно: поза, масштаб, ориентация, выражение, этническая принадлежность и цвет кожи.

Процедура выделения участков кожи может основываться на цвете кожи, благодаря которому возможно локализовать области лица без учета текстуры и геометрических свойств. Но данная реализация доступна только при рассмотрении цветных изображений. Внешние факторы, такие как окклюзия, сложный фон, непостоянные условия освещения и качество изображения, могут существенно усложнить решаемую задачу.

Задача обнаружения человеческого лица полностью и его областей решается не только в вышеперечисленных случаях, но является основным этапом, предшествующим этапу обнаружения микролицевых движений на изображении для последующего анализа. В данном разделе представлены две конфигурации алгоритма поиска лица и глаз на изображении. Первая конфигурация алгоритма основана на алгоритме Виолы-Джонса, вторая – на алгоритме оценки яркости. Произведен сравнительный анализ результатов, полученных в каждой из конфигураций.

Одним из алгоритмов поиска областей лица и глаз является алгоритм Виолы-Джонса, предложенный в 2001 году Полом Виолой и Майклом Джонсоном. Данный алгоритм считается наиболее оптимальным по соотношению

затрачиваемых вычислительных ресурсов и времени [112]. В нем используются признаки Хаара, с помощью которых происходит поиск нужного объекта, в данном случае – лица и глаз.

Все признаки поступают на вход классификатора, который дает результат «истина» или «ложь». Обучение классификатора происходит медленно, но получение результатов поиска быстрое. Алгоритм Виолы-Джонса в исходном или модифицированном виде используется для поиска лица и глаз в работах [113,114].

Второй алгоритм основан на яркостной характеристике кожи в цветовом пространстве YCrCb и RGB. Данный подход используется в работах [111, 113,115-117].

Цветовое пространство RGB представляет три основных цвета: красный, зеленый и синий. RGB является аппаратно-ориентированной моделью, применяемой в цветных мониторах. YCrCb – это аппаратно-ориентированная модель, но в отличие от RGB в данной модели компонента яркости отделена от компоненты цвета. Значение Y представляет собой яркость, а значения Cr и Cb, также известные как сигналы цветового различия, представляют компоненты цветности изображения.

Поскольку обработка цвета происходит быстрее, чем обработка других признаков, данные о цвете лица не зависят от его угла поворота. Таким образом, для определения области лица используется цвет кожи. Одним из вариантов цветового пространства без освещения является хроматическое цветовое пространство. Изображение преобразуется из цветового пространства RGB в YCbCr в соответствии с уравнением (13):

$$\begin{cases} Y = 0,299R - 0,587G - 0,114B \\ Cb = R - Y \\ Cr = B - Y \end{cases} . \quad (13)$$

Алгоритм поиска областей лица и глаз в цветовых пространствах YCrCb и RGB состоит из набора последовательных операций обработки и фильтрации входного изображения. Схема алгоритма представлена на рисунке 14.

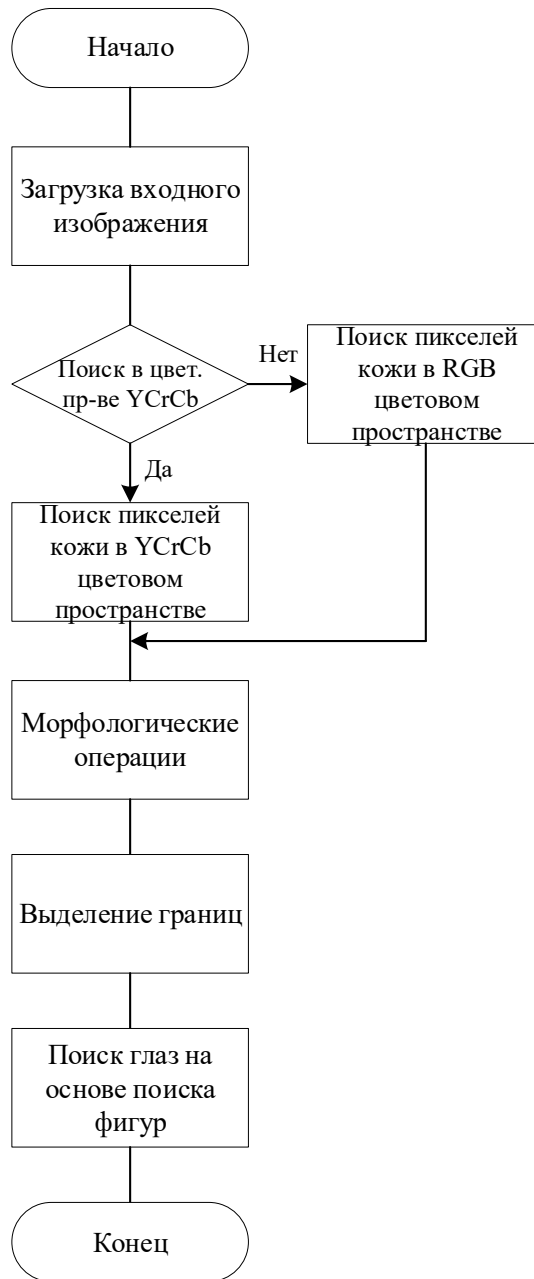


Рисунок 14 – Схема алгоритма поиска лица и глаз

На вход подается исходное изображение, к которому в зависимости от выбранной конфигурации алгоритма применяется алгоритм поиска пикселей кожи на основе анализа яркостей, базирующийся на цветовом пространстве YCrCb, или

алгоритм Виолы-Джонса, основанный на цветовом пространстве RGB. После чего реализуются морфологические операции наращивания, в процессе которых увеличивается интересующая область и эрозии, уточняется исследуемая область на изображении путем удаления всех объектов, меньших, чем объект интереса. Затем выполняется выделение границ, а завершающим шагом является поиск глаз на основе метода поиска фигур.

Для проведения экспериментов использовался набор данных фронтальных изображений лиц IMM. Исходные изображения представлены на рисунке 15.

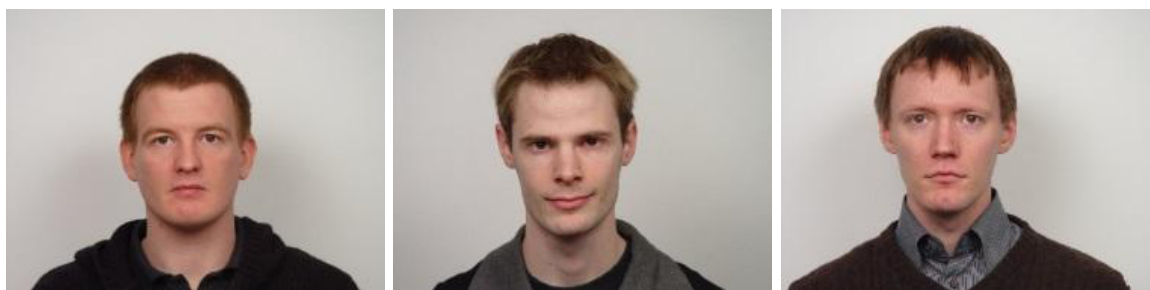


Рисунок 15 – Примеры изображений из набора данных IMM

Результаты применения алгоритма, основанного на анализе яркостей, представлены на рисунке 16.

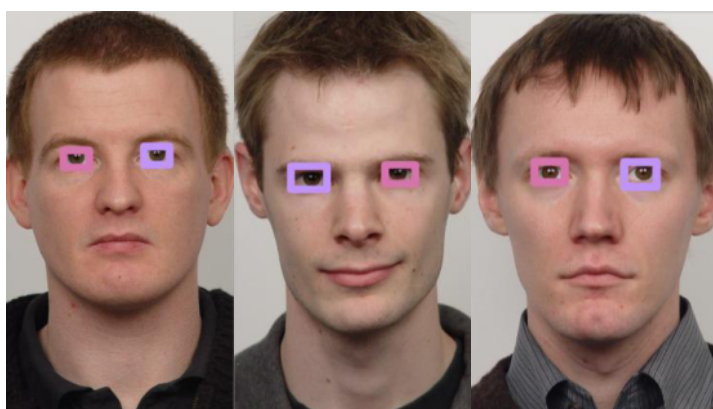


Рисунок 16 – Результаты алгоритма на основе анализа яркостей

Результаты применения алгоритма, основанного на алгоритме Виолы-Джонса, представлены на рисунке 17.

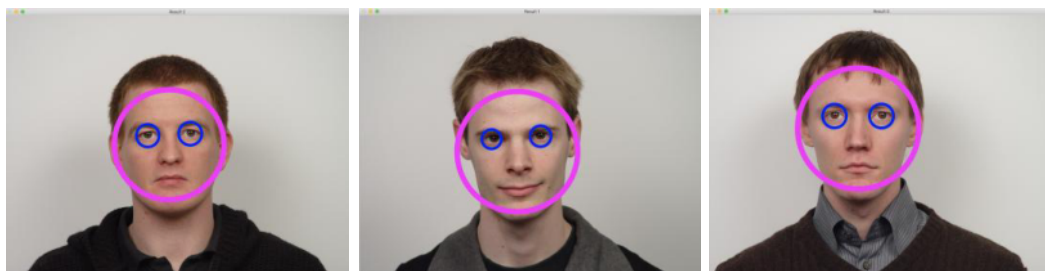
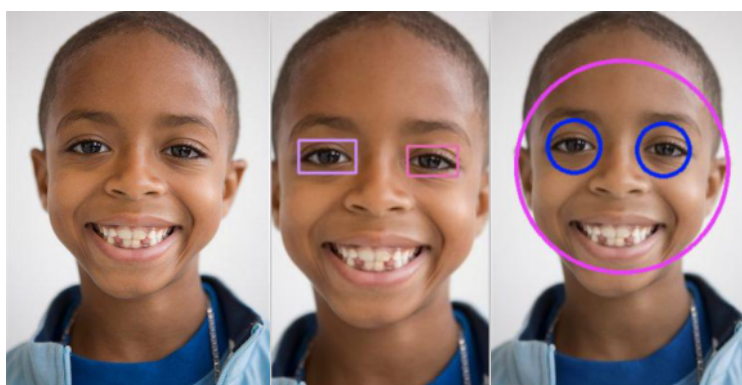


Рисунок 17 – Результаты алгоритма на основе алгоритма Виолы-Джонса

Конфигурации алгоритма поиска лица и глаз применимы для обработки изображений людей с темным цветом кожи. Результаты применения алгоритма для изображений людей с темным цветом кожи представлены на рисунке 18.



а

б

в

Рисунок 18 – Результаты поиска лица и глаз на изображениях людей с темным цветом кожи: а – исходное изображение, б – конфигурация алгоритма на основе анализа яркостей, в – конфигурация алгоритма на основе алгоритма Виола-Джонса

Результаты экспериментов показали, что обе конфигурации одинаково хорошо справляются с задачей поиска лица и глаз на цветных изображениях, при условии, если цвет кожи контрастирует с цветом фона и лицо смотрит прямо в камеру. При оценке времени выполнения каждого из алгоритмов на разных разрешениях изображений было рассчитано среднеквадратичное отклонение  $\sigma$ . Сравнение характеристик алгоритмов представлено в таблице 7.

Таблица 7 – Сравнение характеристик алгоритмов

Разрешение изображения	Алгоритм на основе анализа яркостей		Алгоритм Виолы-Джонса	
	$\bar{t}, ms$	$\sigma, ms$	$\bar{t}, ms$	$\sigma, ms$
664x1000	63,57	3,08	190,41	11,06
810x1071	88,48	7,15	227,81	33,36
2560x1920	466,78	2,78	437,19	20,33
2560x1920	469,28	4,34	735,16	15,88
2560x1920	469,19	2,02	461,19	34,02
2560x1920	464,33	5,04	356,91	19,27
2560x1920	485,89	9,31	420,29	6,82

Результаты экспериментов показали, что конфигурация алгоритма, основанная на анализе яркостей, имеет преимущество в быстродействии на небольших изображениях, разрешение которых не превышает 2560x1920. Но в экспериментах на изображениях с разрешением 2560x1920 алгоритм Виолы-Джонса работает быстрее. Также следует отметить, что алгоритм Виолы-Джонса работает как с цветными, так и с черно-белыми изображениями, в то время как алгоритм, основанный на анализе яркостей, имеет ограничение, связанное с возможностью работы только с цветными изображениями [118].

Обе конфигурации алгоритма могут быть применены для предварительной обработки изображений в конвейере обнаружения микровыражений лица.

## 2.2 Алгоритм выбора антропометрических точек вокруг глаз

Лицо играет важную роль в визуальной коммуникации. Глядя на лицо, человек может автоматически извлечь множество невербальных сообщений, таких как намерения и эмоции другого человека. В компьютерном зрении для автоматического извлечения лицевой информации требуется локализация антропометрических точек лица. Поэтому, одним из самых главных и фундаментальных этапов, следующим за поиском лица, является определение антропометрических точек на лице. Информация о лице, полученная через расположение вычисленных точек, используется для взаимодействия человека и



компьютера, контроля безопасности, в медицинских приложениях для обнаружения и распознавания микровыражений лица.

Выделяют категории методов решения задачи поиска антропометрических точек на лице: целостные методы, методы, основанные на ограниченной локальной модели, регрессионные методы.

Наиболее известным целостным методом является модель активного внешнего вида (Active Appearance Model, AAM), из семейства CLM – ASM.

AAM была предложена Тейлором и Кутсом. Это статистическая модель, которая аппроксимирует изображения лица с помощью небольшого количества параметров, контролируя как внешний вид, так и вариации формы. Во время построения модели AAM последовательно строит глобальную модель формы лица и целостную модель внешнего вида лица на основе анализа главных компонент PCA. На этапе обнаружения AAM определяет местоположения ориентиров, подгоняя изученные модели внешнего вида и формы к тестовому изображению.

Методы CLM вычисляют местоположения ориентиров лица на основе глобальных моделей формы лица, а также независимой локальной информации о внешнем виде вокруг каждого ориентира. Методы из семейства CLM представляют собой либо детерминированные, либо вероятностные методы. Одной из самых известных и общепринятых моделей является ASM [119].

Модель ASM по сравнению с AAM имеет следующие преимущества: более точна при локализации точки или контура, менее чувствительна к колебаниям освещенности; лучше адаптирована под приложения, требующие точного контурирования. Со времени первого использования ASM появилось множество расширений для повышения ее надежности, производительности и эффективности [120]. Однако, несмотря на последние изменения, построение данной модели остается нерешенной задачей, так как не существует единой модели формы, которая могла бы отразить все естественные вариации формы лица. Данные модели прогнозируют положение лицевых ориентиров или их групп на основе обучения по заранее размеченным вручную изображениям, но этот процесс

требует значительных временных затрат для подготовки обучающих наборов данных.

В то же время существуют методы, выполняющие поиск каждой точки самостоятельно, то есть без предварительного обучения. Преимущество независимых точечных ориентиров лица состоит в том, что нет необходимости в этапе инициализации. Однако одним из основных недостатков такого подхода является проблема двусмысленности, обусловленная наличием нескольких позиций, похожих на целевой ориентир.

В данной диссертационной работе разработан алгоритм обнаружения антропометрических точек вокруг глаз на основе геометрии изображения глаза без предварительного обучения. Результаты, полученные на основе применения разработанного алгоритма, сравниваются с результатами работы модели ASM.

Первым шагом ASM является создание обучающей выборки, которая представляет собой размеченные вручную по схеме разметки лица, представленной на рисунке 19, наборы изображений. Общее количество лицевых ориентиров на схеме соответствует значению 79. Форма объекта отображается через ориентиры, представляющие собой единую цепочку последовательных точек в количестве, достаточном для охвата всей формы и деталей лица.

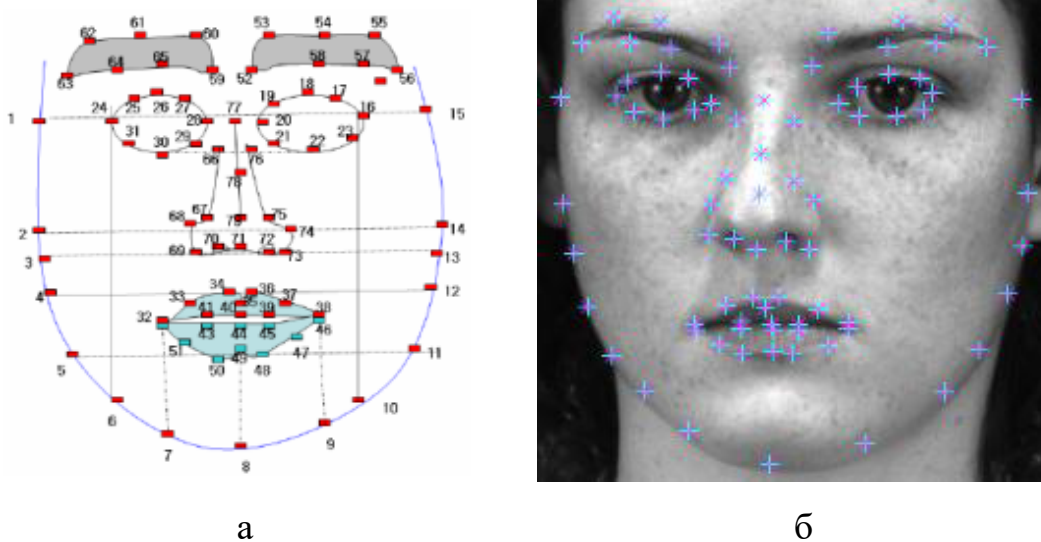


Рисунок 19 – Антропометрические точки лица:  
а – схема разметки точек лица, б – размеченное лицо

ASM находит изменения в обучающих данных с помощью анализа главных компонент, который позволяет модели автоматически распознавать, является ли контур приемлемым контуром объекта.

После создания модели исходный контур деформируется путем поиска наилучшего соответствия вычисленной усредненной модели контрольным точкам. Процесс является итерационным (рисунок 20), перемещение контрольных точек ограничивается «нормальным» контуром объекта, который вычисляется на основе обучающих данных [121].



Рисунок 20 – Процесс поиска наилучшего соответствия построенной модели тестовому изображению

Результат вычисления антропометрических точек на лице с помощью ASM представлен на рисунке 21 [119].



Рисунок 21 – Результат применения ASM

ASM имеет следующие недостатки:

- положение результирующих антропометрических точек сильно разнится со схемой разметки;
- при увеличении количества обучающих наборов изображений результирующая модель сильно обобщается;
- время вычисления антропометрических точек определяется количеством итераций [120,121].

Таким образом, экспериментальные исследования показали, что ASM наряду со своими достоинствами имеет ряд недостатков. Поэтому, для дальнейшего повышения точности обнаружения антропометрических точек разработан алгоритм, ориентированный на конкретные участки лица, а именно на глаза.

В разработанном алгоритме поиск ориентиров лица реализуется на основе СКЛиД, в которой имеется 26 зон, склонных к демонстрации микровыражений. Одной из таких зон является область глаз. Схема разработанного алгоритма поиска антропометрических точек вокруг глаз представлена на рисунке 22.

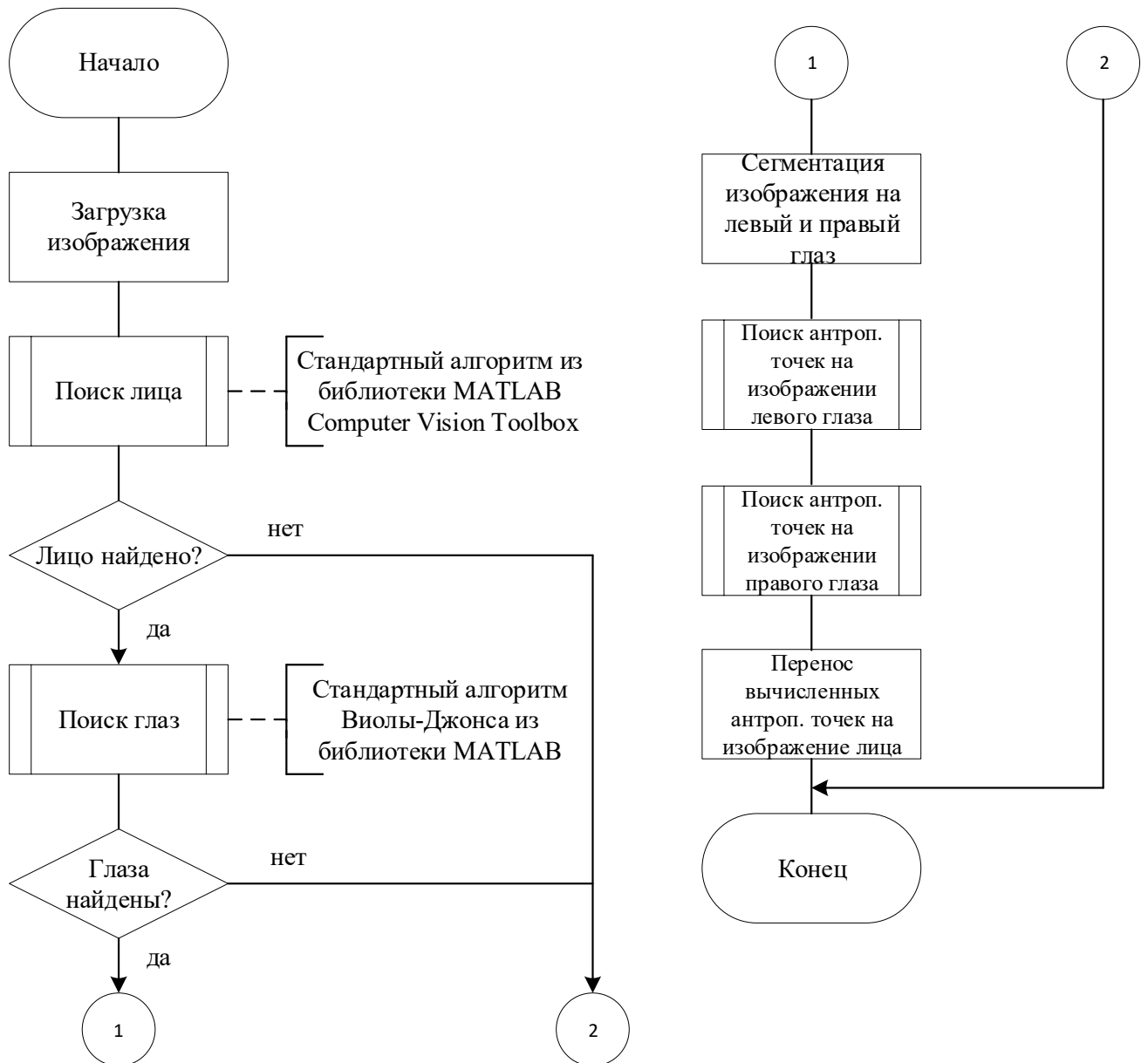


Рисунок 22 – Схема алгоритма поиска антропометрических точек вокруг глаз

Первым шагом является поиск лица на изображении, который реализуется с помощью готового алгоритма из библиотеки MATLAB Computer Vision Toolbox. Если лицо на изображении обнаружено, следующим шагом является поиск глаз на первом изображении последовательности. Данный шаг выполняется с помощью алгоритма Виолы-Джонса, реализованного в MATLAB и показавшего хорошие результаты, представленные в параграфе 2.1. Затем выполняется сегментация

изображения на левый и правый глаз. Последующая обработка реализуется для каждого изображения глаза в отдельности.

Вычисление антропометрических точек на каждом из глаз реализуется в соответствии со схемой алгоритма, представленной на рисунке 23.

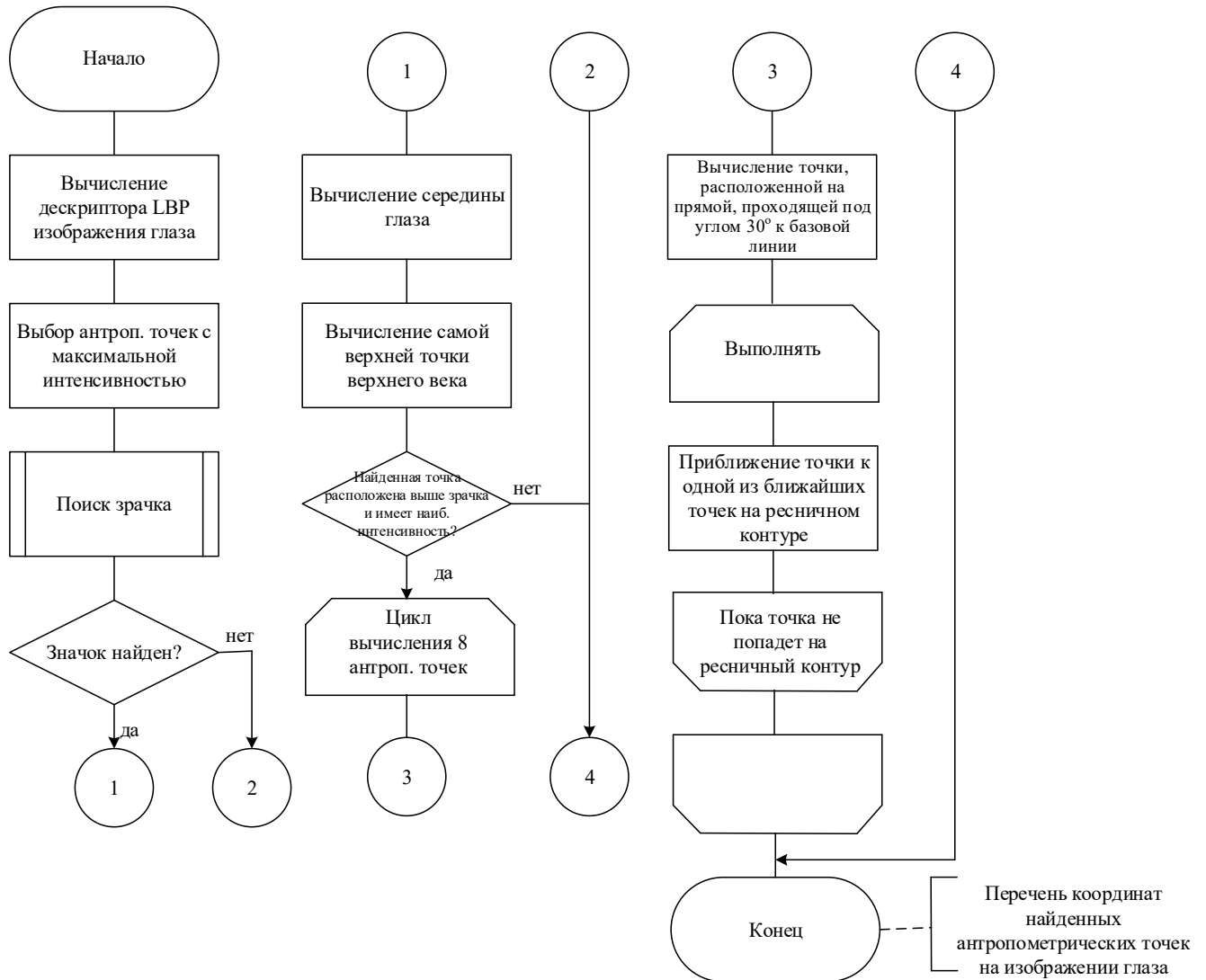


Рисунок 23 – Схема алгоритма поиска антропометрических точек на изображении глаза

Первым шагом является вычисление дескрипторов LBP изображения глаза. Анализ текстуры LBP определяется как мера текстуры, инвариантная по шкале серого, полученная из общего определения текстуры в локальной окрестности. Оператор LBP рассматривается как объединяющий подход к традиционно

расходящимся статистическим и структурным моделям анализа текстур. Результат выполнения данного шага представлен на рисунке 24.

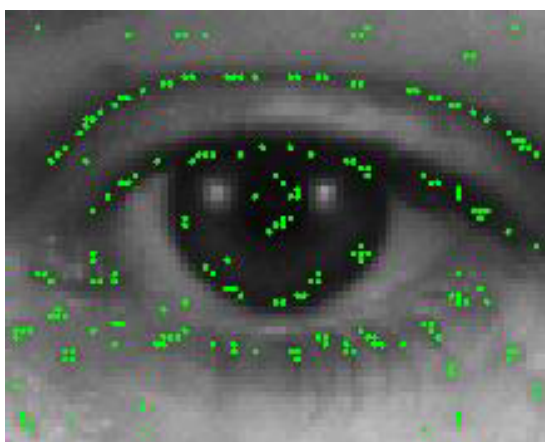


Рисунок 24 – Вычислены точки с наибольшей интенсивностью

Вторым шагом является поиск внешнего и внутреннего уголков глаза. В качестве точек, удовлетворяющих данному условию, выбираются крайние точки по горизонтали с максимальной интенсивностью (рисунок 25).

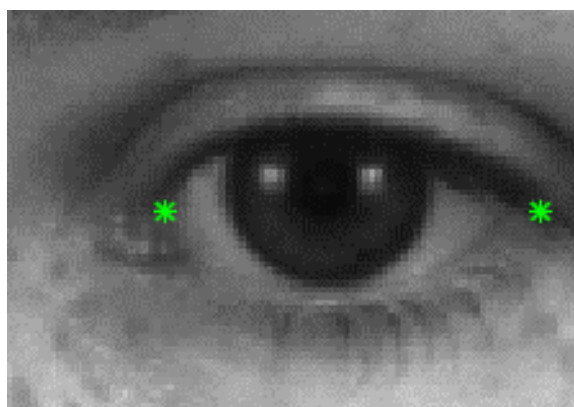


Рисунок 25 – Вычисленные уголки глаза

Третьим шагом является поиск зрачка. Данный шаг реализуется путем применения медианного фильтра для удаления зашумленности, преобразования полученного черно-белого изображения в бинарное изображение, инвертирование изображения и поиск объектов окружностей на изображении.

Результат выполнения, описанных выше шагов представлен на рисунках 26 и 27.



Рисунок 26 – Инвертированное изображение

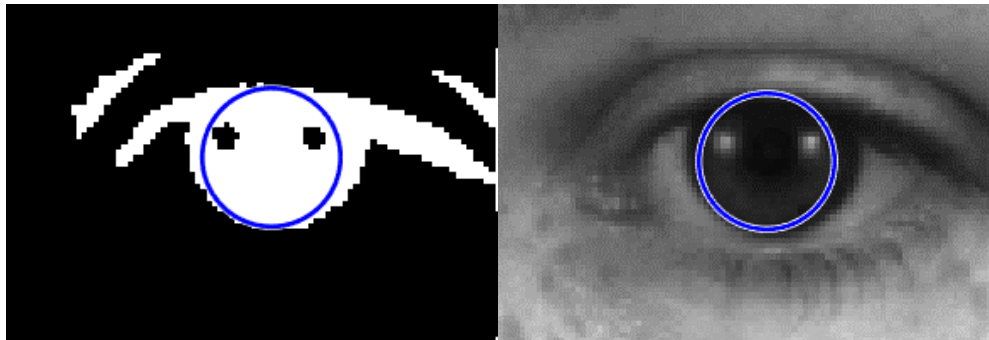


Рисунок 27 – Найденный зрачок

Четвертым шагом является вычисление геометрического центра глаза, которое реализуется путем вычисления середины между уголками глаза (Рисунок 28).

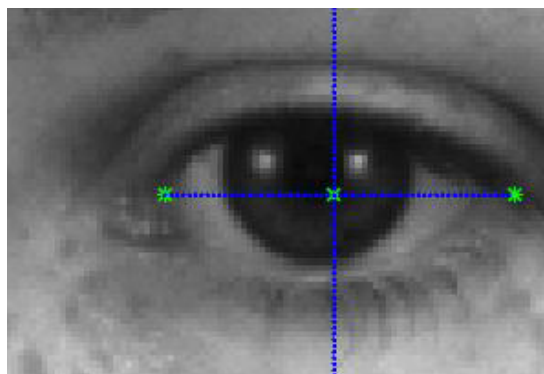


Рисунок 28 – Геометрический центр глаза



Пятым шагом является вычисление антропометрических точек на верхнем веке.

1. Вычисление самой верхней точки верхнего века по отношению к точке геометрического центра глаза. Выбор соответствующей точки реализуется на основе оценки диаметра окружности, описывающей видимую часть зрачка, и карты интенсивностей глаза, полученной на первом шаге. Самая верхняя точка верхнего века должна удовлетворять следующим условиям: она должна быть расположена выше зрачка и соответствовать одной из точек с наибольшей интенсивностью, лежащих на ресничном контуре, то есть она должна сама лежать на ресничном контуре верхнего века (рисунок 29).

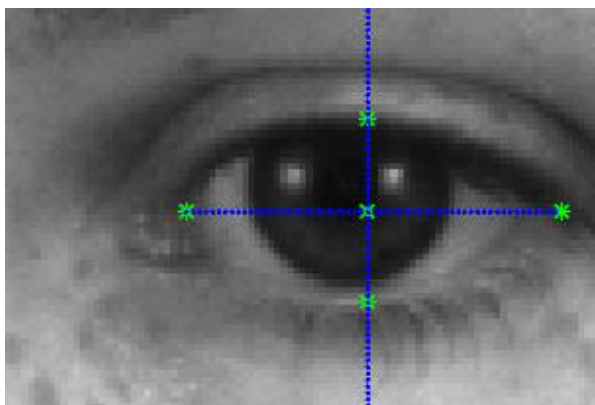


Рисунок 29 – Определение антропометрических точек на верхнем и нижнем веке

2. Вычисление антропометрических точек, расположенных в верхнем правом секторе глаза (рисунок 30). Из геометрического центра глаза проводится луч до самой верхней точки верхнего века под углом  $30^\circ$  к базовой линии, перпендикулярной центру. Затем вычисляется оценка попадания данной точки на ресничный контур, а именно в точку с наибольшей интенсивностью. Если точка не попала на ресничный контур, выполняется ее приближение и снова оценка попадания на ресничный контур. Считается, что точка попала на ресничный контур, если она имеет наименьшее расстояние с точкой интенсивности, расположенной на ресничном контуре. После того, как первая точка будет

определена, аналогичным образом выполняется постановка остальных точек. Все точки на верхнем веке вычисляются примерно на одинаковом расстоянии между собой.

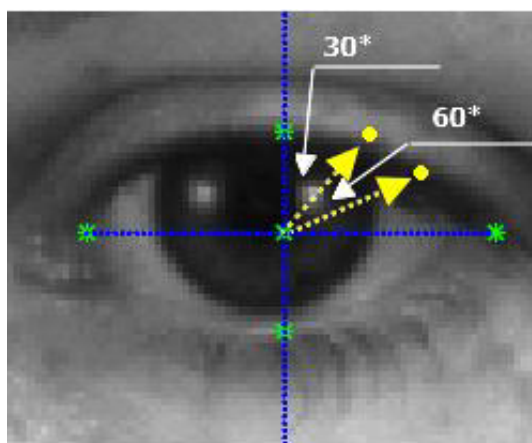


Рисунок 30 – Вычисление антропометрических точек в первом секторе

Шестым шагом является вычисление антропометрических точек на нижнем веке. Данный шаг выполняется аналогично предыдущему шагу.

Результат выполнения описанных выше шагов представлен на рисунке 31.

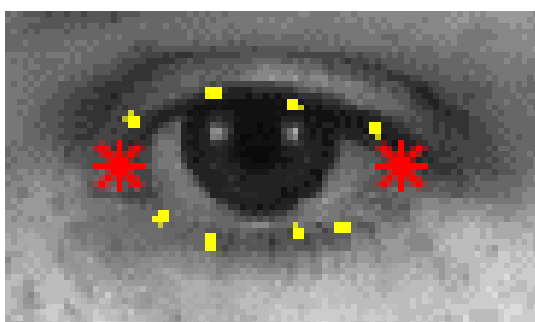


Рисунок 31 – Антропометрические точки, вычисленные на верхнем и нижнем веке левого глаза

Седьмым шагом является перенос вычисленных антропометрических точек на исходное изображение лица.

Результат переноса вычисленных антропометрических точек на исходное изображение представлен на рисунке 32.



Рисунок 32 – Результат применения разработанного алгоритма выбора антропометрических точек вокруг глаз

После того как антропометрические точки на глазах на первом изображении последовательности найдены, выполняется отслеживание данных точек на последующих кадрах последовательности с помощью алгоритма Канада-Лукаса-Томаси (Kanade-Lucas-Tomasi, KLT). Таким образом, антропометрические точки определяются только один раз, затем алгоритм KLT отслеживает найденные точки на всех изображениях последовательности [123].

Для сравнения полученных данных с помощью разработанного алгоритма и модели ASM с точки зрения точности обнаружения антропометрических точек вокруг глаз было выбрано образцовое изображение, которым являлось то же самое изображение, что было передано на вход, но размеченное вручную.

Для сравнения данных выбиралась очередная точка, полученная с помощью модели ASM, ее координаты сравнивались со считанными координатами образцового изображения. Для всех точек образцового изображения  $p_r$  с координатами  $(x_r, y_r)$  и точек ASM  $p_i$  с координатами  $(x_i, y_i)$  рассчитывалось расстояние по формуле (14):

$$S_{ir} = \sqrt{(x_r - x_i)^2 + (y_r - y_i)^2} . \quad (14)$$

Минимальное расстояние из рассчитанных расстояний выбиралось по формуле (15):

$$S_{i_{\min}} = \Delta_i = \min_r(S_{ir}) . \quad (15)$$

Минимальное расстояние принимается как показатель среднеквадратичного отклонения между координатами антропометрических точек, обнаруженных с помощью алгоритма ASM и координатами точек образцового изображения. После нахождения отклонений для точек, найденных на входном изображении, рассчитывалось среднее отклонение по формуле (16):

$$\Delta = \frac{\sum_{i=0}^n \Delta_i}{n} . \quad (16)$$

Аналогичное сравнение выполнялось для разработанного алгоритма.

В результате для каждого анализируемого изображения было найдено среднеквадратичное отклонение вычисленных координат антропометрических точек вокруг глаз с помощью разработанного алгоритма и модели ASM от координат антропометрических точек образцового изображения. Результаты исследования представлены в таблице 8.

Таблица 8 – Результирующие среднеквадратичные отклонения

№	ASM	Разработанный алгоритм
1	7,00	3,08
2	9,41	3,91
3	4,65	2,36
4	6,57	2,09
5	7,38	5,49

Из анализа полученных данных следует, что разработанный алгоритм показывает меньшее в 2,2 раза среднеквадратичное отклонение в расположении антропометрических точек вокруг глаз, по сравнению с моделью ASM.

На основе полученных результатов следует, что разработанный алгоритм обладает следующими преимуществами:

- обеспечивает более точную постановку антропометрических точек по сравнению с ASM;
- не требует обучения;
- не требует дополнительных действий человека (разметка изображений для дальнейшего обучения);
- работает с черно-белыми изображениями.

Но, так как в разработанном алгоритме поиска антропометрических точек вокруг глаз для поиска лица и глаз используется метод Виолы-Джонса, данный алгоритм имеет недостаток, который заключается в неспособности определить оба глаза при фиксации лица с некоторым углом поворота [124].

### **2.3 Классификация современных методов обнаружения антропометрических точек**

Замечено, что для конвейера анализа микролицевых движений необходимо более точное обнаружение антропометрических точек лица, поэтому в дальнейшем рассматриваются другие методы вычисления ориентиров.

Методы, основанные на регрессии, напрямую изучают сопоставление внешнего вида изображения с местоположениями ориентиров и не строят глобальную модель формы. Они считаются более перспективными по сравнению с методами, представленными в параграфе 2.2. Методы, основанные на регрессии, можно разделить на методы прямой регрессии, методы каскадной регрессии и методы регрессии, основанные на глубоком обучении [125]. Методы каскадной регрессии показывают лучшие результаты, чем методы прямой регрессии.

Процедура обнаружения ориентиров лица с помощью метода каскадой регрессии представлена в [92]. Для обнаружения ориентиров лица используется

каскад функций регрессии. Пусть  $x_i \in \mathbb{R}^2$ , где  $x$  и  $y$ -координаты  $i$ -го лицевого ориентира на изображении  $I$ . Тогда вектор  $S = (x_1^T, x_2^T, \dots, x_p^T)^T \in \mathbb{R}^{2p}$  обозначает координаты всех  $p$  лицевых ориентиров на  $I$ . В дальнейшем вектор  $S$  будет называться формой. Параметр  $\hat{S}^{(t)}$  обозначает текущую оценку формы  $S$ . Каждый регрессор  $r_t$  в каскаде предсказывает вектор обновления на основе изображения и  $\hat{S}^{(t)}$ , который добавляется к текущей оценке формы для ее уточнения (17):

$$\hat{S}^{(t+1)} = \hat{S}^{(t)} + r_t(I, \hat{S}^{(t)}). \quad (17)$$

Регрессор  $r_t$  делает свои прогнозы на основе значения интенсивности пикселей, вычисленных из изображения  $I$ , индексированных относительно текущей оценки формы  $\hat{S}^{(t)}$ .

Диапазон выходных данных лежит в линейном подпространстве обучающих данных. Поэтому, отсутствует необходимость налагать дополнительные ограничения на прогнозы, что значительно упрощает метод. Начальная форма выбирается как средняя форма обучающих данных, центрированная и масштабированная в соответствии с выходными данными ограничивающего прямоугольника обычного детектора лиц.

Обучение регрессора выполняется следующим образом. Предположим, имеются данные для обучения  $(I_1, S_1), \dots, (I_n, S_n)$ , где каждый  $I_i$  представляет собой изображение лица, а  $S_i$  — вектор формы лица. Чтобы изучить первую функцию регрессии  $r_0$  в каскаде, создаются триплеты обучающих данных, которые содержат изображение лица, начальную оценку формы и целевой шаг обновления, то есть  $(I_{\pi_i}, \hat{S}_i^{(0)}, \Delta S_i^{(0)})$ , где  $\pi_i \in \{1, \dots, n\}$ ,  $\hat{S}_i^{(0)} \in \{S_1, \dots, S_n\} \setminus S_{\pi_i}$ ,  $\Delta S_i^{(0)} = S_{\pi_i} - \hat{S}_i^{(0)}$  для  $i = 1, \dots, N$ .

Общее количество триплетов равно  $N = nR$ , где  $R$  — количество инициализаций, используемых для каждого изображения  $I_i$ . Каждая начальная

оценка формы изображения выбирается из  $\{S_1, \dots, S_n\}$ . Таким образом, становится известной функция регрессии  $r_0$ . Набор обучающих триплетов  $(I_{\pi_i}, \hat{S}_i^{(1)}, \Delta S_i^{(1)})$  обновляется для предоставления обучающих данных для следующего регрессора  $r_1$  в каскаде:

$$\hat{S}_i^{(t+1)} = \hat{S}_i^{(t)} + r_i(I_{\pi_i}, \hat{S}_i^{(t)}) , \quad (18)$$

$$\Delta S_i^{(t+1)} = S_{\pi_i} - \hat{S}_i^{(t+1)} . \quad (19)$$

Процесс повторяется до тех пор, пока не образуется каскад  $T$  регрессоров  $r_0, r_1, \dots, r_{T-1}$  с приемлемым уровнем точности. Каждый регрессор  $r_i$  изучается с использованием алгоритма повышения градиента. В формулировке алгоритма присутствует параметр скорости обучения  $0 < \nu < 1$ , также известный как коэффициент сжатия. Установка  $\nu < 1$  помогает бороться с переобучением и обычно приводит к тому, что регрессоры обучаются гораздо лучше, чем те, которые были обучены при  $\nu = 1$ . В каждом узле дерева регрессии принимается решение на основе определения порога разницы между интенсивностями пикселей. Пиксели, используемые в тестовом изображении, находятся в позициях  $u$  и  $v$ , если они определены в системе координат усредненной формы. Для изображения лица произвольной формы необходимо проиндексировать точки, которые имеют то же положение относительно формы, что  $u$  и  $v$  для средней формы. Изображение перед вычислением ориентиров деформируется до средней формы на основе текущей оценки формы. Поскольку используется только очень разреженное представление изображения, гораздо эффективнее деформировать расположение точек, а не все изображение. Более того, грубое приближение деформации может быть выполнено с использованием глобального преобразования подобия.

Пусть  $k_u$  - индекс лицевого ориентира средней формы, ближайший к  $u$ , смещение от  $u$  определяется по формуле (20):

$$\delta x_u = u - \bar{x}_{k_u}. \quad (20)$$

Тогда для формы  $S_i$ , определенной на изображении  $I_i$ , положение в  $I_i$ , которое аналогично  $u$  в изображении средней формы, определяется выражением (21):

$$u' = x_{i,k_u} + \frac{1}{s_i} R_i^T \delta x_u, \quad (21)$$

где  $S_i$  и  $R_i$  — масштаб и матрица вращения преобразования подобия, которое преобразует  $S_i$  в  $\bar{S}$ , среднюю форму.

Масштаб и вращение минимизируются (22):

$$\sum_{j=1}^p \left\| \bar{x}_j - (s_i R_i x_{i,j} + t_i) \right\|^2, \quad (22)$$

сумма квадратов между ориентирами лица средней формы,  $\bar{x}_j$  и точками искривленной формы,  $v'$  определяется аналогично.

Формально каждое разделение представляет собой решение, включающее 3 параметра  $\theta = (\tau, u, v)$  и применяется к каждому процессу обучения и тестирования (23):

$$h(I_{\pi i}, \hat{S}_i^{(t)}, \theta) = \begin{cases} 1, & I_{\pi i}(u') - I_{\pi i}(v') > \tau \\ 0, & \text{остальные случаи} \end{cases}, \quad (23)$$

где  $u_0$  и  $v_0$  определяются с использованием матрицы масштаба и вращения, которая лучше всего преобразует  $\hat{S}_i^{(t)}$  в  $\bar{S}$ .



Вычисление преобразования подобия выполняется только один раз на каждом уровне каскада.

Алгоритм каскадной регрессии представляет собой алгоритм машинного обучения, который реализован в библиотеке `dlib` и предварительно обучен. OpenFace — набор библиотек, моделей и методов для обнаружения, отслеживания положения, идентификации лиц на изображениях с использованием глубоких нейронных сетей, использующих каскадную регрессию, реализованную в библиотеке `dlib`. OpenFace представляет собой набор предварительно обученных моделей нейронных сетей. Для обнаружения лица используется библиотека `OpenCV`. Затем обнаружение ориентиров выполняется с помощью библиотеки `dlib`, которая позволяет распознать 68 лицевых ориентиров (рисунок 33).

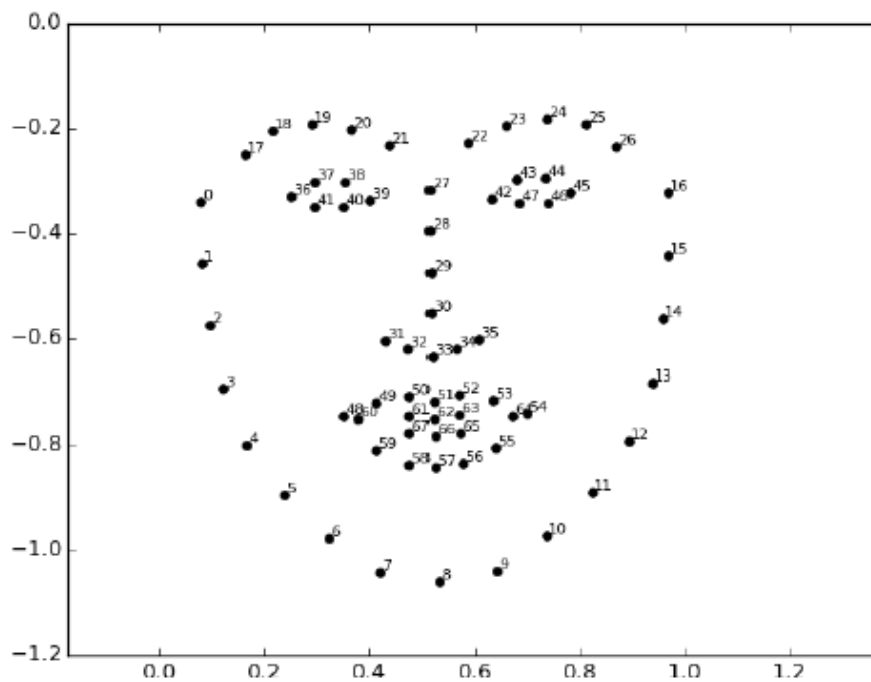


Рисунок 33 – 68 лицевых ориентиров созданных OpenFace

Другой современный метод обозначения ориентиров лица, основанный на глубоком обучении, представлен в [92]. Данная модель на основе нейронной сети для получения приближенного трехмерного представления человеческого лица создает плотную сетку из 468 ориентиров лица. Входными данными для модели

являются изображения, выходными – сетка лица, состоящая из обнаруженных ориентиров лица.

Аннотирование данного количества точек выполняется следующим образом.

1. Кадр с камеры обрабатывается детектором лиц, который создает рамку, ограничивающую лицо, и несколько ориентиров, которые используются для поворота рамки, чтобы совместить линию, соединяющую геометрические центры глаз, с горизонтальной осью.

2. Выделенное лицо, вырезается из исходного изображения и изменяется его размер, чтобы сформировать входные данные для нейронной сети прогнозирования сетки (размер варьируется от 256x256 пикселей в полной модели до 128x128 в самой маленькой). Модель создает вектор трехмерных координат ориентиров, который впоследствии сопоставляется с исходной системой координат изображения.

Координаты  $x$  и  $y$  вершин соответствуют местоположениям точек в 2D-плоскости и соответствующим координатам пикселей изображения.  $Z$ -координаты интерпретируются как глубина относительно базовой плоскости, проходящей через центр масс сетки.

Рецептивные поля нейронов охватывают большие площади входного изображения. Когда рецептивное поле достигает границы изображения, его относительное местоположение во входном изображении становится неявно доступным для модели. Тогда нейроны более глубоких слоев способны различать признаки, например, относящиеся ко рту и глазам. Данная модель реализована в библиотеке TensorFlow.

Программным решением, использующим данную модель является MediaPipe Face Mesh, который позволяет обнаружить 468 ориентиров на лице, расположенных в фиксированных квадрантах. Топология сетки представлена на рисунке 34.

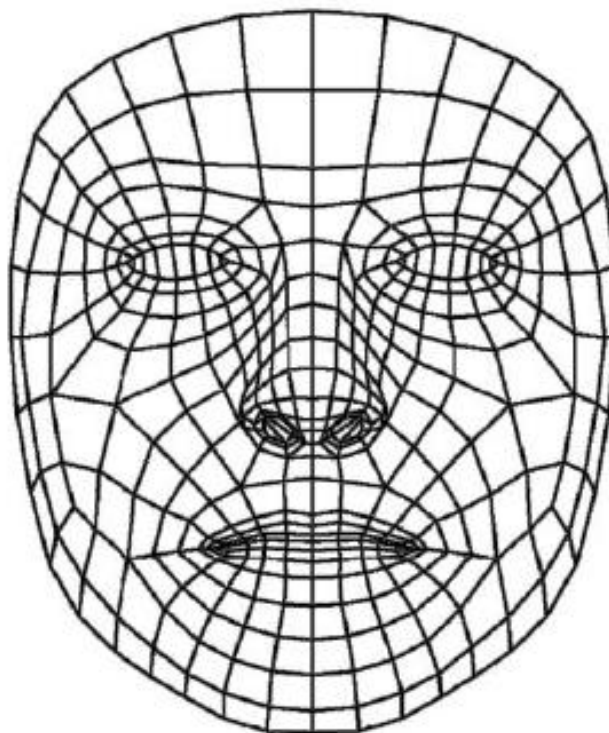


Рисунок 34 – Топология сетки восстанавливаемой MediaPipe Face Mesh

В работе [94] было проведено исследование представленных выше современных методов, применяемых к решению задачи вычисления ориентиров лица. В качестве исходных данных использовался набор данных SAMM. Данные, полученные в результате применения каждой из технологий, были нормализованы, так как MediaPipe Face Mesh обнаруживает 468 ориентиров, а OpenFace – 68. В качестве эталонного изображения было принято изображение, аннотированное вручную. После чего, вычислялось отклонение ориентиров, полученных OpenFace и MediaPipe Face Mesh по отношению к эталонным ориентирам. Результаты показали, что MediaPipe Face Mesh обнаруживает лицевые ориентиры точнее, чем OpenFace. Поэтому, для дальнейшей работы на этапе обнаружения антропометрических точек в конвейере анализа микролицевых движений лица предпочтительнее использовать метод глубокого обучения. Более точные результаты на данном этапе могут в конечном итоге повысить правильность распознавания скрытых эмоций человека.

## 2.4 Подготовка набора векторов признаков для обнаружения микролицевых движений

Для обнаружения микролицевых движений очень важно локализовать области лица, в которых микро движение имело место. Области возможного возникновения микролицевых движений в диссертационной работе выбираются в соответствии со СКЛиД. Такими областями являются: правая щека, подбородок, левая щека, правая бровь, левая бровь, нос (прямая), нос (нижняя часть), правый глаз, левый глаз, губы (внутренняя часть), губы (внешняя часть).

Каждая эмоция в СКЛиД выражается движением мышц лица. Движение описывается ЕД. Выражение эмоции представляет собой комбинацию различных ЕД, но набор ЕД, которые могут иметь место для определенной эмоции, фиксирован. Например, эмоция радости в общем рассмотрении выражается в подъеме щек (ЕД 6), растяжении уголков губ (ЕД 12), совмещении верхнего и нижнего века глаза (ЕД 7), но в отдельном случае она может выражаться только ЕД 12. Также, ЕД может быть описана ее интенсивностью, то есть силой действия, которая приводит к силе выраженности изменения внешнего вида. Интенсивность оценивается в терминах А,В,С,Д,Е и обозначает силу действия от едва заметного до самого высокого уровня.

Для подготовки набора векторов признаков в качестве исходных данных использовался набор данных SAMM и файл разметки данного набора по СКЛиД, который прилагается с набором данных. Фрагмент содержимого файла разметки представлен на рисунке 35.

Subject	Filename	Inducement Code	Onset Frame	Apex Frame	Offset Frame	Duration	Micro	Action Units	Estimated Emotion	Objective Classes	Notes
006	006_1_2	1	5562	5588	5632	71	Micro - 1/2	4+7	Anger	3	
006	006_1_3	1	3912	3948	3988	77	Micro - 1/2	4	Anger	3	While blinking
006	006_1_4	1	2324	2368	2403	80	Micro - 1/2	4+7	Anger	3	
006	006_1_5	1	5343	5388	5424	82	Micro - 1/2	4+7	Anger	3	
006	006_1_6	1	7160	7197	7259	100	Micro - 1/2	4+7+43	Anger	7	Double AU4
006	006_2_4	2	180	217	274	95	Micro - 1/2	4+7	Anger	3	
006	006_3_4	3	639	668	702	64	Micro - 1/2	15A+17	Sadness	5	
006	006_3_5	3	1787	1829	1856	70	Micro - 1/2	25+26	Surprise	2	
006	006_5_10	5	1436	1465	1514	79	Micro - 1/2	L20+21	Fear	6	look to the neck
006	006_5_11	5	2130	2190	2226	97	Micro - 1/2	4	Anger	3	while blinking
006	006_5_9	5	5258	5282	5351	94	Micro - 1/2	5	Surprise	2	

Рисунок 35 – Содержимое файла разметки набора SAMM

В файле содержится информация об индексах начального, пикового и конечного кадров микролицевого движения в последовательности, наименовании эмоции, длительности микролицевого движения, выраженной в количестве кадров, ЕД и их интенсивности.

В соответствии с описанными в файле ЕД выбираются антропометрические точки, принадлежащие областям, в которых имеют место данные ЕД. В таблице 9 представлены наборы антропометрических точек в каждой из областей, значимых для обнаружения микролицевых движений. Общее количество точек - 62. Данные точки были выбраны из 468 точек лица с трехмерными координатами, аннотированные библиотекой TensorFlow фреймворка MediaPipe Face Mesh, описание которого представлено в параграфе 2.3. Сетевая топология лица TensorFlow предоставляет информации больше, чем необходимо, поэтому следует выбрать только те точки, которые необходимы для минимизирования избыточности информации без ущерба для решаемой задачи. Схема выбранных точек представлена на рисунке 36.

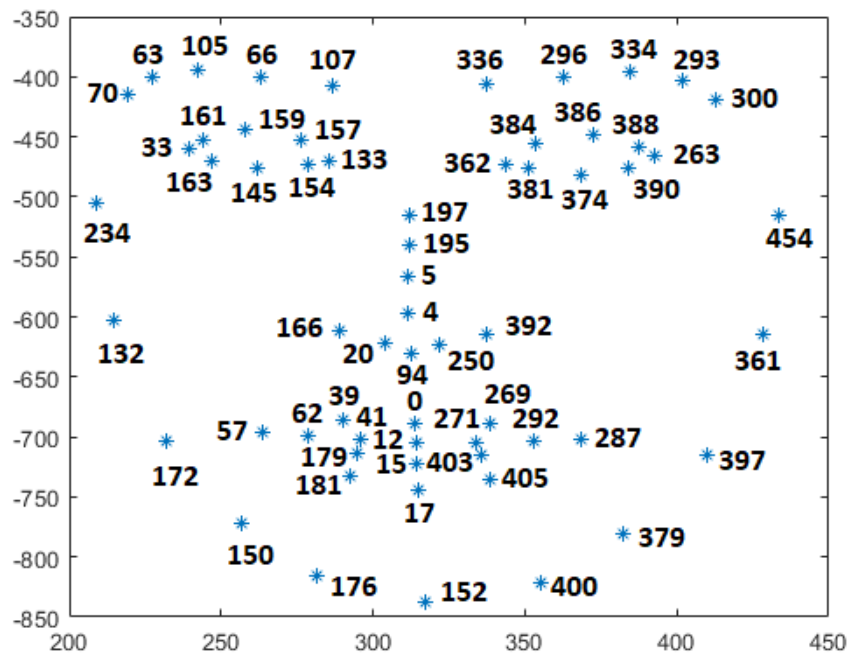


Рисунок 36 – Антропометрические точки, выбранные из множества точек TensorFlow



Описанный процесс подготовки векторов признаков позволяет точно определить наличие или отсутствие микролицевого движения в определенной области лица, что является важным шагом для точного выявления проявляемой эмоции. Но данный метод имеет недостаток, заключающийся в сложности точного позиционирования точки, подверженной движению, что может в дальнейшем привести к неверному определению ЕД и в результате - к неверной трактовке испытываемой собеседником эмоций [126].

Таблица 10 – Примеры наборов векторов признаков

Номер набора данных	ЕД	Правая щека	Подбородок	Левая щека	Правая бровь	...	Губы (внутренняя часть)	Губы (внешняя часть)	Эмоция
006_3_5	25+ 26	00000	0	00000	00000	...	00000111	11111111	Удивление
006_5_10	L20+ 21	00000	0	00000	00000	...	11100000	11000000	Страх
007_6_2	12	00000	0	00000	00000	...	10001101	11001101	Радость
018_5_1	17	00000	1	00000	00000	...	00000000	00000000	Грусть

## 2.5 Выводы по второй главе

Во второй главе рассмотрены алгоритмы выбора областей возможного проявления микролицевого движения. Для локализации областей лица и глаз были рассмотрены алгоритм Виолы-Джонса и алгоритм поиска пикселей кожи в YCrCb цветовом пространстве. Алгоритм Виолы-Джонса выполняет поиск областей с помощью признаков Хаара, в то время как второй алгоритм основан на яркостной характеристике кожи. По результатам экспериментов оба алгоритма показали достаточно хорошие результаты для решения представленной задачи. Алгоритм Виолы-Джонса показывает приемлемые результаты при работе как с черно-белыми, так и с цветными изображениями, но при повороте головы

определяет только один глаз, в то время как алгоритм, основанный на поиске пикселей кожи в цветовом пространстве YCrCb, способен выполнить поиск областей только на изображениях, в которых цвет кожи контрастирует с цветом фона, при повороте головы обнаруживает оба глаза.

Так как глаза являются одними из ключевых маркеров скрываемых или подавляемых эмоций, следующей задачей, рассмотренной в данной главе, являлась задача поиска антропометрических точек вокруг глаз. При решении данной задачи были рассмотрены уже разработанные классические алгоритмы, результаты выполнения которых представлены в параграфе 2.2. В связи с тем, что полученные результаты оказались не вполне удовлетворительными для решения рассматриваемой задачи, был разработан собственный алгоритм поиска антропометрических точек вокруг глаз, который показал в 2,2 раза меньшее среднеквадратичное отклонение вычисленных координат точек по сравнению с найденными алгоритмом ASM координатами точек от координат антропометрических точек вокруг глаз образцового изображения.

Для улучшения показателей поиска антропометрических точек были рассмотрены современные методы, среди которых библиотека TensorFlow фреймворка MediaPipe Face Mesh, основанная на алгоритме глубокого обучения, и библиотека dlib, основанная на алгоритме машинного обучения. В результате экспериментов были сделаны выводы, что библиотека TensorFlow определяет ориентиры точнее относительно искомой позиции по сравнению с библиотекой dlib. Поэтому для выполнения этапа поиска антропометрических точек в конвейере определения микролицевых движений была выбрана библиотека TensorFlow.

Для повышения точности определения микролицевых движений описан процесс подготовки набора векторов признаков, благодаря которому возможно локализовать конкретную точку, участвующую в микродвижении. Таким образом, это позволит более точно определить проявляемые ЕД, что повлечет за собой более точную классификацию скрываемых эмоций.



### **3 Модификация алгоритмов пространственно-временных дескрипторов признаков**

#### **3.1 Подходы к выбору алгоритмов для обнаружения микролицевых движений**

Извлечение признаков играет важную роль в обработке и распознавании изображений, в частности – в задачах обнаружения микролицевых движений. Данные задачи подразделяются на три подзадачи: извлечение антропометрических точек, извлечение информативных признаков и их классификация.

На основе выводов, представленных в главе 2, в диссертационной работе для конвейера обнаружения микродвижений лица извлечение антропометрических точек реализуется с помощью фреймворка TensorFlow. Извлеченные точки используются в качестве центров областей, основанных на областях СКЛиД, и необходимы для дальнейшего обнаружения микролицевых движений. Количество точек - 62.

Подготовительным этапом для извлечения информативных признаков является этап вычисления трехмерных блоков [127] для каждой из 62 вычисленных точек каждой последовательности изображений. Данный этап был реализован с помощью одного из следующих алгоритмов. В обоих алгоритмах выбирается квадратная окрестность размером  $65 \times 65$  на каждом изображении последовательности для создания одного слоя блока. В алгоритме «Прямоугольные параллелепипеды» центр квадрата постоянно находится в положении антропометрической точки на первом изображении последовательности. Таким образом, блок представляет собой прямоугольный параллелепипед, вырезанный из последовательности изображений, рассматриваемый как трехмерный массив пикселей размером  $65 \times 65 \times N$ , где  $N$  - это количество изображений в последовательности. В алгоритме «Скользящие блоки» центр квадрата находится в положении рассматриваемой антропометрической точки на текущем изображении. Таким образом, окрестность движется или

скользит относительно данного положения антропометрической точки. Затем все квадратные слои соединяются вместе, чтобы сгенерировать трехмерный блок того же самого размера, как в первом алгоритме. Один или другой тип построения трехмерного блока может быть использован в зависимости от специфики. Если лицо в целом не очень сильно сдвигается, то предпочтительно использовать «Прямоугольные параллелепипеды», в противном случае – алгоритм «Скользящие блоки».

Этап извлечения признаков является одним из самых важных и заключается в вычислении дескрипторов признаков по выбранному алгоритму. Это позволяет преобразовать большой объем информации, содержащейся в некоторой области изображения или в последовательности фрагментов изображений, в сравнительно небольшой вектор признаков, описывающий наиболее существенные для решаемой задачи характеристики рассматриваемой области. Таким образом, пространственно-временной дескриптор признаков можно рассматривать как некоторую функцию  $f(s)$ , применяемую к последовательности изображений  $S$ , результатом выполнения которой является вектор признаков  $\vec{V}$ , содержащий пространственную и временную информацию.

Выбор дескриптора зависит от специфики решаемой задачи. В настоящей работе рассматривается задача обнаружения микролицевых движений по последовательности изображений лица человека. Наиболее широко используемыми алгоритмами для решения данной задачи являются: алгоритм вычисления дескрипторов признаков локальных бинарных шаблонов по трем ортогональным плоскостям LBP-TOР, алгоритм гистограммы направленного оптического потока HOOF, алгоритм трехмерной гистограммы направленных градиентов 3DНОG. В диссертационной работе рассмотрены алгоритм LBP-TOР и вариации алгоритма HOOF.

### 3.2 Исследование влияния значений параметров алгоритма LBP-TOP на диагностическую эффективность обнаружения микролицевых движений

Основными задачами исследования алгоритма LBP-TOP являются:

- исследование влияния изменения настраиваемых параметров алгоритма LBP-TOP на диагностическую эффективность обнаружения микролицевых движений;
- модификация траектории расположения пикселей в окрестности вокруг центрального пикселя для плоскостей  $XOT$ ,  $YOT$  со стандартной круговой на эллиптическую;
- исследование влияния изменения траектории пикселей в окрестности вокруг центрального пикселя на диагностическую эффективность обнаружения микролицевых движений;
- формирование рекомендаций по подбору значений параметров алгоритма LBP-TOP для повышения диагностической эффективности обнаружения микролицевых движений.

Дескриптор LBP-TOP учитывает информацию в пространственной и временной областях. Рассматриваются три ортогональные плоскости  $XOY$ ,  $XOT$ ,  $YOT$ , для каждой из которых вычисляется дескриптор LBP, что позволяет учесть изменения, происходящие от кадра к кадру.

Принцип формирования дескриптора LBP-TOP представлен на рисунке 37.

Шаги алгоритма вычисления дескриптора LBP-TOP представлены ниже.

Шаг 1. На вход поступает трехмерный блок, сформированный из последовательности изображений.

Шаг 2. Преобразование трехмерного блока в три двумерных блока соответственно для каждой из плоскостей  $XOY$ ,  $XOT$ ,  $YOT$ .

Шаг 3. Вычисление дескриптора LBP для каждого из трех полученных на предыдущем шаге двумерных блоков.

Для вычисления дескриптора LBP для каждого пикселя изображения определяется количество пикселей в окрестности центрального пикселя,

расположенных по кругу.  $P$  – количество выбранных пикселей,  $R$  – радиус окружности, на которой на равных расстояниях друг от друга располагаются указанные  $P$  пикселей, как показано на рисунке 38. Если точка не попадает в центр, используется билинейная интерполяция.

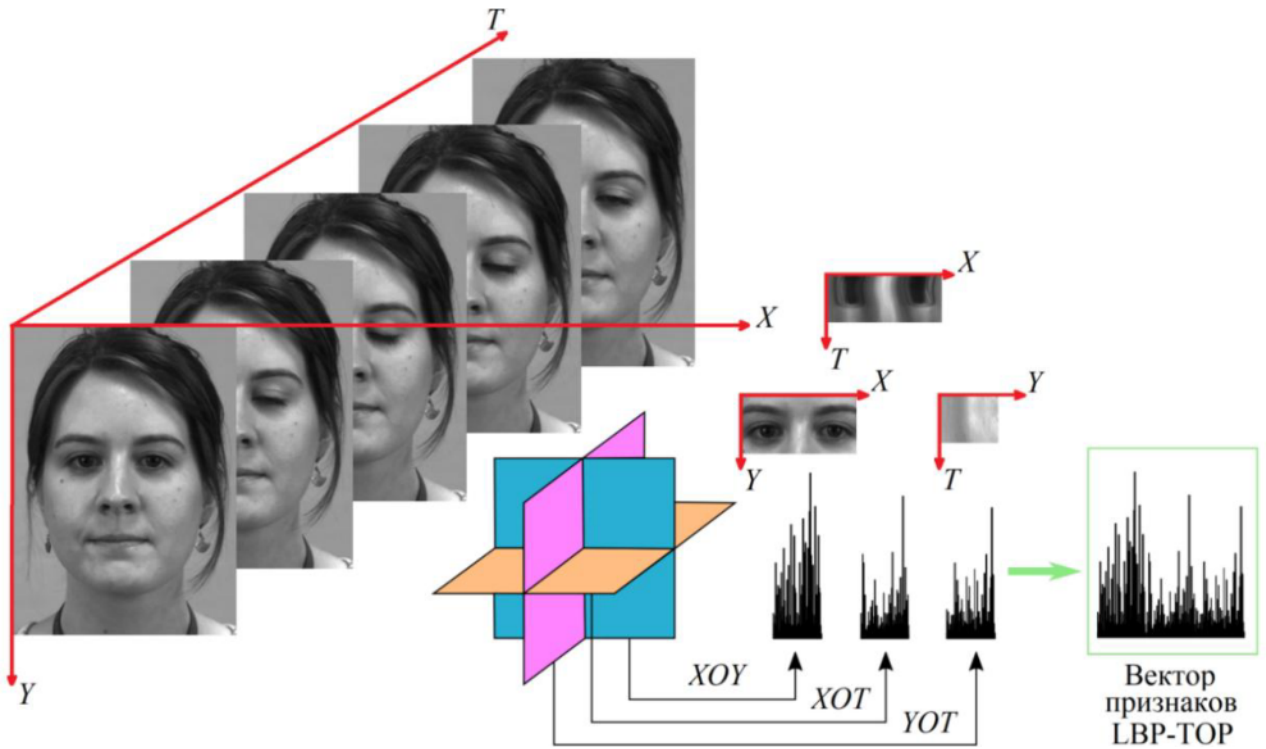


Рисунок 37 – Визуальное представление принципа формирования дескриптора LBP-TOP

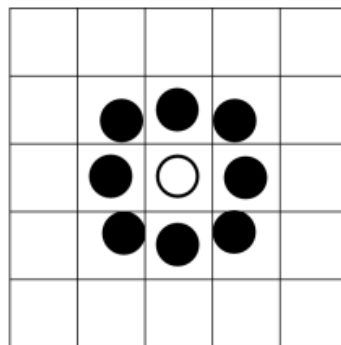


Рисунок 38 – Определение пикселей в окрестности вокруг центрального пикселя для параметров  $P=4$ ,  $R=1$

На рисунке 39 изображен процесс формирования локального бинарного шаблона в окрестности одного пикселя для параметров  $P=8$ ,  $R=1$ . Обозначим значение яркости центрального пикселя  $g_c$ , а яркости пикселей, расположенных на окружности –  $g_p$ , где  $p=0, 1, \dots, P-1$ . Тогда формирование локального бинарного шаблона можно описать следующими формулами (24, 25):

$$B = b(s(g_0 - g_c), s(g_1 - g_c), \dots, s(g_{P-1} - g_c)), \quad (24)$$

$$\text{где } s(n) = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0; \end{cases}$$

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p. \quad (25)$$

130	117	105
133	125	108
135	123	110

а

1	0	0
1		0
1	0	0

б

1	2	4
128		8
64	32	16

в

Рисунок 39 – Пример вычисления локального бинарного шаблона в окрестности одного пикселя для параметров  $P=8$ ,  $R=1$ , полученное значение равно  $11000001_2$  или  $193_{10}$ : а – яркости пикселей в окрестности, б – результат сравнения с порогом, в – весовые коэффициенты

По формуле (24) можно рассчитать бинарное значение, по формуле (25) – десятичное. После вычисления десятичных значений для всех пикселей изображения или его рассматриваемого фрагмента строится гистограмма всех этих значений. Гистограмма представляет собой вектор признаков LBP [128]. Схема алгоритма вычисления дескриптора LBP представлена на рисунках 40 и 41.

При вычислении пикселей в окрестности вокруг центрального пикселя в плоскостях YOT, XOT рассматривается эллиптическая траектория расположения пикселей в окрестности вокруг центрального пикселя, так как в пределах одних и

тех же временных интервалов текстура может резко меняться при высоком разрешении кадров и низкой частоте записи.

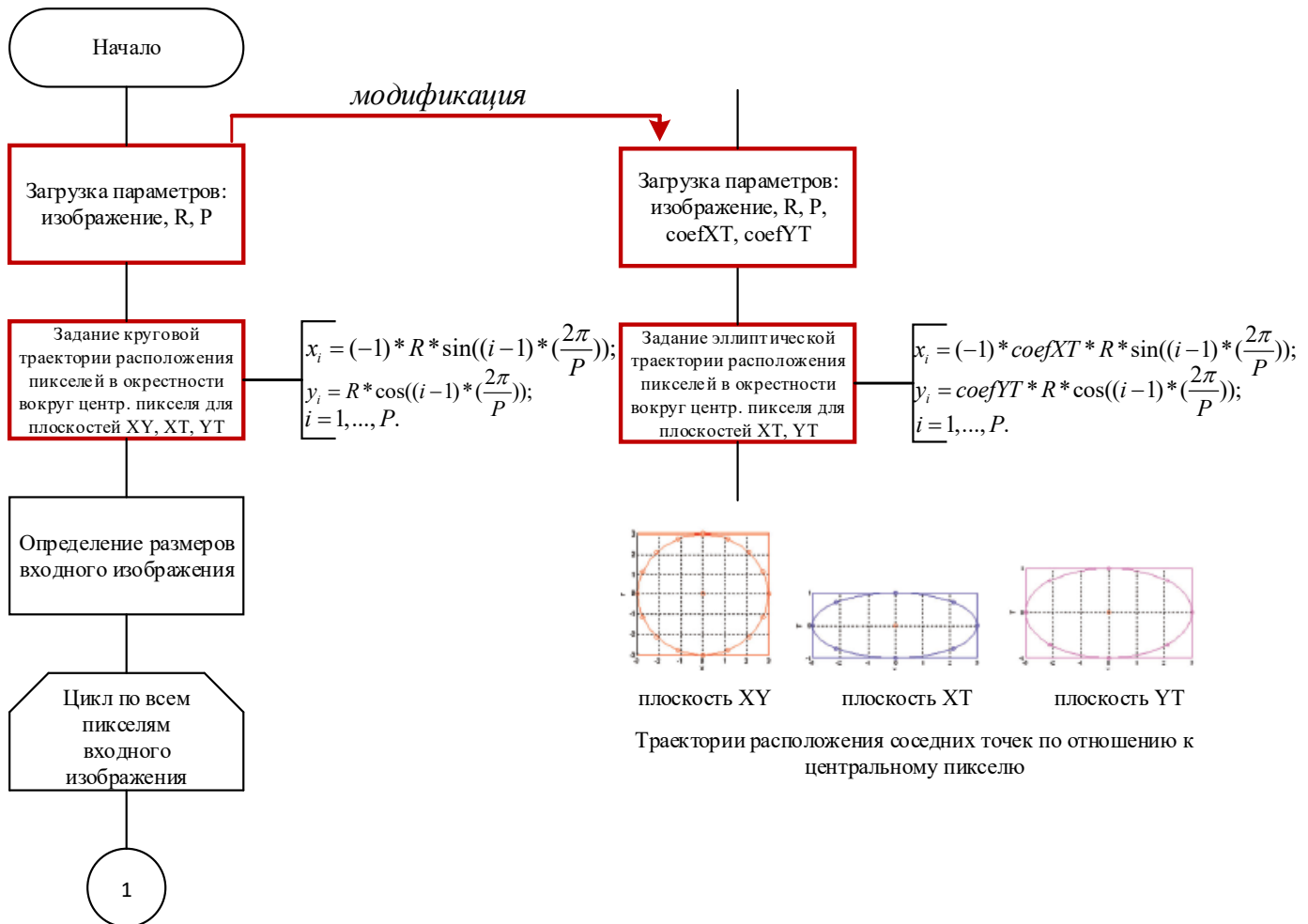


Рисунок 40 – Схема алгоритма LBP

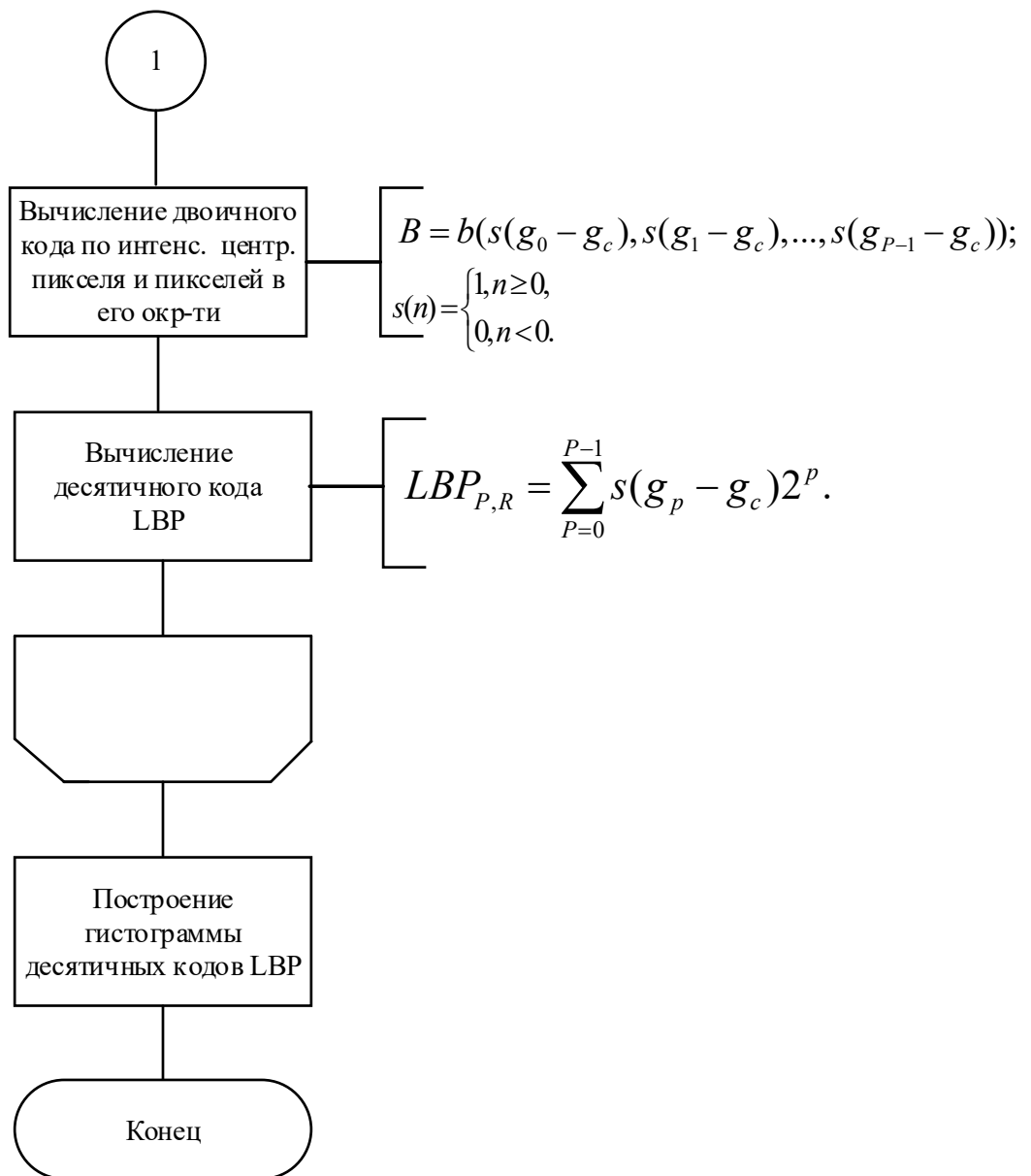


Рисунок 41 – Схема алгоритма LBP (продолжение)

Шаг 4. Преобразование каждого из полученных трех векторов признаков LBP в трехмерный блок для дальнейшего вычисления гистограмм.

Шаг 5. Разбиение полученных трехмерных блоков на меньшие блоки в соответствии с входными параметрами алгоритма LBP-TOP.

Шаг 6. Каждый меньший блок разворачивается в вектор, по которому вычисляется гистограмма с ранее заданным количеством интервалов.

Шаг 7. Конкатенация вычисленных по меньшим блокам гистограмм в единую гистограмму признаков по каждой из плоскостей  $XOY$ ,  $XOT$ ,  $YOT$ .

Шаг 8. Конкатенация вычисленных трех гистограмм по ортогональным плоскостям в единую гистограмму признаков LBP-TOP.

На данном шаге вся полученная локальная информация по трем плоскостям объединяется и, таким образом, получается статистика по  $XOY$ ,  $XOT$ ,  $YOT$  плоскостям для формирования информации о микролицевом движении и выражении лица.

Гистограмма выражается формулой (26):

$$H_{i,j} = \sum_{x,y,t} I\{f_j(x,y,t) = i\}, i = 0, \dots, n_j - 1; j = 0, 1, 2, \quad (26)$$

где  $n_j$  – количество разных отметок, созданных алгоритмом LBP на  $j$  плоскости, где  $j = 0$  для  $XOY$ ,  $j = 1$  для  $XOT$ ,  $j = 2$  для  $YOT$ ;  $f_j(x, y, t)$  выражает LBP код центрального пикселя на  $j$ -й плоскости [129].

Схема алгоритма LBP-TOP представлена на рисунках 42 и 43.

Данный метод упрощает извлечение признаков при большем количестве пикселей в окрестности вокруг центрального пикселя, устойчив к ошибкам в выравнивании лица [7].



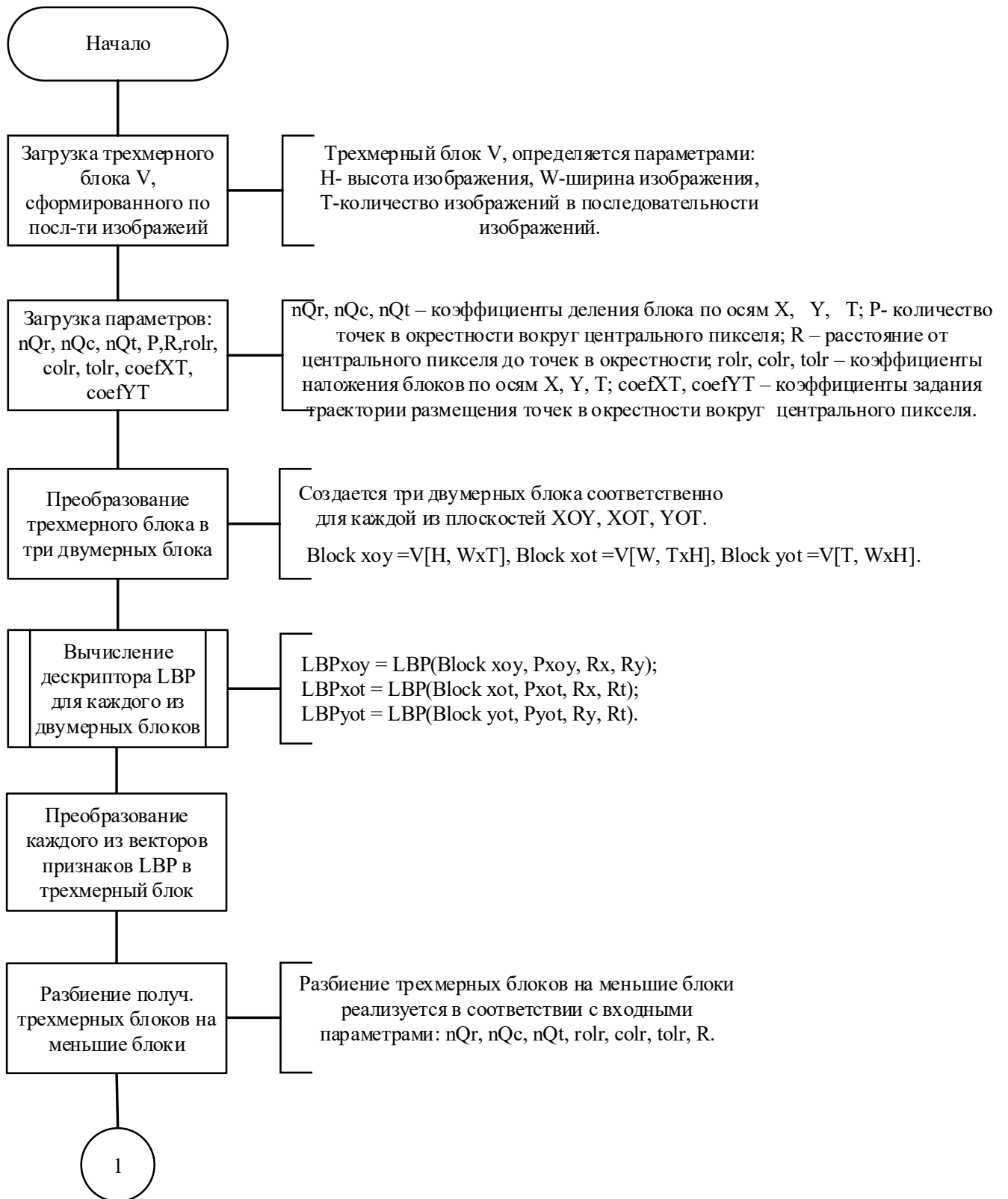


Рисунок 42 – Схема алгоритма LBP-TOP

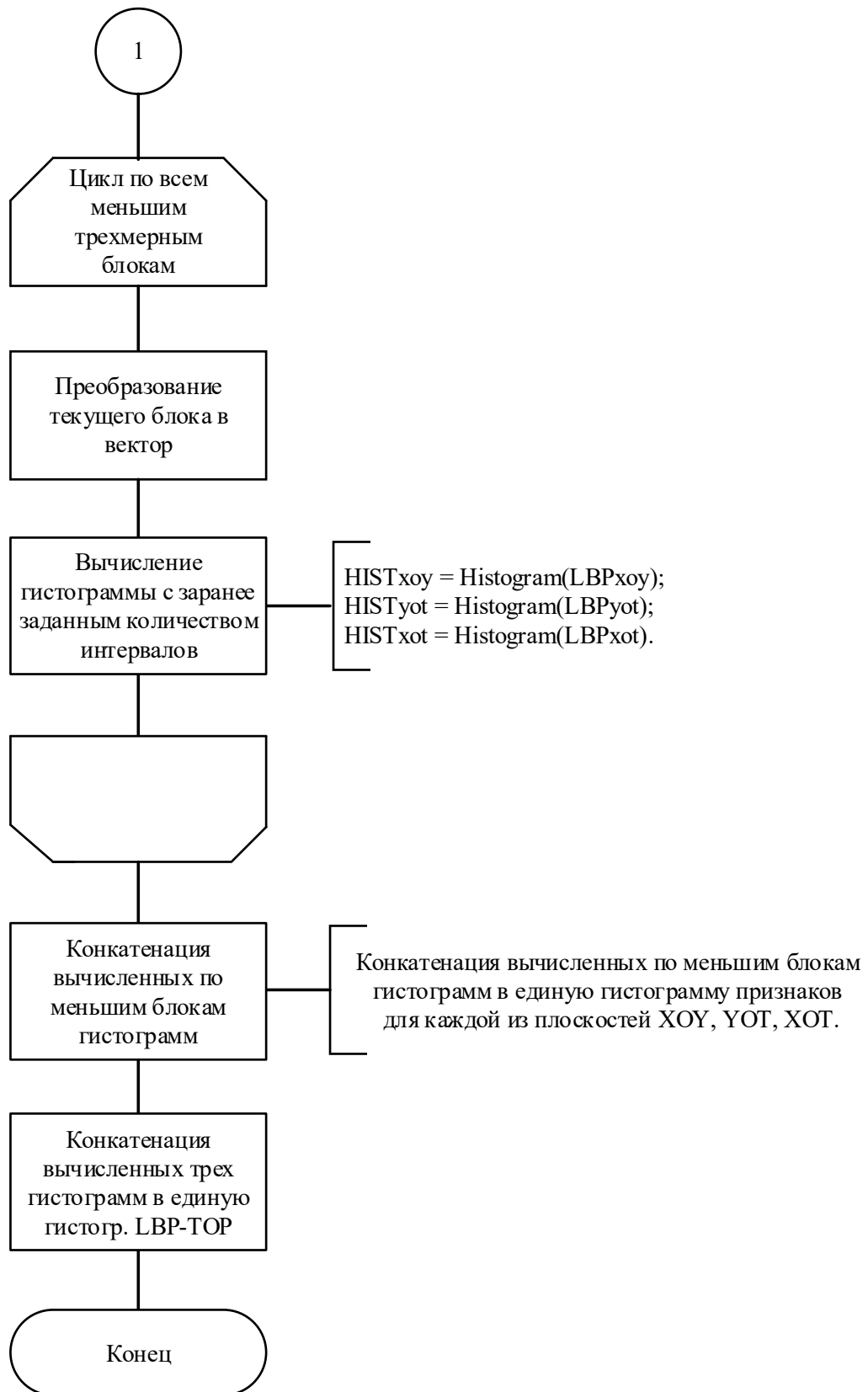


Рисунок 43 – Схема алгоритма LBP-TOP (продолжение)

В качестве исходных данных для проведения исследования, позволяющего сравнить диагностическую эффективность обнаружения микролицевых движений при использовании стандартного и модифицированного алгоритмов LBP-TOP, выбран набор данных SAMM. Около 30% набора было размечено с точки зрения наличия или отсутствия микролицевого движения в каждой из антропометрических точек на основе списка комбинаций единиц движений, прилагаемого к набору данных. Разметка набора данных реализовывалась в соответствии с алгоритмом, описанным в параграфе 2.4. Макролицевые движения, например, моргание или поворот головы, отнесены к классу отсутствия микролицевых движений. Оценка диагностической эффективности обнаружения микролицевых движений выполнялась с помощью классификаторов SVM и MLP. Поэтому, размеченный набор использовался для получения обучающего и тестового наборов данных в следующем соотношении: 80% данных служат для обучения, 20% данных - для тестирования. Задача классификации состояла в определении, присутствовало ли или отсутствовало микролицевое движение в области рассматриваемой антропометрической точки.

Настраиваемыми параметрами алгоритма LBP-TOP являются:

- количество выбранных пикселей вокруг центрального пикселя;
- расстояние от центрального пикселя до выбранного пикселя;
- коэффициенты деления блоков изображений на блоки меньшего размера;
- коэффициенты перекрытия блоков между собой;
- коэффициенты расчета траектории расположения пикселей в окрестности вокруг центрального пикселя для каждой из плоскостей.

Исходные значения перечисленных выше параметров:

- количество пикселей в окрестности вокруг центрального пикселя  $P=8$ ;
- расстояние от центрального пикселя до выбранного пикселя  $R=1$ ;
- коэффициенты деления блоков изображений на блоки меньшего размера отсутствуют;
- коэффициенты перекрытия блоков между собой отсутствуют;

– траектория размещения выбранных пикселей вокруг центрального пикселя – круговая для каждой из плоскостей.

При исходных значениях входных параметров алгоритма LBP-TOP и алгоритма формирования блоков «Прямоугольные параллелепипеды» диагностическая эффективность обнаружения микролицевых движений на части набора данных SAMM изменяется от 55% до 89%. Диагностическая эффективность представляет собой значение Accuracy обнаружения микролицевых движений, вычисляемое по формуле (27):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} , \quad (27)$$

где  $TP$  – «True Positive», количество случаев корректного обнаружения наличия микролицевых движений;  $TN$  – «True Negative», количество случаев корректного обнаружения отсутствия микролицевых движений;  $FP$  – «False Positive», количество случаев ложного обнаружения наличия микролицевых движений при их фактическом отсутствии;  $FN$  – «False Negative», количество случаев ложного обнаружения отсутствия микролицевых движений при их фактическом наличии [130].

Рассмотрим влияние изменения настраиваемых параметров алгоритма вычисления дескриптора LBP-TOP на диагностическую эффективность обнаружения микролицевых движений на части набора данных SAMM.

Первым исследуемым параметром алгоритма является количество выбранных пикселей  $P$  вокруг центрального пикселя. Влияние изменения данного параметра и параметра расстояния  $R$  от центрального пикселя до пикселей в его окрестности на диагностическую эффективность обнаружения микролицевых движений представлено в таблице 11.

Можно заметить, что увеличение величины  $P$  при использовании в качестве классификатора SVM существенным образом не сказалось на результате. Изменение расстояния  $R$  от центрального пикселя до пикселей в его окрестности

привело к увеличению диагностической эффективности алгоритма при обнаружении микролицевых движений только при использовании ядра Sigmoid.

Таким образом, по таблице 11 можно проанализировать влияние обоих настраиваемых параметров алгоритма LBP-TOP на диагностическую эффективность обнаружения микролицевых движений. Из полученных результатов следует, что изменение стандартных значений обоих параметров не привело к значительному повышению диагностической эффективности обнаружения микролицевых движений.

Таблица 11 – Результаты, полученные при изменении количества выбранных пикселей вокруг центрального пикселя и расстояния от центрального пикселя до пикселей в его окрестности

Классификатор	Тип ядра	$R=1$			$R=2$			$R=3$		
		$P=8$	$P=12$	$P=16$	$P=8$	$P=12$	$P=16$	$P=8$	$P=12$	$P=16$
SVM	Linear	74%	74%	74%	74%	74%	74%	74%	74%	74%
	RBF	89%	92%	76%	87%	74%	74%	84%	89%	74%
	Sigmoid	55%	50%	50%	76%	74%	50%	76%	74%	74%
MLP		95%	98%	90%	99%	98%	97%	98%	98%	96%

Рассмотрим изменение коэффициентов деления блоков изображений по каждому из направлений, где  $n_{Qr}$  – коэффициент деления блока по оси  $X$ ;  $n_{Qc}$  – коэффициент деления блока по оси  $Y$ ;  $n_{Qt}$  – коэффициент деления блока по оси  $T$ . Было принято, что количество выбранных пикселей  $P=8$ , расстояние от центрального пикселя до выбранного пикселя  $R=1$ . Перекрытие между результирующими блоками отсутствует. Результаты оценки диагностической эффективности обнаружения микролицевых движений при изменении коэффициентов деления блоков алгоритма LBP-TOP представлены в таблице 12.

Из полученных результатов видно, что добавление деления блоков по каждому из направлений увеличивает диагностическую эффективность до 95%. Наилучший результат получается при следующих значениях входных параметров:  $n_{Qr}=5$ ,  $n_{Qc}=5$ ,  $n_{Qt}=5$ . При использовании классификатора MLP диагностическая

эффективность обнаружения микролицевых движений составляет 98% при отсутствии деления блоков, то есть  $nQr = 1$ ,  $nQc = 1$ ,  $nQt = 1$ .

Рассмотрим влияние изменения коэффициентов перекрытия блоков ( $rolr$  – коэффициент перекрытия блоков по горизонтали;  $colr$  – коэффициент перекрытия блоков по вертикали;  $tolr$  – коэффициент перекрытия блоков по диагонали) алгоритма LBP-TOP на диагностическую эффективность обнаружения микролицевых движений. Было принято, что количество выбранных пикселей  $P = 8$ , расстояние от центрального пикселя до выбранного пикселя  $R = 1$ . Исходные блоки изображений на меньшие блоки не делились. Результаты влияния изменения коэффициентов перекрытия блоков на диагностическую эффективность представлены в таблице 12.

Из полученных результатов видно, что добавление перекрытия между блоками также увеличивает диагностическую эффективность обнаружения микролицевых движений до 95%. Следует отметить, что перекрытие блоков или его отсутствие при использовании классификатора MLP не влияет на диагностическую эффективность, которая составляет 98%.

Рассмотрим изменение коэффициентов расчета траектории расположения пикселей в окрестности вокруг центрального пикселя для плоскостей  $XOT$  и  $YOT$ . Так как в пределах непродолжительного временного интервала возможно появление резкого микродвижения мышц лица, следует назначать разные радиусы в пространственной и во временной областях для выбора пикселей вокруг центрального пикселя. При данном подходе круговая траектория преобразуется в эллиптическую траекторию. Изменение данного параметра проводилось при следующих значениях входных параметров алгоритма LBP-TOP:

- количество выбранных пикселей  $P = 8$ ;
- расстояние от центрального пикселя до выбранного пикселя  $R = 1$  для плоскости  $XOY$ ;
- коэффициенты разбиения блока на меньшие блоки:  $nQr = 4$ ,  $nQc = 4$ ,  $nQt = 3$ ;
- коэффициенты перекрытия блоков между собой:  $rolr = 0,1$ ;  $colr = 0,1$ ;  $tolr = 0,1$ .

Результаты вычисления диагностической эффективности обнаружения микролицевых движений на части набора данных SAMM при изменении коэффициентов расчета траектории расположения пикселей в окрестности вокруг центрального пикселя алгоритма LBP-TOP представлены в таблице 12.

Таблица 12 – Результаты, полученные при изменении коэффициентов деления блоков, перекрытия блоков и расчета траектории расположения пикселей вокруг центрального пикселя алгоритма LBP-TOP

Настраиваемые параметры алгоритма LBP-TOP	Тип ядра классификатора SVM			MLP
	Linear	RBF	Sigmoid	
Коэффициенты деления блоков				
$n_{Qr} = 1, n_{Qc} = 1, n_{Qt} = 1$	74%	89%	55%	98%
$n_{Qr} = 3, n_{Qc} = 1, n_{Qt} = 1$	81%	87%	60%	97%
$n_{Qr} = 3, n_{Qc} = 2, n_{Qt} = 1$	89%	92%	66%	96%
$n_{Qr} = 3, n_{Qc} = 2, n_{Qt} = 2$	92%	89%	66%	97%
$n_{Qr} = 3, n_{Qc} = 3, n_{Qt} = 2$	95%	89%	74%	97%
$n_{Qr} = 4, n_{Qc} = 4, n_{Qt} = 3$	92%	92%	76%	95%
$n_{Qr} = 5, n_{Qc} = 5, n_{Qt} = 5$	95%	92%	76%	95%
Коэффициенты перекрытия блоков между собой	Linear	RBF	Sigmoid	MLP
$rolr = 0, colr = 0, tolr = 0$	74%	89%	55%	98%
$rolr = 0,1, colr = 0, tolr = 0$	92%	92%	76%	98%
$rolr = 0,1, colr = 0,1, tolr = 0$	95%	92%	74%	98%
$rolr = 0,2, colr = 0,1, tolr = 0$	95%	92%	74%	98%
$rolr = 0,1, colr = 0,1, tolr = 0,1$	95%	92%	74%	98%
Коэффициенты расчета траектории расположения пикселей вокруг центрального пикселя	Linear	RBF	Sigmoid	MLP
$coefXT = 0,1, coefYT = 0,3$	95%	92%	76%	95%
$coefXT = 0,1, coefYT = 0,5$	95%	92%	79%	96%
$coefXT = 0,1, coefYT = 0,7$	92%	92%	79%	98%
$coefXT = 0,1, coefYT = 0,9$	92%	92%	79%	97%
$coefXT = 0,2, coefYT = 0,4$	95%	92%	74%	98%
$coefXT = 0,3, coefYT = 0,6$	95%	92%	74%	96%
$coefXT = 0,4, coefYT = 0,9$	95%	92%	76%	96%

Анализ полученных результатов позволяет сделать вывод, что наибольшее влияние на диагностическую эффективность обнаружения микролицевых движений оказывают следующие параметры алгоритма LBP-TOP:

- коэффициенты деления блоков на блоки меньшего размера;
- коэффициенты перекрытия блоков между собой;
- коэффициенты изменения траектории расположения пикселей в окрестности вокруг центрального пикселя.

Разработана методика подбора значений настраиваемых параметров для модифицированного алгоритма LBP-TOP, представленная далее.

1. Параметры алгоритма LBP-TOP: расстояние от центрального пикселя до соседних пикселей ( $R$ ), количество соседних пикселей вокруг центрального пикселя ( $P$ ) значительно не повышают диагностическую эффективность обнаружения микролицевых движений, поэтому данные значения параметров  $R=1, P=8$  можно оставить без изменений.

2. Установить величину деления блоков, по умолчанию приняты следующие значения: по оси  $X$ :  $nQr = 1$ , по оси  $Y$ :  $nQc = 1$ , по оси  $T$ :  $nQt = 1$ , где  $nQr, nQc, nQt \in N$ .

3. Установить коэффициенты перекрытия блоков между собой по отношению к величине не перекрытого блока по оси  $X$ :  $rolr = 0$ , по оси  $Y$ :  $colr = 0$ , по оси  $T$ :  $tolr = 0$ , где  $nQr, nQc, nQt \in N$ .

4. Установить коэффициенты расчета траектории расположения пикселей вокруг центрального пикселя для плоскости  $XOT$ :  $coefXT = 0.1$ , для плоскости  $YOT$ :  $coefYT = 0.3$ , где  $rolr, colr, tolr \in R$ .

5. Если выполнены шаги 1-4, но существенное увеличение диагностической эффективности на выбранном наборе данных не обнаружено, следует повторить шаг 2, но с увеличением величины деления блоков по каждой из осей на величину шага, равного 1.



6. Если изменение, внесенное на шаге 5, не повлияет на результат выполнения алгоритма, увеличить значения коэффициентов перекрытия блоков на 0,1.

7. Изменить траекторию размещения соседних точек вокруг центрального пикселя по направлениям  $XТ$ ,  $YТ$ , причем для  $coefXТ$  установить меньшее значение, чем для  $coefYТ$ , по причине более высокой скорости изменений параметров во временной плоскости.

При корректном задании параметров алгоритма извлечения информативных признаков оба классификатора достаточно эффективно справляются с задачей обнаружения микролицевых движений.

Получены следующие результаты исследования:

– проведено экспериментальное исследование влияния выбора настраиваемых параметров алгоритма вычисления пространственно-временного дескриптора LBP-TOP на диагностическую эффективность обнаружения микролицевых движений;

– выполнена модификация алгоритма LBP-TOP, состоящая в замене траектории расположения пикселей в окрестности вокруг центрального пикселя с круговой на эллиптическую;

– показано, что изменение настраиваемых параметров алгоритма LBP-TOP оказывает влияние на диагностическую эффективность обнаружения микролицевых движений;

– разработана методика подбора параметров алгоритма LBP-TOP для повышения диагностической эффективности обнаружения микролицевых движений.

Эксперименты доказывают, что модифицированный алгоритм LBP-TOP обеспечивает достаточные результаты для решения задачи обнаружения микролицевых движений при правильной настройке параметров алгоритма. Перспективно применение модифицированной реализации с эллиптическими траекториями вместо стандартных круговых.

### 3.3 Сравнительный анализ алгоритмов дескрипторов FHOOF и FHOFO

В параграфе 3.1 рассмотрен дескриптор локального бинарного шаблона по трем ортогональным плоскостям для обнаружения микролицевых движений. Для данной цели также может быть использован другой пространственно-временной дескриптор, а именно – дескриптор гистограмм ориентированного оптического потока (HOOF).

Ниже представлены шаги алгоритма вычисления пространственно-временного дескриптора HOOF для фрагмента изображения  $I$  размером  $N \times M$ .

Шаг 1. Вычислить вектор оптического потока  $v = [v_x, v_y]^T$  для каждого местоположения пикселя изображения.

Шаг 2. Вычислить величину движения  $(\rho = \sqrt{v_x^2 + v_y^2})$  и направление движения  $(-\pi \leq \theta \leq \pi)$  для каждого пикселя на основе векторов оптического потока, полученных на шаге 1, путем преобразования из евклидовых координат в полярные. Таким образом, для каждого пикселя будет получена пара из величины и направления движения  $(\rho_i, \theta_i)$ .

Шаг 3. Сгруппировать все векторы потока в соответствии с углом поворота и взвесить в соответствии с их величиной.

Шаг 4. Вычислить гистограмму с количеством интервалов  $n$  на основе векторов оптического потока.

Гистограммы генерируются на основе направления оптического потока, поэтому если вектор расположен близко к границе интервалов гистограммы, то он оказывает влияние только на собственный интервал гистограммы (например, вектора в 5-м и 6-м интервалах гистограммы, изображенные на рисунке 6).

Таким образом, дескрипторы НООФ для фрагмента изображения  $I$  размером  $N \times M$  представлены формулой (28):

$$H(I) = [h_1, h_2, \dots, h_n], \quad h_i = \frac{h'_i}{\sum_{k=1}^n h'_k}, \quad h'_i = \sum_{\forall \theta_j \in (i^{\text{th}} \text{bin})} \rho_j, \quad (28)$$

где  $n$  – количество интервалов гистограммы,  $\rho_j$  – величина движения,  $\theta_j$  – направление движения в местоположении  $j$ -го пикселя.

В формуле (28)  $h'_i$  представляет собой значение  $i$ -го интервала гистограммы направления движения, взвешенного по соответствующим величинам движения.

$h'_i$  можно записать формулой (29):

$$h'_i = \sum_{j=1}^{M \times N} \rho_j \mu_{ij}, \quad (29)$$

где  $\mu_{ij}$  – значение принадлежности к интервалу, которое определяет принадлежность  $\theta_j$  к  $i$ -му интервалу.  $\mu_{ij} \in [0; 1]$ .

Границы диапазона четко определены для назначения интервалов,  $\mu_{ij}$  определяется формулой (30):

$$\mu_{ij} = \begin{cases} 1, & \text{если } -\pi + \frac{2\pi}{n}(i-1) \leq \theta_j < -\pi + \frac{2\pi}{n}i, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (30)$$

Дескрипторы НООФ объединяют по всем фрагментам изображения для получения модели движения, но, если модели движения похожи, они попадут в

один интервал гистограммы, в противном случае – в разные интервалы, и модели считаются различными. Несмотря на то что вектора 5 и 6 (Рисунок 7) расположены очень близко, величина углового порога определит их в разные интервалы. Чтобы не связывать величину углового порога с интервалом, был разработан нечеткий подход, который позволяет определить вектора оптического потока в оба интервала. Таким образом, вектор 6, который ранее вносил вклад только в 6-й интервал, при нечетком подходе будет вносить вклад как в 5, так и в 6 интервалы. На основе этого получается плавная гистограмма переходов.

Для повышения эффективности алгоритма НООФ были разработаны его нечеткие модификации, а именно – алгоритмы дескрипторов FНООФ и FНОФО. При данном подходе гистограммы имеют свойство круговой непрерывности. Вклад угла распространяется в окружающие его интервалы на основе функции принадлежности. Для каждого направления вычисляются значения принадлежности  $\mu_{ij}$  для каждого интервала. Поскольку данная процедура требует больших вычислительных ресурсов, сначала создается гистограмма с большим количеством интервалов, затем преобразуется в гистограмму с меньшим количеством интервалов.

Ниже представлены шаги алгоритмов вычисления дескрипторов FНООФ и FНОФО.

Шаг 1 и шаг 2 соответствуют алгоритму вычисления дескриптора НООФ, представленному ранее.

Шаг 3. Сгруппировать все векторы потока в соответствии с углом поворота и добавить веса в соответствии с величинами этих векторов.

Принадлежность вектора потока  $\theta_j$  к  $i$ -му интервалу гистограммы определяется путем распределения значения принадлежности  $\theta_j$  к  $i$ -му интервалу гистограммы с использованием распределения Гаусса. Оно присваивает более высокие значения принадлежности тем гистограммам, что более близки к центрам интервалов, в то время как значения принадлежности постепенно уменьшаются по мере удаления интервалов от центра. Дисперсия функции принадлежности по

Гауссу определяет вес различных интервалов гистограммы. Чем больше значение дисперсии, тем больше будет вклад угла в соседние интервалы гистограммы. Чем меньше дисперсия, тем меньший вклад в соседние интервалы гистограммы.

Значение принадлежности  $\theta_j$  к  $i$ -му интервалу гистограммы, определяется формулой (31):

$$\mu_{ij} = \exp\left(\frac{-(\theta_j - c_i)^2}{2\sigma^2}\right), \quad (31)$$

где  $c_i$  – центр  $i$ -го интервала,  $\sigma$  – дисперсия функции принадлежности Гаусса.

Шаг 4. Вычислить гистограмму с количеством интервалов  $n'$  на основе векторов оптического потока.

Построение векторов признаков для дескриптора FHOOF по изображению  $I$  размером  $N \times M$  имеет следующий вид (32, 33):

$$F(I) = [f_1, f_2, \dots, f_n], f_i = \frac{f'_i}{\sum_{k=1}^n f'_k}, \quad (32)$$

$$f'_i = \sum_{j=1}^{M \times N} \rho'_j \mu_{ij}, \rho'_j = \begin{cases} \rho_j, & \text{если } \rho_j > T, \\ 0 & \text{в остальных случаях,} \end{cases} \quad (33)$$

где  $\rho_j$  – величина движения в местоположении  $j$ -го пикселя,  $T$  – порог, ниже которого векторы оптического потока не учитываются,  $\mu_{ij}$  – значение принадлежности  $\theta_j$  к  $i$ -му интервалу гистограммы.

Шаг 5. Преобразовать точную гистограмму с большим количеством интервалов в грубую гистограмму с меньшим количеством интервалов.

Величина направления движения  $\theta_j$  разбивается на подгруппы по  $n'$  интервалам точной гистограммы в диапазоне  $[-\pi, \pi)$ . Затем интервалы точной гистограммы отображаются в интервалы грубой гистограммы.

$M$  – матрица принадлежности, в которой каждый элемент  $m_{ij}$  содержит значение отображения  $j$ -го точного углового интервала к  $i$ -му грубому угловому интервалу. Выбрав число грубых угловых интервалов равным  $n$ , получим (34):

$$F_{n \times 1} = M_{n \times n'} H_{n' \times 1}. \quad (34)$$

Количества интервалов точной и грубой гистограмм являются настраиваемыми параметрами [122].

Различие между алгоритмами вычисления дескрипторов FHOOF и FHOFO заключается в следующем: дескриптор FHOOF учитывает амплитуду и ориентацию вектора оптического потока, описывающего движение между кадрами, FHOFO – только ориентацию [131].

Схема алгоритмов FHOOF и FHOFO представлена на рисунке 44.

В диссертационной работе было проведено исследование влияния значений настраиваемых параметров алгоритмов дескрипторов FHOOF и FHOFO на диагностическую эффективность обнаружения микролицевых движений на части набора данных SAMM. Используется реализация данных дескрипторов, имеющаяся в открытом доступе и выполненная на языке MATLAB.

Были выполнены следующие изменения в алгоритмах FHOOF и FHOFO:

- изменена функция принадлежности с Гауссовой на колоколообразную;
- изменено количество интервалов точной и грубой гистограмм;
- изменено значение дисперсии.

На вход поступает последовательность изображений  $\{I_k\}$ ,  $k = \overline{1, N}$ .

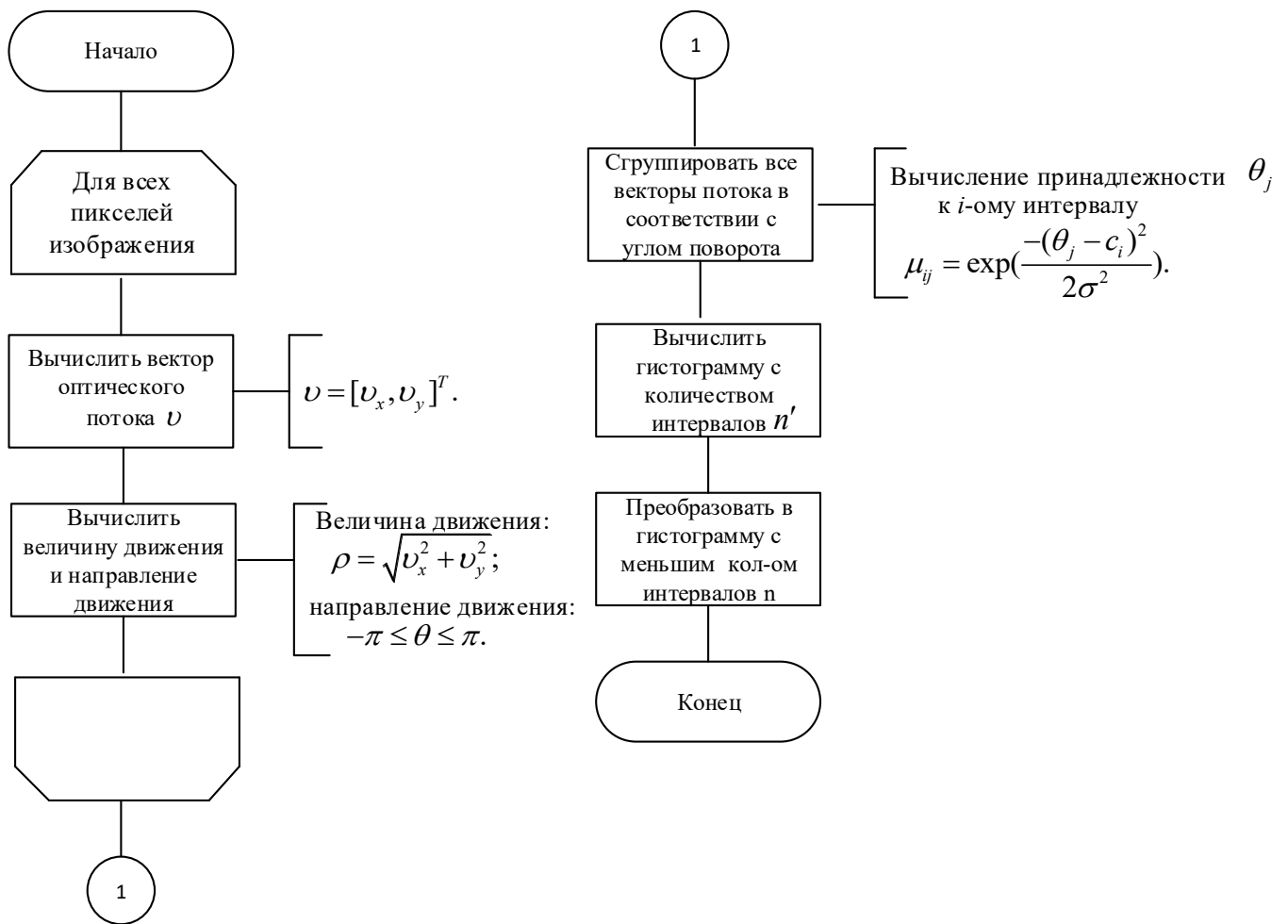


Рисунок 44 – Схема алгоритмов FHOOF и FHOFO

Первый и последний этапы исследования аналогичны этапам эксперимента, описанного в параграфе 3.1. Следующим этапом является извлечение пространственно-временных признаков с помощью алгоритмов вычисления дескрипторов FHOOF или FHOFO. Затем реализуется выравнивание длин вычисленных дескрипторов, так как длины данных дескрипторов признаков зависят от количества кадров в последовательности, отображающей процесс возникновения, пика и затухания микролицевого движения.

На завершающем этапе вектора признаков FHOOF и FHOFO подаются в классификатор SVM с различными функциями ядра или в классификатор MLP. Классификатор MLP настроен на 2000 эпох обучения и 5 скрытых слоев.

По умолчанию для следующих параметров алгоритмов FHOOF и FHOFO были заданы значения:

- количество интервалов нечеткой гистограммы –  $n' = 36$ ;
- количество интервалов точной гистограммы –  $n = 8$ ;
- функция принадлежности – Гауссова;
- дисперсия –  $\sigma = 10$ .

Исследования проводились на части набора данных SAMM. Точность обнаружения микролицевых движений оценивалась по формуле (27).

Вычисленные значения диагностической эффективности обнаружения микролицевых движений на части набора данных SAMM при использовании алгоритмов дескрипторов FHOOF и FHOFO без изменения значений параметров представлены в таблице 13. Наилучшие результаты для каждого из дескрипторов выделены жирным шрифтом.

Таблица 13 – Результаты вычисления диагностической эффективности обнаружения микролицевых движений на части набора данных SAMM при использовании алгоритмов FHOOF и FHOFO с параметрами по умолчанию

Классификатор	Дескриптор признаков	
	FHOOF	FHOFO
SVM: kernel - linear	75%	75%
SVM: kernel - rbf	81%	80%
SVM: kernel - sigmoid	74%	74%
MLP	<b>96%</b>	<b>97%</b>

Из полученных результатов видно, что оба дескриптора показывают близкие значения.

В таблице 14 представлены вычисленные значения диагностической эффективности обнаружения микролицевых движений на части набора данных SAMM при различных значениях настраиваемых параметров алгоритмов дескрипторов FHOOF и FHOFO.



Таблица 14 – Диагностическая эффективность обнаружения микролицевых движений на части набора данных SAMM при различных значениях настраиваемых параметров FHOOF и FHOFO и изменении применяемых классификаторов (а-г)

(а) Результаты FHOOF при алгоритме формирования блоков «Прямоугольные параллелепипеды»

n'	n	σ	μ	Классификатор			
				SVM			MLP
				Linear	RBF	Sigmoid	
46	9	10	gbell	77%	77%	74%	97%
			gauss	75%	<b>81%</b>	74%	98%
60	10	10	gbell	77%	78%	74%	98%
			gauss	75%	80%	74%	95%
70	10	12	gbell	77%	78%	74%	<b>99%</b>
			gauss	75%	80%	74%	97%
80	14	10	gbell	75%	78%	74%	<b>99%</b>
			gauss	75%	78%	74%	96%

(в) Результаты FHOFO при алгоритме формирования блоков «Прямоугольные параллелепипеды»

n'	n	σ	μ	Классификатор			
				SVM			MLP
				Linear	RBF	Sigmoid	
46	9	10	gbell	75%	76%	74%	96%
			gauss	75%	79%	74%	96%
60	10	10	gbell	75%	76%	74%	<b>98%</b>
			gauss	75%	<b>80%</b>	74%	96%
70	10	12	gbell	75%	76%	74%	<b>98%</b>
			gauss	75%	79%	74%	97%
80	14	10	gbell	75%	76%	70%	<b>98%</b>
			gauss	75%	<b>80%</b>	74%	97%

(б) Результаты FHOOF при алгоритме формирования блоков «Скользящие блоки»

n'	n	σ	μ	Классификатор			
				SVM			MLP
				Linear	RBF	Sigmoid	
46	9	10	gbell	79%	81%	70%	95%
			gauss	75%	82%	74%	95%
60	10	10	gbell	77%	82%	71%	95%
			gauss	81%	<b>87%</b>	74%	<b>96%</b>
70	10	12	gbell	76%	82%	71%	95%
			gauss	74%	80%	74%	95%
80	14	10	gbell	80%	80%	72%	95%
			gauss	76%	80%	74%	95%

(г) Результаты FHOFO при алгоритме формирования блоков «Скользящие блоки»

n'	n	σ	μ	Классификатор			
				SVM			MLP
				Linear	RBF	Sigmoid	
46	9	10	gbell	80%	<b>82%</b>	74%	95%
			gauss	74%	81%	74%	93%
60	10	10	gbell	75%	<b>82%</b>	74%	<b>96%</b>
			gauss	75%	79%	74%	95%
70	10	12	gbell	76%	<b>82%</b>	75%	95%
			gauss	75%	79%	74%	94%
80	14	10	gbell	79%	81%	74%	<b>96%</b>
			gauss	75%	80%	74%	<b>96%</b>

Из полученных результатов следует, что диагностическая эффективность обнаружения микролицевых движений при использовании алгоритмов FHOOF и FHOFO и классификатора SVM с ядром RBF выше по сравнению с другими ядрами классификатора. Классификатор MLP показывает наилучшие результаты при любых значениях настраиваемых параметров и любом из рассматриваемых алгоритмов дескрипторов по сравнению с классификатором SVM. Диапазон изменения диагностической эффективности обнаружения микролицевых движений на части набора данных SAMM при использовании алгоритма вычисления дескриптора FHOOF и применения классификатора MLP изменяется от 95% до 99%, классификатора SVM – от 71% до 87%. Диапазон оцениваемой

диагностической эффективности обнаружения микролицевых движений на части набора данных SAMM при использовании алгоритма вычисления дескриптора FHOFO и применения классификатора MLP изменяется от 93% до 98%, классификатора SVM – от 70% до 82%. Наибольшая диагностическая эффективность обнаружения микролицевых движений достигается при использовании классификатора MLP и алгоритма построения блоков «Прямоугольные параллелепипеды».

Для настройки значений параметров алгоритмов с целью повышения диагностической эффективности обнаружения микролицевых движений была разработана методика подбора значений параметров, представленная далее.

1. Выбрать алгоритм формирования блоков, так как наибольшее влияние на диагностическую эффективность обнаружения микролицевых движений при использовании алгоритмов FHOOF и FHOFO оказывает алгоритм формирования блоков. При использовании классификатора многослойного перцептрона следует выбрать алгоритм «Прямоугольные параллелепипеды», при использовании метода опорных векторов с радиальной функцией ядра - «Скользящие блоки».

2. Выбрать функцию принадлежности ( $\mu$ ) гауссова или колоколообразная, по умолчанию - гауссова. При использовании в качестве классификатора метода опорных векторов следует выбрать функцию принадлежности колоколообразную, многослойного перцептрона – гауссову.

3. Установить значение дисперсии ( $\sigma$ ), причем, чем меньше значение дисперсии, тем меньший вклад в соседние интервалы вектора оптического потока, по умолчанию  $\sigma = 10$ , где  $\sigma \in N$ .

4. Установить количество интервалов нечеткой гистограммы ( $n'$ ), по умолчанию  $n' = 36$ , где  $n' \in N$ .

5. Установить количество интервалов уточненной гистограммы ( $n$ ), по умолчанию  $n = 8$ , где  $n \in N$ .

6. Если выполнены шаги 1-5, но увеличения диагностической эффективности обнаружения микролицевых движений при использовании

алгоритмов не обнаружено, следует увеличить значение дисперсии  $\sigma$  на 1. Если требуемое значение диагностической эффективности обнаружения микролицевых движений при использовании алгоритмов не получено выполнить шаги 7 и 8 или повторить шаг 6.

7. Увеличить количество интервалов уточненной гистограммы ( $n$ ). Выполнить классификацию одним из классификаторов, при недостаточном увеличении диагностической эффективности обнаружения микролицевых движений перейти к шагу 8.

8. Увеличить количество интервалов нечеткой гистограммы ( $n'$ ). При необходимости повторить шаг 6.

При значениях параметров: количество интервалов нечеткой гистограммы  $n'=70$ , количество интервалов четкой гистограммы  $n=10$ , функция принадлежности  $\mu = gbell$ , дисперсия  $\sigma = 12$  было измерено время вычисления обоих дескрипторов на 50% набора данных. Эксперименты проводились на ЭВМ с процессором Intel Core i5-3317U с тактовой частотой 1,7ГГц и объемом оперативной памяти 8ГБ. Оптимизация алгоритмов по времени не проводилась.

Результаты эксперимента приведены в таблице 15.

Таблица 15 – Время вычисления дескрипторов FHOOF и FHOFO

Дескриптор признаков	Алгоритм формирования блоков	Время вычисления дескрипторов признаков, минуты
FHOOF	Прямоугольные параллелепипеды	30
	Скользящие блоки	32
FHOFO	Прямоугольные параллелепипеды	12
	Скользящие блоки	16

Из анализа результатов следует, что скорость вычисления дескриптора FHOFO, учитывающего только ориентацию вектора оптического потока, вдвое выше по сравнению со скоростью вычисления дескриптора FHOOF,

учитывающего ориентацию и амплитуду вектора оптического потока, независимо от алгоритма формирования блоков. С другой стороны, диагностическая эффективность обнаружения микролицевых движений при использовании алгоритма FHOOF выше по сравнению с алгоритмом FHOFO [131, 132].

Таким образом, оба алгоритма FHOOF и FHOFO показывают достаточные результаты для применения в конвейере обнаружения микролицевых движений. Параметром, оказывающим наибольшее влияние на диагностическую эффективность обнаружения микролицевых движений, является алгоритм формирования блоков.

### 3.4 Общее сравнение алгоритмов LBP-TOP, FHOOF и FHOFO

Исследование диагностической эффективности обнаружения микролицевых движений и быстродействия алгоритмов проводилось на ЭВМ с процессором Intel Core i5-3317U с тактовой частотой 1,7ГГц и объемом оперативной памяти 8ГБ.

Результаты исследования диагностической эффективности обнаружения микролицевых движений при использовании стандартных алгоритмов LBP-TOP, FHOOF, FHOFO на всем наборе данных SAMM представлены в таблице 16.

Таблица 16 – Результаты оценки диагностической эффективности обнаружения микролицевых движений на наборе данных SAMM при использовании алгоритмов LBP-TOP, FHOOF, FHOFO и стандартных значениях параметров

Дескриптор Классификатор		LBP-TOP		FHOOF		FHOFO	
		Прямоуг. паралл.	Скользящие блоки	Прямоуг. паралл.	Скользящие блоки	Прямоуг. паралл.	Скользящие блоки
SVM	Linear	87%	87%	75%	74%	74%	74%
	RBF	87%	87%	81%	82%	74%	74%
	Sigmoid	87%	87%	74%	74%	63%	63%
MLP		88%	88%	95%	94%	96%	93%
Среднее время вычисления дескрипторов для всего набора, минуты		58		68		40	

Исследование модифицированного алгоритма LBP-TOP на наборе данных SAMM проводилось при следующих значениях параметров:

- расстояние от центрального пикселя до выбранного пикселя  $R=1$ ;
- количество выбранных пикселей  $P=8$ ;
- коэффициент деления блока по оси  $X$   $nQr=4$ ;
- коэффициент деления блока по оси  $Y$   $nQc=4$ ;
- коэффициент деления блока по оси  $T$   $nQt = 3$ ;
- коэффициент перекрытия блоков по горизонтали  $rolr = 0.1$ ;
- коэффициент перекрытия блоков по вертикали  $colr = 0.1$ ;
- коэффициент перекрытия блоков по диагонали  $tolr = 0.1$
- коэффициенты расчета траектории расположения пикселей вокруг центрального пикселя  $coefXT = 0,1$ ,  $coefYT = 0,7$ .

При данных значениях параметров алгоритма LBP-TOP диагностическая эффективность обнаружения микролицевых движений на части набора данных была наибольшей. На всем наборе данных SAMM при модифицированных значениях параметров алгоритма LBP-TOP и использовании классификатора MLP диагностическая эффективность обнаружения микролицевых движений составила 98%. При использовании классификатора SVM – 92% с функцией ядра RBF и алгоритмом формирования трехмерных блоков «Скользящие параллелепипеды», тогда как при стандартной реализации алгоритма LBP-TOP диагностическая эффективность на всем наборе составляла 87%. При этом время вычисления дескрипторов в модифицированной версии увеличилось примерно в 1,5 раза.

Из ранее проведенного исследования алгоритма FHOOF на части набора данных следует, что диагностическая эффективность обнаружения микролицевых движений зависит не только от значений настраиваемых параметров, но и от алгоритма формирования блоков.

При алгоритме формирования блоков «Прямоугольные параллелепипеды» для алгоритма FHOOF были выбраны следующие значения параметров, показавшие наилучшие результаты на части набора данных:

- количество интервалов нечеткой гистограммы  $n'=70$ ;
- количество интервалов четкой гистограммы  $n=10$ ;
- функция принадлежности  $\mu = gbell$ ;
- дисперсия  $\sigma = 12$ .

При данных значениях параметров алгоритма FHOOF и использовании классификатора MLP диагностическая эффективность обнаружения микролицевых движений на наборе данных SAMM составляет 98%, что является более высоким показателем по сравнению с диагностической эффективностью, вычисленной при использовании стандартной реализацией алгоритма, при которой значение диагностической эффективности составило 95%.

При алгоритме формирования блоков «Скользящие параллелепипеды» для алгоритма FHOOF были выбраны следующие значения параметров, показавшие наилучшие результаты на части набора данных:

- количество интервалов нечеткой гистограммы  $n'=60$ ;
- количество интервалов четкой гистограммы  $n=10$ ;
- функция принадлежности  $\mu = gauss$ ;
- дисперсия  $\sigma = 10$ .

Диагностическая эффективность обнаружения микролицевых движений на наборе данных SAMM при использовании классификатора MLP составила 96%, при использовании классификатора SVM - 85%. В зависимости от выбранной функции ядра наибольшее значение достигалось при использовании функции ядра RBF.

Алгоритм FHOFO, также, как и алгоритм FHOOF исследовался при различных значениях параметров для каждой из реализаций алгоритма формирования трехмерных блоков.

Значения параметров, при которых было получено наибольшее значение диагностической эффективности обнаружения микролицевых движений на части набора данных, для алгоритма FHOFO при алгоритме формирования блоков «Прямоугольные параллелепипеды»:

- количество интервалов нечеткой гистограммы  $n'=80$ ;
- количество интервалов четкой гистограммы  $n=14$ ;
- функция принадлежности  $\mu = gauss$ ;
- дисперсия  $\sigma = 10$ .

Наибольшая диагностическая эффективность обнаружения микролицевых движений на наборе данных SAMM была получена при использовании классификатора SVM с ядром RBF – 83%. Причем, при изменении алгоритма формирования блоков на «Скользящие параллелепипеды» диагностическая эффективность не изменилась. При использовании классификатора MLP диагностическая эффективность составила 97%.

Значения параметров, при которых было получено наибольшее значение диагностической эффективности обнаружения микролицевых движений на части набора данных, для алгоритма FHOFO при алгоритме формирования блоков «Скользящие параллелепипеды»:

- количество интервалов нечеткой гистограммы  $n'=70$ ;
- количество интервалов четкой гистограммы  $n=10$ ;
- функция принадлежности  $\mu = gbell$ ;
- дисперсия  $\sigma = 12$ .

Наибольшая диагностическая эффективность обнаружения микролицевых движений на всем наборе SAMM при использовании классификатора SVM с ядром RBF соответствует значению 84%, при использовании классификатора MLP – 97%.

Результаты исследования диагностической эффективности обнаружения микролицевых движений при использовании алгоритмов LBP-TOP, FHOOF, FHOFO с модифицированными значениями параметров на всем наборе данных SAMM представлены в таблице 17.

Алгоритмом с наибольшей длительностью выполнения независимо от наличия модификации является FHOOF. Алгоритмом с наименьшим временем выполнения является FHOFO, но диагностическая эффективность обнаружения

микролицевых движений меньше по сравнению со случаем использования алгоритма LBP-TOP. Таким образом, диагностическая эффективность обнаружения микролицевых движений при использовании алгоритма LBP-TOP наибольшая.

Таблица 17 – Результаты экспериментальных исследований диагностической эффективности обнаружения микролицевых движений на наборе данных SAMM при использовании алгоритмов LBP-TOP, FHOOF, FHOFO с модифицированными значениями параметров

Дескриптор Классификатор		LBP-TOP		FHOOF		FHOFO	
		Прямоуг. паралл.	Скользящие блоки	Прямоуг. паралл.	Скользящие блоки	Прямоуг. паралл.	Скользящие блоки
SVM	Linear	89%	92%	76%	79%	74%	84%
	RBF	90%	92%	81%	85%	83%	84%
	Sigmoid	74%	79%	63%	74%	74%	62%
MLP		96%	98%	98%	96%	97%	97%
Среднее время вычисления дескрипторов для всего набора, минуты		95		114		45	

### 3.5 Выводы по третьей главе

Глава 3 посвящена модификации алгоритмов вычисления пространственно-временных дескрипторов признаков LBP-TOP, FHOOF, FHOFO и анализу влияния настраиваемых параметров алгоритмов на диагностическую эффективность обнаружения микролицевых движений. В результате исследования каждого из этих алгоритмов были найдены комбинации настраиваемых параметров, обеспечивающие наибольшее количество корректно обнаруженных антропометрических точек, соответствующих позициям возникновения микродвижений в последовательности изображений.

В результате модификации структуры алгоритма и изменения значений настраиваемых параметров диагностическая эффективность обнаружения микролицевых движений при использовании алгоритма LBP-TOP увеличилась с



88% до 98%. В результате применения разработанной методики подбора параметров диагностическая эффективность обнаружения микролицевых движений при использовании алгоритмов FHOOF и FHOFO увеличилась до 98% и 97% соответственно, в то время как диагностическая эффективность обнаружения микролицевых движений при стандартной реализации алгоритмов составляет 95% и 96% соответственно. Но с увеличением диагностической эффективности увеличилось и время вычисления дескрипторов признаков с помощью алгоритмов LBP-TOP и FHOOF примерно в 1,5 раза, FHOFO – примерно в 1,1 раза при обработке всего набора данных SAMM.

Все предложенные модифицированные алгоритмы показали возможность их применения в конвейере обнаружения микролицевых движений. Алгоритмом с наименьшим временем выполнения является FHOFO, но его точность обнаружения микролицевых движений меньше по сравнению с алгоритмами LBP-TOP и FHOOF, выбор конкретного типа дескриптора зависит от приоритетов решаемой задачи.

Наилучшие значения диагностической эффективности обнаружения микролицевых движений для всех модифицированных алгоритмов были получены при последующем использовании классификатора MLP. Таким образом, показано, что потенциал методов машинного обучения для решения задачи классификации остается высоким, что делает возможным их дальнейшее использование за счет оптимизации процедуры выбора областей интереса на изображении и согласования дескрипторов признаков с используемым классификатором. При этом объединение методов глубокого обучения для обнаружения антропометрических точек, классических математических алгоритмов для вычисления информативных признаков и классификаторов на основе машинного обучения для классификации позволяет исследователю успешно использовать более простые методы для обеспечения контроля за процедурой обнаружения микролицевых движений, чем последовательное использование нейросетевых методов.

## **4 Разработка программного комплекса для экспериментальных исследований пространственно-временных дескрипторов**

### **4.1 Перспективы диагностики психоэмоционального состояния человека на основе микровыражений лица**

Психическое здоровье можно понимать как состояние оптимального благополучия, при котором люди способны справляться со своими ежедневными стрессорами, удовлетворять свои эмоциональные потребности и вносить ценный вклад в профессиональную и общественную сферы деятельности. На психическое благополучие человека могут влиять разные элементы, включая факторы окружающей среды, социально экономические обстоятельства, личные качества и генетическую предрасположенность. Такие симптомы, как стресс, беспокойство и меланхолия являются ключевыми признаками ухудшения психического здоровья. Растущая осведомленность о психическом здоровье подчеркивает острую необходимость в эффективных и надежных методах выявления психических проблем и своевременного вмешательства. Исследования показывают, что наиболее сильными предикторами суицидальных мыслей и поведения были предыдущие симптомы депрессии. Несмотря на усилия по выявлению и лечению психических заболеваний, их количество увеличивается. Текущие методы оценки для диагностики психических заболеваний в основном полагаются на тяжесть симптомов. Поэтому, существует острая необходимость в объективном и эффективном методе ранней диагностики психических отклонений [133].

Традиционные методы самоотчета и клинические интервью для здоровья трудоемки и субъективны. Методы магнитно-резонансной томографии (МРТ) играют важную роль в диагностике психических заболеваний. Они неинвазивны и имеют относительно хорошее пространственное и временное разрешение для поиска связей между мозгом и психологическими изменениями. Но, данные методы дороги, требуют много времени.

Вместе с тем, существует потенциал для раннего выявления и лучшей диагностики психических проблем, заключающийся в понимании

микровыражений лица для идентификации истинных эмоций человека и являющийся возможной альтернативой для оценки психического здоровья.

Связь между выражением лица человека и психическим здоровьем является очень сложной. Индивидуальные и культурные различия, социальные контексты и опыт могут влиять на интерпретацию эмоций. Способность человека скрывать эмоции делает этот анализ еще более сложным. Например, человек, испытывающий депрессию, может показать видимость счастья, чтобы скрыть внутреннее смятение. Важно использовать объективные методы для расшифровки языка микровыражений лица, чтобы понять состояние психического здоровья. Несмотря на культурные и языковые различия между людьми, большинство скрытых эмоций проявляется через похожие микромимические изменения лица, которые содержат важную информацию о психическом и когнитивном состояниях человека. Таким образом, расшифровывая микровыражения лица можно выйти за рамки субъективных интерпретаций и получить понимание связи между внутренним состоянием и внешними проявлениями личности. Возможность выявлять психологами потенциальные психические расстройства на ранней стадии может помочь в разработке эффективных индивидуальных планов лечения [134].

На сегодняшний день существует возможность разработки компьютерных диагностических систем путем интеграции различных математических методов и методов машинного обучения, способных выполнить диагностику психоэмоционального состояния человека. Данные системы смогут сэкономить финансы и время как пациента, так и психиатра. Более того, может быть обеспечена конфиденциальность, необходимая в данной сфере. Следовательно, указанные системы могут быть рекомендованы как эффективные, безопасные, неинвазивные, работающие в реальном времени и облегчающие раннее выявление психических заболеваний.

В диссертационной работе создан программный комплекс, содержащий модифицированные алгоритмы вычисления дескрипторов признаков LBP-TOP,

FHOOF, FHOFO по антропометрическим точкам, расположенным в областях возможного возникновения микролицевого движения в соответствии со СКЛиД. Для обучения и исследования использовался набор данных SAMM. Диагностическая эффективность обнаружения микролицевых движений при использовании модифицированных в диссертационной работе алгоритмов составляет от 95% до 98%. Проведенное исследование показало, что при наличии сдерживаемых эмоций программный комплекс обнаруживает антропометрические точки, задействованные в микролицевых движениях. Следовательно, данные алгоритмы могут использоваться в системах обнаружения и анализа микролицевых движений человека для оценки психических проблем и, при необходимости, своевременного подбора протокола лечения.

В следующем разделе представлено описание разработанного программного комплекса, обеспечивающего исследование модифицированных алгоритмов LBP-TOP, FHOOF, FHOFO.

#### **4.2 Основные функции разработанного программного комплекса**

Процесс обнаружения микролицевых движений состоит из следующих этапов: вычисление антропометрических точек на лице, вычисление дескрипторов признаков и их классификация. Одним из основных этапов является этап вычисления дескрипторов признаков, для вычисления которых был разработан программный комплекс, позволяющий исследовать пространственно-временные дескрипторы признаков LBP-TOP, FHOOF, FHOFO.

Реализация программного комплекса, состоящего из алгоритмов, обеспечивающих исследование пространственно-временных дескрипторов признаков для дальнейшего обнаружения микролицевых движений, была выполнена в среде MATLAB® с использованием скриптов, разработанных на языке программирования Python.

Программный комплекс обеспечивает выполнение следующих основных функций:

1. Загрузка набора данных.

2. Поиск антропометрических точек на изображениях последовательностей.
3. Вычисление дескрипторов признаков LBP-TOP, FHOOF, FHOFO.
4. Вычисление диагностической эффективности обнаружения микролицевых движений при использовании алгоритмов LBP-TOP, FHOOF, FHOFO на загруженном наборе данных.
5. Вычисление вектора признаков, отображающего, в каких антропометрических точках присутствовало микролицевое движение [129, 135].

### 4.3 Результаты экспериментов с применением разработанного программного комплекса

Применение программного комплекса реализуется в несколько этапов, представленных ниже.

Этап 1. Вычисление антропометрических точек на лице.

На этапе обнаружения и вычисления антропометрических точек на лице использовался фреймворк MediaPipe Face Mesh, который позволяет обнаружить 468 точек, из которых выбираются 62 точки, соответствующие областям возможного возникновения микролицевых движений в соответствии со СКЛид. Схема маски лица с 62 точками для выбранного изображения 011\_03080.jpg представлена на рисунке 45.

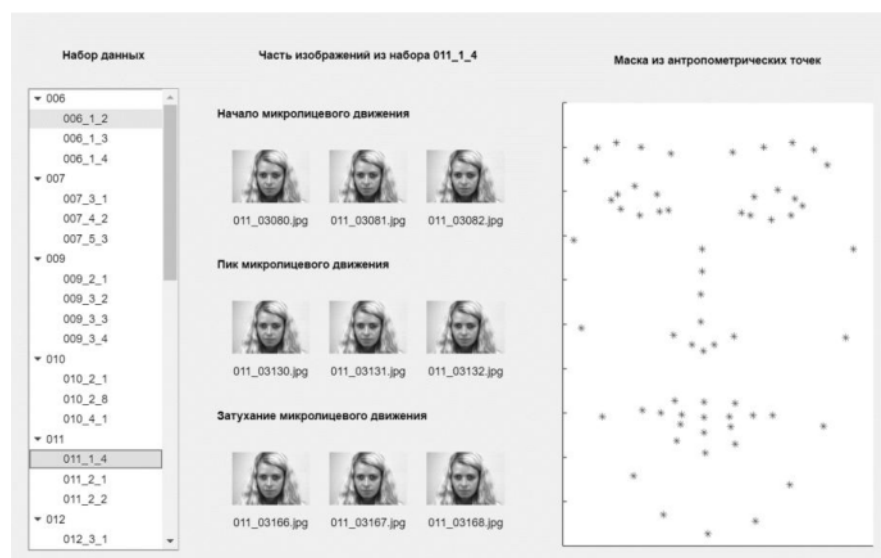


Рисунок 45 – Отображение части изображений из набора данных SAMM и маски по выбранному изображению

Этап 2. Вычисление дескрипторов признаков.

Вычисление дескрипторов признаков выполняется одним из алгоритмов LBP-TOP, FNOOF или FNOFO, представленных в диссертационной работе.

Для вычисления дескрипторов с помощью модифицированного алгоритма LBP-TOP были указаны значения параметров:

- количество выбранных пикселей  $P = 8$ ;
- расстояние от центрального пикселя до выбранного пикселя  $R = 1$ .
- коэффициент деления блока по оси X  $nQr = 4$ ;
- коэффициент деления блока по оси Y  $nQc = 4$ ;
- коэффициент деления блока по оси T  $nQt = 3$ ;
- коэффициент перекрытия блоков по горизонтали  $rolr = 0.1$ ;
- коэффициент перекрытия блоков по вертикали  $colr = 0.1$ ;
- коэффициент перекрытия блоков по диагонали  $tolr = 0.1$ ;
- коэффициент плоскости XT  $coefXT = 0.2$ ;
- коэффициент плоскости YT  $coefYT = 0.4$ .

Алгоритм построения блоков «Скользящие блоки». Блоки изображений делятся на блоки меньшего размера с наложением, траектория размещения точек вокруг центрального пикселя эллиптическая.

Интерфейс программного комплекса для исследования пространственно-временного дескриптора признаков LBP-TOP представлен на рисунке 46.

**2. Вычисление дескрипторов**

Алгоритм построения блоков: Скользящие блоки

LBP-TOP FHOOF/FHOFO

**Введите значения настраиваемых параметров алгоритма вычисления дескриптора LBP-TOP**

Количество выбранных пикселей:  $P =$

Расстояние от центрального пикселя до выбранного пикселя:  $R =$

*Коэффициенты деления блоков изображений на блоки меньшего размера:*

Коэффициент деления блока по оси X:  $nQr =$

Коэффициент деления блока по оси Y:  $nQc =$

Коэффициент деления блока по оси T:  $nQt =$

*Коэффициенты перекрытия блоков между собой:*

Коэффициент перекрытия блоков по горизонтали:  $rolr =$

Коэффициент перекрытия блоков по вертикали:  $colr =$

Коэффициент перекрытия блоков по диагонали:  $tolr =$

*Коэффициенты размещения выбранных пикселей вокруг центрального пикселя:*

Коэффициент плоскости XT:  $coefXT =$

Коэффициент плоскости YT:  $coefYT =$

Вычислить

Сохранить дескрипторы в JSON-файл

Рисунок 46 – Заданные значения параметров алгоритма LBP-TOP

Для вычисления дескрипторов признаков FHOOF были указаны следующие значения параметров (рисунок 47):

- количество интервалов нечеткой гистограммы  $n' = 46$ ;
- количество интервалов четкой гистограммы  $n = 9$ ;
- функция принадлежности – Гауссова;
- дисперсия  $\sigma = 10$ ;
- алгоритм формирования трехмерных блоков вокруг выбранных ключевых точек - «Скользющие блоки».

**2. Вычисление дескрипторов**

Алгоритм построения блоков Скольльзящие блоки

LBP-TOP FHOOF/FHOFO

**Введите значения настраиваемых параметров алгоритма вычисления дескрипторов FHOOF/FHOFO**

Дескриптор

FHOOF  
 FHOFO

Количество интервалов нечеткой гистограммы  $n' =$

Количество интервалов точной гистограммы  $n =$

Дисперсия  $\sigma =$

Функция принадлежности Гауссова

**Примечание** *Дескриптор FHOOF учитывает амплитуду и ориентацию вектора оптического потока, описывающего движение между кадрами.*

*Дескриптор FHOFO учитывает только ориентацию вектора оптического потока, описывающего движение между кадрами.*

Вычислить

Сохранить дескрипторы в JSON-файл

Рисунок 47 – Параметры алгоритма FHOOF

Для вычисления дескрипторов признаков FHOFO были введены следующие значения параметров (рисунок 48):

- количество интервалов нечеткой гистограммы  $n' = 70$ ;
- количество интервалов четкой гистограммы  $n = 10$ ;
- функция принадлежности – Гауссова;
- дисперсия  $\sigma = 12$ ;
- алгоритм формирования трехмерных блоков вокруг выбранных ключевых точек - «Скольльзящие блоки».



**2. Вычисление дескрипторов**

Алгоритм построения блоков Скользящие блоки

**LBP-TOP** **FHOOF/FHOFO**

**Введите значения настраиваемых параметров алгоритма вычисления дескрипторов FHOOF/FHOFO**

Дескриптор  
 FHOOF  
 FHOFO

Количество интервалов нечеткой гистограммы  $n' =$

Количество интервалов точной гистограммы  $n =$

Дисперсия  $\sigma =$

Функция принадлежности Гауссова

**Примечание** *Дескриптор FHOOF учитывает амплитуду и ориентацию вектора оптического потока, описывающего движение между кадрами.*

*Дескриптор FHOFO учитывает только ориентацию вектора оптического потока, описывающего движение между кадрами.*

Рисунок 48 – Параметры алгоритма FHOFO

Этап 3. Вычисление диагностической эффективности обнаружения микролицевых движений.

Для вычисления диагностической эффективности обнаружения микролицевых движений были выбраны классификаторы SVM и MLP. В классификаторе SVM рассматриваются три варианта ядра, а именно:

- линейное (Linear);
- радиальное (RBF);
- сигмоидное (Sigmoid).

Количество эпох обучения классификатора MLP - 2000, количество скрытых слоев – 5 [130].

Диагностическая эффективность обнаружения микролицевых движений при использовании алгоритма LBP-TOP на 1/3 набора SAMM и значениях параметров, перечисленных ранее, представлена на рисунке 49.

3. Вычисление диагностической эффективности	
SVM. Kernel-Linear	76%
SVM. Kernel-RBF	75%
SVM. Kernel-Sigmoid	74%
MLP	93%

Вычислить

Рисунок 49 – Диагностическая эффективность обнаружения микролицевых движений при использовании алгоритма LBP-TOP

Результаты вычисления диагностической эффективности обнаружения микролицевых движений при использовании алгоритма вычисления дескрипторов FHOOF представлены на рисунке 50 [136].

3. Вычисление диагностической эффективности	
SVM. Kernel-Linear	74%
SVM. Kernel-RBF	81%
SVM. Kernel-Sigmoid	74%
MLP	93%

Вычислить

Рисунок 50 – Диагностическая эффективность обнаружения микролицевых движений при использовании алгоритма FHOOF

#### 4.4 Экспериментальные исследования на собственном наборе изображений

Совместно с сотрудниками центра практической психологии РязГМУ был создан собственный набор последовательностей изображений: с нейтральным выражением лица, выражением эмоции радости и эмоции грусти. При создании набора в качестве эмоционального стимула выступала музыка, которую подбирал для себя испытуемый самостоятельно по классам «нейтральная фоновая»,

«вызывающая радость» и «вызывающая грусть», чтобы учесть индивидуальные особенности восприятия.

Участникам эксперимента был продемонстрирован процесс прохождения эксперимента. Наблюдатель оповестил участников, что в любой момент эксперимент можно прервать из-за потенциальной возможности чрезмерного стимулирования эмоций. После каждого стимула участника спрашивали, готов ли он продолжать эксперимент, что гарантировало, компенсацию испытываемых эмоций. Запись участников производилась только тогда, когда воспроизводился стимул. Однако из-за лабораторных условий, в которых находились участники, возможно, отсутствовал должный уровень расслабления, так как присутствовал эффект наблюдения, несмотря на то, что наблюдатель находился вне поля зрения участника, чтобы максимизировать шансы естественного подавления эмоций для создания максимально комфортных условий.

Эксперимент был организован в комнате, которая позволила свести взаимодействие с участником к минимуму и позволила чувствовать себя комфортно в лабораторной обстановке (рисунок 51). В экспериментальной комнате присутствовало все необходимое оборудование для съемки высокоскоростного видео. При настройке среды камера и стул для каждого участника регулировались в зависимости от роста человека, чтобы обеспечить верное захватывание лица.

Для эксперимента использовалась камера GoPro HERO 13 Black Edition со следующими характеристиками:

- число мегапикселей матрицы 27.6 Мп;
- физический размер матрицы 1/1.9”;
- количество объективов: 1;
- максимальное число кадров при максимальном Ultra HD 240 кадр./сек;
- максимальное число кадров при Full HD 240 кадр./сек;
- разрешение 2704x1520.

В связи с отсутствием встроенной памяти видео были получены на карту памяти microSDXC.

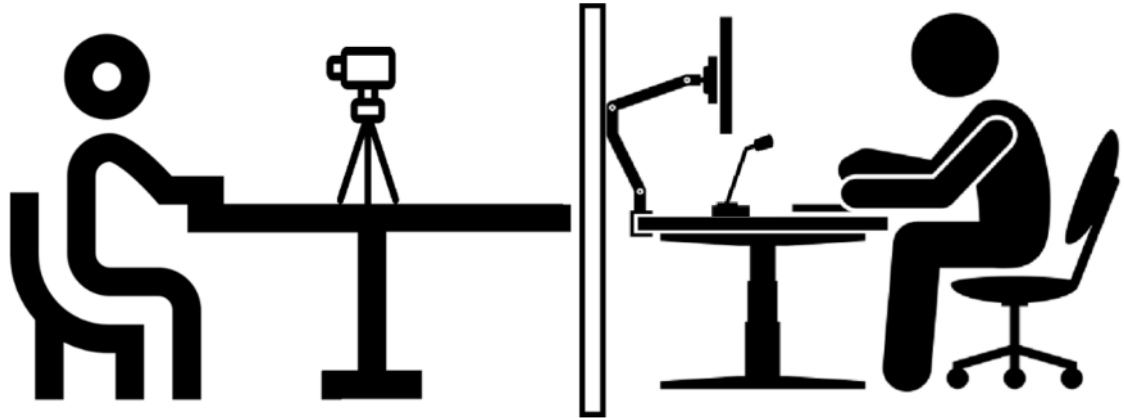


Рисунок 51 – Модель эксперимента: слева представлен участник эксперимента, справа – наблюдатель

Примеры изображений из собственного набора данных представлены на рисунках 52 и 53.



Рисунок 52 - Предварительно выбранная последовательность кадров из видефрагмента с отсутствием микровыражений лица



Рисунок 53 - Предварительно выбранная последовательность кадров из видефрагмента с проявлением микровыражений эмоции радости

Для определения наличия микролицевых движений в каждой из последовательностей изображений собственного набора данных был вычислен с помощью разработанного программного комплекса вектор признаков по разметке лица, представленной на рисунке 36.

Для последовательности изображений с нейтральным лицом, представленной на рисунке 52, вычисленный вектор признаков имеет вид:  
[0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1].

Для последовательности изображений с микровыражением эмоции радости, представленной на рисунке 53, вычисленный вектор признаков имеет вид:  
[0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1].

Диагностическая эффективность обнаружения микролицевых движений при использовании модифицированного и стандартного алгоритмов LBP-TOP на собственном наборе данных представлена в таблице 18.

Таблица 18 – Диагностическая эффективность обнаружения микролицевых движений при использовании стандартного и модифицированного алгоритмов на собственном наборе данных

Дескриптор Классификатор		LBP-TOP	
		Стандартный алгоритм	Модифицированный алгоритм
SVM	Linear	83%	84%
	RBF	62%	68%
	Sigmoid	57%	57%

С помощью программного комплекса для последовательностей изображений с эмоциональной и нейтральной окраской было вычислено количество антропометрических точек, задействованных в микродвижении в каждой из выбранных последовательностей собственного набора. Обучение было

реализовано на части набора данных SAMM. Результаты эксперимента представлены в таблице 19.

Таблица 19 – Доля выявленных микродвижений антропометрических точек на выбранных последовательностях собственного набора изображений

Номер участника эксперимента	Доля антропометрических точек, задействованных в микродвижениях	
	Видеофрагменты, созданные при отсутствии эмоционального стимула	Видеофрагменты, созданные при наличии эмоционального стимула
1	24%	55%
2	16%	34%
3	24%	78%
4	8%	52%
Средняя доля антропометрических точек, задействованных в микродвижениях	18%	55%

Средняя доля выявленных микродвижений антропометрических точек по выбранным последовательностям изображений с эмоциональной окраской из собственного набора составляет 55% от общего количества антропометрических точек, сгруппированных в областях лица, соответствующих микровыражению, отражающему проявление определенной эмоции, в то время, как на последовательностях изображений с подлинно нейтральным выражением лица доля выявленных микродвижений антропометрических точек составляет в среднем 18% от общего количества точек, не соответствующих какому-либо конкретному микровыражению лица. На основе полученных результатов можно сделать вывод, что при наличии сдерживаемых эмоций обнаруживаются

антропометрические точки, задействованные в микролицевых движениях, таким образом, разработанный программный комплекс позволяет выявлять наличие скрытых эмоций.

#### **4.5 Выводы по четвертой главе**

Глава 4 содержит информацию о перспективах диагностики психоэмоционального состояния человека на основе автоматизированного анализа микровыражений лица, разработанном программном комплексе и его основных функциях, а также о результатах применения программного комплекса на собственном наборе экспериментальных данных.

На основе алгоритмических решений, предложенных в главах 2 и 3, разработан и апробирован программный комплекс, предоставляющий возможности обнаружения антропометрических точек, задействованных в микролицевых движениях, и проведения исследований оценки влияния настраиваемых параметров алгоритмов вычисления пространственно-временных дескрипторов признаков на диагностическую эффективность обнаружения микролицевых движений. Программный комплекс позволяет загрузить пользовательский набор последовательностей изображений с микролицевыми движениями, вычислить пространственно-временные дескрипторы по выбранному алгоритму LBP-TOP, FHOOF или FHOFO при заданных пользователем значениях настраиваемых параметров.

Проведены экспериментальные исследования по обнаружению микролицевых движений на собственном наборе данных. Результаты эксперимента подтвердили существенно значимую разницу в количестве микролицевых движений в момент эмоционально окрашенного воздействия по сравнению с моментами отсутствия стимуляции или воздействия эмоционально-нейтральным стимулом.

Таким образом, показано, что совокупность предложенных в диссертации алгоритмических и программных решений позволяет достоверно обнаруживать признаки подавляемых эмоций на видеопоследовательности изображений.

## ЗАКЛЮЧЕНИЕ

В диссертационной работе проведены исследование и модификация алгоритмов построения пространственно-временных дескрипторов признаков, разработаны методики подбора параметров, позволяющие повысить диагностическую эффективность обнаружения микролицевых движений.

Получены следующие результаты диссертационной работы.

1. Разработан новый алгоритм обнаружения антропометрических точек вокруг глаз, обеспечивающий в 2,2 раза меньшее среднеквадратичное отклонение вычисленных координат антропометрических точек от искомым по сравнению с алгоритмом модели активной формы.

2. Модифицирован алгоритм вычисления дескриптора локального бинарного шаблона по трем ортогональным плоскостям путем изменения траектории расположения окрестных точек вокруг центрального пикселя с круговой на эллиптическую, и разработана методика подбора параметров данного алгоритма, которые позволили повысить диагностическую эффективность обнаружения микролицевых движений до 98% при использовании на наборе данных SAMM.

3. Разработана методика подбора параметров для алгоритмов дескриптора нечеткой гистограммы направленного оптического потока и дескриптора нечеткой гистограммы ориентаций оптического потока, которая позволила повысить диагностическую эффективность обнаружения микролицевых движений до 98% и 97% соответственно на наборе данных SAMM.

4. Разработан программный комплекс, предоставляющий возможности обнаружения антропометрических точек, задействованных в микролицевых движениях, и проведения исследований оценки влияния настраиваемых параметров пространственно-временных дескрипторов признаков (LBP-TOР, FHOOF, FHOFO) на диагностическую эффективность обнаружения микролицевых движений.



Совокупность предлагаемых в диссертационной работе алгоритмов, методик и программных средств создает предпосылки для разработки автоматизированных систем для объективной оценки эмоциональных реакций на основе обработки видеопоследовательности изображений лица в задачах диагностики нарушений эмоциональной сферы.

**СПИСОК ЛИТЕРАТУРЫ**

1. Megrhi S., Souidène W., Beghdadi A. Spatio-temporal SURF for Human Action Recognition // *Advances in Multimedia Information Processing*, vol. 8294. – 2003. – Pp. 505-516.
2. Jiang Y.G., Ngo C.W., Yang J. Towards optimal Bag-of-features for object categorization and semantic video retrieval // *Proceedings of the 6th ACM international conference on Image and video retrieval*. 2007. – Pp. 494-501.
3. Hu W., Tan T., Wang L., Maybank S. A Survey on Visual Surveillance of Object Motion and Behaviors // *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 34. – 2004. – Pp. 334-352.
4. Viola P., Jones M. Rapid Object Detection using a Boosted Cascade of Simple Features // *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 12001. – 2001. – Pp. 511-518.
5. Zhao G. and Pietikainen M. Principal appearance and motion from boosted spatiotemporal descriptors // *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. 2008. – Pp. 1-8.
6. Sergeeva A.D. (Chernenko A.D.), Melnik O.V., and Sablina V.A. Facial Micro Expression Analysis for Deception Detection: Methods and Applications, *Proceedings of 3rd ISPRS International Workshop, Photogrammetric and computer vision techniques for video surveillance, biometrics and biomedicine, ISPRS WG 11/10,11/5 Workshop, 2019, Moscow, State Research Institute of Aviation Systems, 13-15 May 2019*. – Pp. 44-45.
7. Саблина В.А., Черненко А.Д. Распознавание выражений лица с помощью дескриптора локальных бинарных шаблонов по трем ортогональным плоскостям // *Биотехнические, медицинские и экологические системы, измерительные устройства и робототехнические комплексы – Биомедсистемы-2021: сб. тр. XXXIV Всерос. науч.-техн. конф. студ., мол. ученых и спец., 8-10 декабря 2021 г. / под общ. ред. В.И. Жулева. – Рязань: ИП Коняхин А.В. (Book Jet), 2021. С. 287-290.*

8. Peteri R. and Chetverikov D. Dynamic Texture Recognition Using Normal Flow and Texture Regularity // *Pattern Recognition and Image Analysis*, vol. 3523. – 2005. – Pp. 223-230.
9. Chetverikov D. and Peteri R. A Brief Survey of Dynamic Texture Description and Recognition // *Computer Recognition Systems*, vol. 30. – 2005. – Pp. 17-26.
10. Doretto G., Chiuso A., Wu Y. and Soatto S. Dynamic Textures // *International Journal of Computer Vision*, vol. 51. - 2003. – Pp. 91-109.
11. Ke S-R., Thuc H.L.U., Lee Y-J., Hwang J-N., Yoo J-H., Choi K-H. A Review on Video-Based Human Activity Recognition // *Computers*. 2013. – Pp. 88-131.
12. Jose C.M., Enrique C.J., Fernández-Caballero A. A survey of video datasets for human action and activity recognition // *Computer Vision and Image Understanding*, vol. 117. – 2013. – Pp. 633-659.
13. Caetano C., Santos J. A., Schwartz W. R. Optical Flow Co-occurrence Matrices: A Novel Spatiotemporal Feature Descriptor // *International Conference on Pattern Recognition*. – 2016. – Pp. 1947-1952.
14. Bastos I., Soares L.R. and Schwartz W.R. Pyramidal Zernike Over Time: A Spatiotemporal Feature Descriptor Based on Zernike Moments // *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. – 2018. – Pp. 77-85.
15. Xu Y., Quan Y., Ling H. and Ji H. Dynamic Texture Classification Using Dynamic Fractal Analysis // *2011 International Conference on Computer Vision*. – 2011. – Pp. 1219-1226.
16. Saisan P., Doretto G., Wu Y. and Soatto S. Dynamic texture recognition // *Proceedings of the IEEE Computer Society Conference and Computer Vision and Pattern Recognition*, vol. 2. – 2001. – Pp. 58–63.
17. Chan A.B. and Vasconcelos N. Probabilistic Kernels for the Classification of Auto-Regressive Visual Processes // *Proceedings of the IEEE Computer Vision and Pattern Recognition*, vol. 1. – 2005. – Pp. 846–851.

18. Ravichandran A., Chaudhry R., and Vidal R. View-invariant dynamic texture recognition using a bag of dynamical systems // Proceedings of the IEEE Computer Vision and Pattern Recognition. – 2009. – Pp. 1651–1657.
19. Li C., Zhong Q., Xie D. and Pu S. Collaborative Spatiotemporal Feature Learning for Video Action Recognition // Proceedings of the IEEE Computer Vision and Pattern Recognition. – 2019. – Pp. 7864-7873.
20. Laptev I. On Space-Time Interest Points // International Journal of Computer Vision, vol. 64. – 2005. – Pp.107–123.
21. Laptev I. and Lindeberg T. Space-time interest points // Proceedings of the Ninth IEEE International Conference on Computer Visio, vol. 1. – 2003. – Pp. 432-439.
22. Harris C. and Stephens M. A combined corner and edge detector // Proceeding of the 4<sup>th</sup> Alvey Vision Conference. – 1988. – Pp. 23.1–23.6.
23. Wang H., Ullah M.M., Klaser A., Laptev I. Evaluation of local spatio-temporal features for action recognition // Proceeding of the British Machine Vision Conference. – 2009. – Pp. 124.1-124.11.
24. Scovanner P., Ali S. and Shah M. A 3-Dimensional SIFT Descriptor and its Application to Action Recognition // Proceeding of the International Conference on Multimedia. – 2007. – Pp. 357-360.
25. Ghamdi M., Zhang L. and Gotoh Y. Spatio-temporal SIFT and Its Application to Human Action Classification // Computer Vision, vol. 7583. – 2012. – Pp. 301-310.
26. Kläser A., Marszalek M. and Schmid C. A Spatio-Temporal Descriptor Based on 3D-Gradients // Proceedings of British Machine Vision Conference, vol. 275. – 2008. – Pp. 1-10.
27. Zhao G. and Pietikainen M. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions // IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29. – 2007. – Pp. 915-928.
28. Ojala T., Pietikainen M. and Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns // IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24. – 2002. – Pp. 971-987.

29. Tan X. and Triggs B. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions // IEEE Transactions on Image Processing, vol. 19. – 2010. – Pp. 1635-1650.
30. Ren J., Jiang X. and Yuan J. Dynamic texture recognition using enhanced LBP features // IEEE International Conference on Acoustics, Speech and Signal Processing. – 2013. – Pp. 2400-2404.
31. Kellokumpu V., Guoying Z. and Pietikäinen M. Human Activity Recognition Using a Dynamic Texture Based Method // Proceedings of the British Machine Vision Conference. – 2008. – 10 p.
32. Chen Y., Guo X., Klein D. Orthogonal combination of local binary patterns for dynamic texture recognition // Proceedings of the Ninth International Symposium on Multispectral Image Processing and Pattern Recognition, vol. 46. – 2015. – Pp. 1949-1963.
33. Lung F.B., Jaward M.H. and Parkkinen J. Spatio-Temporal Descriptor for Abnormal Human Activity Detection // Proceedings of the 2015 14th IAPR International Conference on Machine Vision Applications (MVA). – 2015. – Pp. 471-474.
34. Zhao G., Pietikinen M., Hadid A. Local Spatiotemporal Descriptors for Visual Recognition of Spoken Phrases // Proceedings of the international workshop on Human-centered multimedia. – 2007. – Pp. 57-66.
35. Zhao G., Barnard M. and Pietikainen M. Lipreading With Local Spatiotemporal Descriptors // IEEE Transactions on Multimedia, vol. 11, no. 7. – 2009. – Pp. 1254-1265.
36. Саблина В.А., Яковлев Н.В. Обзор пространственно-временных дескрипторов признаков для анализа выражений лица // Биотехнические, медицинские и экологические системы, измерительные устройства и робототехнические комплексы – Биомедсистемы-2020: сб. тр. XXXIII Всерос. науч.-техн. конф. студ., мол. ученых и спец., 9-11 декабря 2020 г. / под общ. ред. В.И. Жулева. – Рязань: ИП Коняхин А.В. (Book Jet), 2020. С. 292-295.

37. Mota V., Perez E., Maciel L., Vieira M., and Gosselin P. A tensor motion descriptor based on histograms of gradients and optical flow // *Pattern Recognition Letters*, vol. 39. – 2014. – Pp. 85 – 91.
38. Chaudhry R., Ravichandran A., Hager G. and Vidal R. Histograms of Oriented Optical Flow and Binet-Cauchy Kernels on Nonlinear Dynamical Systems for the Recognition of Human Actions // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. – 2009. – Pp. 1932-1939.
39. Li L., Shao L. and Rockett P. Genetic Programming-Evolved Spatio-Temporal Descriptor for Human-Action Recognition // *Proceedings of the British Machine Vision Conference*. – 2012. – Pp. 18.1-18.12.
40. Cen S., Yu Y., Yan G., Yu M. Sparse Spatiotemporal Descriptor for Micro-Expression Recognition Using Enhanced Local Cube Binary Pattern // *Sensors, Sensor Applications on Emotion Recognition*, vol. 20, issue 16. – 2020. – Number 4437.
41. Davison A.K., Hoon Y.M., Costen N., Tan K. Micro-Facial Movements: An Investigation on Spatio-Temporal Descriptors // *Proceedings of the 13<sup>th</sup> European Conference on Computer Vision*, vol. 8926. – 2015. – Pp. 111-123.
42. Frank M.G., Maccario C.J., and Govindaraju V.I. Behavior and security // *Proceedings of the Protecting airline passengers in the age of terrorism*. – 2009. – Pp. 86-106.
43. O’Sullivan M., Frank M.G., Hurley C.M. and Tiwana J. Police lie detection accuracy: The effect of lie scenario // *Law and Human Behavior*, vol. 33. – 2009. – Pp. 542-544.
44. Frank M., Herbasz M., Sinuk K., Keller A., Nolan C. I see how you feel: Training laypeople and professionals to recognize fleeting emotions // *Proceedings of the International Communication Association*. – 2009. – Pp. 1-35.
45. Cohn J.F., Kruez T.S., Matthews I., Yang Y., Nguyen M.H., Padilla M.T. Detecting Depression from Facial Actions and Vocal Prosody // *Proceedings of the Affective Computing and Intelligent Interaction and Workshops*. – 2009. – Pp. 1-79.

46. Morris M.W., Keltner D. How Emotions Work: The Social Functions of Emotional Expression in Negotiations // *Review of Organization Behavior*, vol. 22. – 2000. – Pp. 1–50.
47. Whitehill J., Serpell Z., Yi-Ching L., Foster A. The Faces of Engagement: Automatic Recognition of Student Engagement from Facial Expressions // *IEEE Transactions on Affective Computing*, vol. 5. – 2014. – Pp. 86–98.
48. Ekman P. Lie Catching and Microexpressions // *The Philosophy of Deception*. – 2009. – Pp. 118-136.
49. Ekman P. An Argument for Basic Emotions // *Cognition and Emotion*, vol. 6. – 1992. – Pp. 169–200.
50. Ekman P. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life* // Phoenix. – 2003. – 285 p.
51. Ekman P. and Rosenberg E.L. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System, Second Edition* // *Series in Affective Science*. – 2005. – 662 p.
52. Yan W-J., Wu Q., Liang J., Chen Y-H. How Fast Are the Leaked Facial Expressions: The Duration of Micro-Expressions // *Journal of Nonverbal Behavior*, vol. 37. – 2013. – Pp. 217-230.
53. Ekman P. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. – 2001. – 390 p.
54. Ekman P. and W.V. Friesen *Nonverbal Leakage and Clues to Deception* // *Psychiatric Foundation*, vol. 32. – 1969. – Pp. 88-106.
55. Shen X., Wu Q. and Fu X. Effects of the duration of expressions on the recognition of microexpressions // *Journal of Zhejiang University SCIENCE B*, vol. 13. – 2012. – Pp. 221-230.
56. Сергеева А.Д. (Черненко А.Д.) Анализ микровыражений на основе системы кодирования лицевых движений // *Материалы V научно-технической конференции магистрантов Рязанского государственного радиотехнического университета*, 2019. – С. 206.

57. Happy S.L. and Routray A. Fuzzy Histogram of Optical Flow Orientations for Micro-Expression Recognition // IEEE Transactions on Affective Computing, vol. 10. – 2019. – Pp. 394-406.
58. Pfister T., Li X., Zhao G., Pietikäinen M. Recognising spontaneous facial micro-expressions // Proceedings of the IEEE International Conference on Computer Vision. – 2011. – Pp. 1449-1456.
59. Polikovskiy S., Kameda Y. and Ohta Y. Facial micro-expressions recognition using high speed camera and 3d-gradient descriptor // Proceedings of the 3<sup>rd</sup> International Conference on Imaging for Crime Detection and Prevention. – 2009. – Pp. 16-21.
60. Cootes T., Taylor C. Active Shape Models – Their Training and Application // Computer Vision and Image Understanding, vol. 61. – 1995. – Pp. 38-59.
61. Davison A.K., Lansley C., Ng C.C., Yap M.H. Objective Micro-Facial Movement Detection Using FACS-Based Regions and Baseline Evaluation // Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition. – 2018. – Pp. 642-649.
62. Ojala T., Pietikainen M., Harwood D. A Comparative Study of Texture Measures with Classification Based on Featured Distributions // Pattern Recognition, vol. 29. – 1996. – Pp. 51-59.
63. Fan J., Arandjelovic O. Employing Domain Specific Discriminative Information to Address Inherent Limitations of the LBP Descriptor in Face Recognition // Proceedings of the International Joint Conference on Neural Networks. – 2018. – Pp. 1–7.
64. Karsten J., Arandjelovic O. Automatic Vertebrae Localization from CT Scans using Volumetric Descriptors // Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society. – 2017. – Pp. 576-579.
65. Черненко А.Д. Перспективы развития современных методов и алгоритмов анализа микровыражений лица// Биотехнические, медицинские и экологические системы, измерительные устройства и робототехнические комплексы – Биомедсистемы-2024 [текст]: сб. тр. XXXVII Всерос. науч.-техн. конф. студ., мол.



ученых и спец., 4-6 декабря 2024 г. / под общ. ред. В.И. Жулева. – Рязань: ИП Коняхин А.В. (Book Jet), 2024. – С. 109-112.

66. Xu F., Zhang J. and Wang J.Z. Microexpression Identification and Categorization using a Facial Dynamics Map // *IEEE Transactions on Affective Computing*, vol. 8. – 2017. – Pp. 254-267.

67. Goshtasby A. Image registration by local approximation methods // *Image and Vision Computing*, vol. 6. – 1988. – Pp. 255-261.

68. Wang S-J., Yan W-J., Sun T., Zhao G. Sparse Tensor Canonical Correlation Analysis for Micro-expression Recognition // *Neurocomputing*, vol. 214. – 2016. – Pp. 218-232.

69. Wang Y., See J., Phan R. and Oh Y-H. LBP with Six Intersection Points: Reducing Redundant Information in LBP-TOP for Micro-expression Recognition // *Computer Vision*, vol. 9003. – 2015. – Pp. 525-537.

70. Wang Y., See J., Phan R., Oh Y-H. Efficient Spatio-Temporal Local Binary Patterns for Spontaneous Facial Micro-Expression Recognition // *PLoS One*, vol. 10, issue 5. – 2015. – e0124674.

71. Yu M., Guo Z., Yu Y., Wang Y. Spatiotemporal Feature Descriptor for Micro-Expression Recognition Using Local Cube Binary Pattern // *IEEE Access*, vol. 7. – 2019. – Pp. 159214-159225.

72. Huang X., Wang S-J., Zhao G., Piteikainen M. Facial Micro-Expression Recognition Using Spatiotemporal Local Binary Pattern with Integral Projection // *Proceedings of the IEEE International Conference on Computer Vision Workshops*. – 2015. – Pp. 1-9.

73. Huang X., Wang S-J., Liu X., Zhao G., Feng X. Discriminative Spatiotemporal Local Binary Pattern with Revisited Integral Projection for Spontaneous Facial Micro-Expression Recognition // *IEEE Transactions on Affective Computing*, vol. 10. – 2017. – Pp. 32-47.

74. Huang X. and Zhao G. Spontaneous facial micro-expression analysis using spatiotemporal local radon-based binary pattern // *Proceedings of the International*

Conference on The Frontiers and Advances in Data Science, vol. 81. – 2017. – Pp. 159-164.

75. Ben X., Jia X., Yan R., Zhang X. Learning effective binary descriptors for micro-expression recognition transferred by macroinformation // Pattern Recognition Letters, vol. 107. – 2017. – Pp. 50-58.

76. Liu Y., Zhang J-K., Yan W-J., Wang W-J. A Main Directional Mean Optical Flow Feature for Spontaneous Micro-Expression Recognition // IEEE Transactions on Affective Computing, vol. 7. – 2016. – Pp. 299-310.

77. Asthana A., Zafeiriou S., Cheng S., Pantic M. Robust Discriminative Response Map Fitting with Constrained Local Models // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2013. – Pp. 3444-3451.

78. Liu Y-J., Li B-J., Lai Y-K. Sparse MDMO: Learning a Discriminative Feature for Spontaneous Micro-Expression Recognition // IEEE Transactions on Affective Computing, vol. 12. – 2018. – Pp. 254-261.

79. Liong S., See J., Wong K., Phan R. Less is more: Micro-expression Recognition from Video using Apex Frame // Signal Processing: Image Communication, vol. 62. – 2018. – Pp. 82-92.

80. Allaert B., Bilasco I.M., Djeraba C. Consistent Optical Flow Maps for Full and Micro Facial Expression Recognition // Proceedings of the International Conference on Computer Vision Theory and Applications. – 2018. – Pp. 235-242.

81. Valstar M. and Pantic M. Fully Automatic Recognition of the Temporal Phases of Facial Actions // IEEE Transactions on Systems, Man, and Cybernetics, vol. 42. – 2012. – Pp. 28-43.

82. Chernenko A., Ashapkina M., Sablina V., and Alpatov A. Physical Activity Set Selection for Emotional State Harmonization Based on Facial Micro-Expression Analysis, Proceedings of the 32<sup>nd</sup> International Conference on Computer Graphics and Vision “GraphiCon”. – 2022. – Pp. 682-691.

83. Мельник О.В., Саблина В.А., Черненко А.Д., Никифоров М.Б., Яковлев Н.В. Видеокомпьютерные технологии для объективной оценки скрытых эмоциональных реакций на основе анализа микровыражений лица // Десятая международная конференция по когнитивной науке: Тезисы докладов. Пятигорск, 26-30 июня 2024 г. В двух частях. Часть I / Отв. Ред. Киреев М.В. – Пятигорск, 2024. – С. 206-207.
84. Davison A.K., Yap M.H., Lansley C. Micro-facial Movement Detection Using Individualised Baselines and Histogram-Based Descriptors // Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. – 2015. – Pp. 1864-1869.
85. Liong S-T., See J., Wong K.S., Ngo A., Oh Y-H. Automatic Apex Frame Spotting in Micro-expression Database // Proceedings of the 2015 3<sup>rd</sup> IAPR Asian Conference on Pattern Recognition. – 2015. – Pp. 665-669.
86. Wang S-J., Wu S., Qian X., Li J. A Main Directional Maximal Difference Analysis for Spotting Facial Movements from Long-term Videos // Neurocomputing, vol. 230. – 2016. – Pp. 382-389.
87. Xia Z., Feng X., Peng J., Peng X., Zhao G. Spontaneous Micro-Expression Spotting via Geometric Deformation Modeling // Computer Vision and Image Understanding, vol. 147. – 2015. – Pp. 87-94.
88. Milborrow S. and Nicolls F. Active Shape Models with SIFT Descriptors and MARS // Proceedings of the 2014 International Conference on Computer Vision Theory and Applications, vol. 2. – 2014. – Pp. 380-387.
89. Saragih J.M., Lucey S., Cohn J.F. Face Alignment through Subspace Constrained Mean-Shifts // Proceedings of the 12<sup>th</sup> International Conference on Computer Vision. – 2009. – Pp. 1034-1041.
90. Face++ Research Toolkit [Электронный ресурс]. – URL: [www.faceplusplus.com](http://www.faceplusplus.com) (дата обращения: 26.03.2025).
91. Cristinacce D. and Cootes T.F. Feature Detection and Tracking with Constrained Local Models // Proceedings of the British Machine Vision Conference. – 2006. – Pp. 929-938.

92. Kartynnik Y., Ablavatski A., Grishchenko I., and Grundmann M. Real-time Facial Surface Geometry from Monocular Video on Mobile GPUs // CVPR Workshop on Computer Vision for Augmented and Virtual Reality, Long Beach, CA, USA. – 2019. – 4 p. [Электронный ресурс]. – URL: <https://arxiv.org/abs/1907.06724> (дата обращения: 26.03.2025).
93. Саблина В.А., Сергеева А.Д. (Черненко А.Д.) Методы распознавания микровыражений лица: обзор // Современные технологии в науке и образовании – СТНО-2019: сб. тр. II междунар. науч.-техн. форума: в 10 т. Т.4./ под общ. ред. О.В. Миловзорова. – Рязань: Рязан. гос. радиотехн. ун-т, 2019. – С. 117-122.
94. Savin A.V., Sablina V.A., Nikiforov M.B. Comparison of Facial Landmark Detection Methods for Micro-Expressions Analysis // Proceedings of the 10<sup>th</sup> Mediterranean Conference on Embedded Computing (MECO). Budva, Montenegro. – 2021. – Pp. 336-339.
95. Kazemi V. and Sullivan J. One Millisecond Face Alignment with an Ensemble of Regression Trees // Proceedings of the Computer Vision and Pattern Recognition Conference. – 2014. – 8 p.
96. Zheng H., Geng X., Yang Z. A Relaxed K-SVD Algorithm for Spontaneous Micro-Expression Recognition // Proceeding of the Pacific Rim International Conference on Artificial Intelligence. – 2016. – Pp. 692-699.
97. Yang J., Zhang L., Xu Y., Yang J. Beyond sparsity: The role of L1-optimizer in pattern classification // Pattern Recognition, vol. 45. – 2011. – Pp. 1104-1118.
98. Zheng H. Micro-Expression Recognition based on 2D Gabor Filter and Sparse Representation // Journal of Physics, vol. 787. – 2017. – 6 p.
99. Zong Y., Huang X., Zheng W., Cui Z., Zhao G. Learning From Hierarchical Spatiotemporal Descriptors for Micro-Expression Recognition // IEEE Transactions on Multimedia, vol. 20. – 2018. – Pp. 3160-3172.
100. Никифоров М.Б., Саблина В.А., Черненко А.Д. Применение алгоритмов пространственно-временных дескрипторов признаков для анализа микровыражений лица // 78-я Научно-техническая конференция

Санкт-Петербургского НТО РЭС им. А.С. Попова, посвященная Дню радио. Сборник материалов. – СПб.: СПбГЭТУ «ЛЭТИ», 2023. – С. 365-370.

101. Мельник О.В., Саблина В.А., Черненко А.Д. Применение дескриптора признаков FHOOF для обнаружения микролицевых движений // Биомедицинская радиоэлектроника. 2023. Т. 26. № 3. С. 61-70.

102. Shreve M., Godavarthy S., Goldgof D., Sarkar S. Macro- and micro-expression spotting in long videos using spatio-temporal strain // Proceedings of the 2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops. – 2011. – Pp. 742-748.

103. Warren G., Schertler E., Bull P. Detecting Deception from Emotional and Unemotional Cues // Journal of Nonverbal Behavior, vol. 33. – 2009. – Pp. 59-56.

104. Yan W-J., Wu Q., Liu Y-J., Wang S-J. CASME database: A Dataset of Spontaneous Micro-Expressions Collected from Neutralized Faces // Proceedings of the 10<sup>th</sup> IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. – 2013. – Pp. 1-7.

105. Li X., Pfister T., Huang X., Zhao G. A Spontaneous Micro-expression Database: Inducement, Collection and Baseline // Proceedings of the 10<sup>th</sup> IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. – 2013. – Pp. 1-6.

106. Wen-Jing Y., Xiaobai L., Su-Jing W., Guoying Z. Casme II: An Improved Spontaneous Micro-Expression Database and the Baseline Evaluation // PLoS One, vol. 9, issue 1. – 2014. – e86041.

107. Davison A., Lansley C., Costen N., Tan K., Yap M.H. SAMM: A Spontaneous Micro-Facial Movement Dataset // IEEE Transactions on Affective Computing, vol. 9. – 2018. – Pp. 116-129.

108. Qu F., Wang S-J., Yan W-J., Fu X. CAS(ME)<sup>2</sup>: A Database for Spontaneous Macro-Expression and Micro-Expression Spotting and Recognition // IEEE Transactions on Affective Computing, vol. 9. – 2017. – Pp. 424-436.

109. Fini M., Kashani M. and Rahmati M. Eye Detection and Tracking in Image with Complex Background // Proceedings of the 3<sup>rd</sup> International Conference on Electronics Computer Technology. – 2011. – Pp. 57-61.
110. Merghani W., Davison A.K., Yan M.H. A Review on Facial Micro-Expressions Analysis: Datasets, Features and Metrics // IEEE Transactions, Computer Vision and Pattern Recognition. – 2018. – 19 p. [Электронный ресурс]. – URL: <https://arxiv.org/abs/1805.02397> (дата обращения: 26.03.2025).
111. Rahman N.A. bin A., Wei K.C. and See J. RGB-H-CbCr Skin Colour Model for Human Face Detection // Computer Science. – 2006. – 6 p.
112. Kuznetsova E.G., Shvets E.A., Nikolaev D.P. Viola-Jones Based Hybrid Framework for Real-Time Object Detection in Multispectral Images // Proc. SPIE 9875, Eighth International Conference on Machine Vision, 98750N. – 2015. – 6 p.
113. Lin K., Huang J., Chen J. and Zhou C. Real-time Eye Detection in Video Streams // Proceedings of the Fourth International Conference on Natural Computation. – 2008. – Pp. 193-197.
114. Alekseev A.V., Orlova Y.A., Rozaliev V.L. Automatic Search of Human Face and Distinguish the Contours of the Mouth, Eyes on the Portrait Images, vol. 157. – 2015. – Pp. 49-53.
115. Nasiri J.A., Khanchi S. and Pourreza H.R. Eye Detection Algorithm on Facial Color Images // Computer Science, Second Asia International Conference on Modelling & Simulation (AMS). – 2008. – Pp. 344-349.
116. Azar A.R. and Khalilzadeh F. Real Time Eye Detection Using Edge Detection and Euclidean Distance // Proceedings of the 2<sup>nd</sup> International Conference on Knowledge-Based Engineering and Innovation (KBEI). – 2015. – Pp. 43-48.
117. Chai D. and Ngan K.N. Face Segmentation Using Skin-Color Map in Videophone Applications // IEEE Transactions on Circuits and Systems for Video Technology, vol. 9. – 1999. – Pp. 551-564.
118. Sergeeva A.D. (Chernenko A.D.), Savin A.V., Sablina V.A. and Melnik O.V. Emotion Recognition from Micro-Expressions: Search for the Face and Eyes //

Proceedings of the 8<sup>th</sup> Mediterranean Conference on Embedded Computing (MECO). Budva, Montenegro. – 2019. – Pp. 632-635.

119. Саблина В.А., Сергеева А.Д. (Черненко А.Д.) Классификация методов обнаружения антропометрических точек лица // Биотехнические, медицинские и экологические системы, измерительные устройства и робототехнические комплексы – Биомедсистемы-2019: сб. тр. XXXII Всерос. науч.-техн. конф. студ., мол. ученых и спец., 4-6 декабря 2019 г. / под общ. ред. В.И. Жулева. – Рязань: ИП Коняхин А.В. (Book Jet), 2019. – С. 203-206.

120. Саблина В.А., Сергеева А.Д. (Черненко А.Д.) Использование метода ASM для обнаружения антропометрических точек лица // Материалы IV Всероссийской научно-технической конференции «Актуальные проблемы современной науки и производства». – Рязань: РГРТУ, 2019. – С. 122-128.

121. Wang Q., Xie L., Zhu B., Yang T. and Y. Zheng Facial Features Extraction based on Active Shape Model // Journal of Multimedia, vol. 8. – 2013. – Pp. 747-754.

122. D’Mello S. Monitoring Affective Trajectories During Complex Learning // Encyclopedia of the Sciences of Learning. – 2007. – Pp. 2325-2328.

123. Саблина В.А., Черненко А.Д. Программа обнаружения антропометрических точек вокруг глаз на изображении лица. Номер государственной регистрации программы для ЭВМ (свидетельства) в ФИПС: 2024665193 // Официальный бюллетень Федеральной службы по интеллектуальной собственности (Роспатент), № 7, 27.06.2024. – 1 с.

124. Sergeeva A.D. (Chernenko A.D.) and Sablina V.A. Eye Landmarks Detection Technology for Facial Micro-Expressions Analysis // Proceedings of the 9<sup>th</sup> Mediterranean Conference on Embedded Computing (MECO). Budva, Montenegro. – 2020. – Pp. 448-451.

125. Sun Y., Wang X. and Tang X. Deep Convolutional Network Cascade for Facial Point Detection // IEEE Conference on Computer Vision and Pattern Recognition. – 2013. – Pp. 3476-3483.

126. Саблина В.А., Черненко А.Д. Подготовка набора векторов признаков для обнаружения микролицевых движений // Биотехнические, медицинские и экологические системы, измерительные устройства и робототехнические комплексы – Биомедсистемы-2022 [текст]: сб. тр. XXXV Всерос. науч.-техн. конф. студ., мол. ученых и спец., 7-9 декабря 2022 г. / под общ. ред. В.И. Жулева. – Рязань: ИП Коняхин А.В. (Book Jet), 2021. – С. 190-192.
127. Sablina V., Burrese G. Micro-Facial Movement Detection Using LBP-TOP Descriptors for Landmark Based Regions // Proceedings of the 10<sup>th</sup> Mediterranean Conference on Embedded Computing. Budva, Montenegro. – 2021. – Pp. 401-404.
128. Мельник О.В., Саблина В.А., Черненко А.Д. Программный комплекс для анализа пространственно-временных дескрипторов признаков микродвижений лица на видеопоследовательности // Труды XVI Международной научной конференции «Физика и радиоэлектроника в медицине и экологии – ФРЭМЭ'2024» – Владимир-Суздаль, Россия, Доклады, 2024. – С. 239-243.
129. Черненко А.Д. Алгоритм вычисления дескриптора локальных бинарных шаблонов по трем ортогональным плоскостям для распознавания микровыражений лица // Современные технологии в науке и образовании – СТНО-2022 [текст]: сб. тр. V междунар. науч.-техн. форума: в 10 т. Т.4./ под общ. ред. О.В. Миловзорова. – Рязань: Рязан. гос. радиотехн. ун-т, 2022. – С. 89-94.
130. Мельник О.В., Саблина В.А., Черненко А.Д. Распознавание микровыражений лица с использованием классификаторов на основе машинного обучения // Модели, системы, сети в экономике, технике, природе и обществе. 2023. №1. С 125-135.
131. Melnik O.V., Nikiforov M.B, Sablina V.A., Chernenko A.D. A Comparative Analysis of the Micro-Facial Movement Detection Accuracy Depending on the Fuzzy Spatio-Temporal Feature Descriptor Selection, the Proceedings of the 2023 Conference on Systems and Technologies of the Digital HealthCare (STDH). Tashkent, Uzbekistan. – 2023. – Pp. 77-80.



132. Мельник О.В., Никифоров М.Б., Саблина В.А., Черненко А.Д. Обнаружение микролицевых движений с помощью пространственно-временных дескрипторов на основе оптического потока // Биомедицинская радиоэлектроника. 2023. Т. 26. №2. С. 5-17.
133. Hu B., Tao Y., Yang M. Detecting depression based on facial cues elicited by emotional stimuli in video // Computer in Biology and Medicine, vol. 165. – 2023. – Number 107457.
134. Sharma D., Singh J., Sehra S.S., Sehra S.K. Demystifying Mental Health by Decoding Facial Action Unit Sequences // Big Data and Cognitive Computing, vol. 8, issue 7. – 2024. – Number 78.
135. Саблина В.А., Черненко А.Д. Программный комплекс вычисления пространственно-временных дескрипторов признаков на основе антропометрических точек для анализа микровыражений лица. Номер государственной регистрации программы для ЭВМ (свидетельства) в ФИПС: 2024665164 // Официальный бюллетень Федеральной службы по интеллектуальной собственности (Роспатент), № 7, 27.06.2024. – 1 с.
136. Мельник О.В., Саблина В.А., Черненко А.Д. Программный комплекс для исследования пространственно-временных дескрипторов признаков // Биомедицинская радиоэлектроника. 2024. Т 27. №.4 С. 48-55.

## ПРИЛОЖЕНИЕ А

Свидетельства о государственной регистрации программ для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



## СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2024665164

**Программный комплекс вычисления пространственно-временных дескрипторов признаков на основе антропометрических точек для анализа микровыражений лица**

Правообладатель: *федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина» (RU)*

Авторы: *Саблина Виктория Александровна (RU), Черненко Анна Дмитриевна (RU)*

Заявка № 2024663201

Дата поступления 06 июня 2024 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 27 июня 2024 г.



Руководитель Федеральной службы  
по интеллектуальной собственности

Ю.С. Zubov

РОССИЙСКАЯ ФЕДЕРАЦИЯ



## СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2024665193

**Программа обнаружения антропометрических точек  
вокруг глаз на изображении лица**

Правообладатель: *федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Рязанский государственный радиотехнический  
университет имени В.Ф. Уткина» (RU)*

Авторы: *Саблина Виктория Александровна (RU), Черненко  
Анна Дмитриевна (RU)*

Заявка № 2024663202

Дата поступления 06 июня 2024 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 27 июня 2024 г.



Руководитель Федеральной службы  
по интеллектуальной собственности

Ю.С. Зубов

## ПРИЛОЖЕНИЕ Б

Документы о практическом использовании результатов  
диссертационного исследования

## АКТ

об использовании результатов диссертационной работы Черненко Анны Дмитриевны на соискание ученой степени кандидата технических наук на тему «Алгоритмы построения пространственно-временных дескрипторов признаков для обнаружения микролицевых движений» в центре практической психологии ФГБОУ ВО «Рязанский государственный медицинский университет имени академика И.П. Павлова» Министерства здравоохранения Российской Федерации

Анализ мимики является одним из инструментов, помогающим психологу оценить текущее эмоциональное состояние человека. При этом микродвижения лицевых мышц, возникающие в случае сознательного или неосознанного подавления эмоциональной реакции, могут быть отслежены только в некоторых случаях специально обученным специалистом-профайлером. Актуальной задачей является создание программно-алгоритмических средств, выявляющих признаки подавляемых эмоций на видеоизображении, что может быть полезно как при диагностике, так и в терапии различных расстройств эмоциональной сферы.

В диссертационной работе Черненко А.Д. разработаны алгоритмы и программный комплекс с понятным графическим интерфейсом для обнаружения микролицевых движений на видеозаписи. Программный комплекс позволяет исследовать влияние выбора дескриптора (LBP-TOP, FHOOF или FHOFO) и его параметров на точность обнаружения микролицевых движений для различных наборов входных данных видеоизображений лица человека и подобрать оптимальное сочетание настроек для проведения экспериментальных исследований.

Исследование микролицевых движений в ответ на воздействие эмоционально окрашенных стимулов открывает новые перспективы в объективной оценке эмоциональных реакций. Предлагаемые Черненко А.Д. в диссертационной работе программно-алгоритмические средства использовались при проведении экспериментальных исследований по оценке воздействия музыкальных произведений на эмоциональную сферу человека. От испытуемых требовалось сохранять максимально нейтральное выражение лица вне зависимости от того, какое воздействие на них оказывает музыка. Результаты эксперимента подтвердили существенно значимую разницу в количестве микролицевых движений в момент эмоционально окрашенного воздействия по сравнению с моментами отсутствия стимуляции или воздействия эмоционально-нейтральным стимулом, что позволяет использовать данный подход для выявления признаков подавляемых эмоций на основе автоматизированного анализа видеозаписи.

Начальник центра практической психологии  
ФГБОУ ВО РязГМУ Минздрава России  
кандидат психологических наук

*В.В. Яковлев* В.В. Яковлев



11.03.2020г

ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ «БИОТЕХПРОДАКТС»

(ООО «БИОТЕХПРОДАКТС»)

390026, г. Рязань, Татарская улица, дом 51/15, кв. 107  
ИНН/КПП 6234162093/623401001 ОГРН 1166234073060

тел.: +7-910-611-25-57  
e-mail: mary62@mail.ru

## АКТ

### о внедрении результатов диссертационной работы

Настоящим актом подтверждается практическое использование обществом с ограниченной ответственностью «БИОТЕХПРОДАКТС» следующих результатов, полученных Черненко Анной Дмитриевной в рамках подготовки ее кандидатской диссертации.

1. Методики подбора параметров алгоритмов вычисления дескрипторов признаков, повышающие диагностическую эффективность дальнейшего анализа микродвижений лица.
2. Программное обеспечение для исследования алгоритмов вычисления пространственно-временных дескрипторов признаков.

Диссертационное исследование Черненко А.Д. посвящено развитию алгоритмов вычисления пространственно-временных дескрипторов признаков для анализа микродвижений лица. Предложенные научно-технические решения пригодны для практического применения в медицине (диагностика психических заболеваний), образовании (диагностика вовлеченности в учебный процесс и заинтересованности учебным материалом), а также в деловых переговорах, международной торговле и межличностном взаимодействии. Организацию «БИОТЕХПРОДАКТС» в первую очередь интересует возможность применения полученных результатов в медицинской сфере. Поскольку эмоции играют важную роль в жизни человека и влияют на весь организм в целом, анализ микродвижений лица полезен для диагностики эмоционального состояния пациента и дальнейшей его реабилитации.

Для оценки наличия микродвижений лица обработка входных изображений реализуется гибридными методами, а именно обнаружение и выбор антропометрических точек реализуется методом глубокого обучения, извлечение пространственно-временных дескрипторов признаков – улучшенными классическими математическими алгоритмами, вычисляющими дескрипторы признаков только в областях возможного возникновения микродвижений лица по системе кодирования лицевых движений, классификация дескрипторов признаков – методами машинного обучения.

Таким образом, технические и методологические аспекты диссертационного исследования Черненко А.Д. позволяют на практике добиться более высокой диагностической эффективности алгоритмов вычисления пространственно-временных дескрипторов признаков для дальнейшего получения более высокой достоверности определения эмоционального состояния пациента и более точного подбора комплекса медицинских мероприятий для нормализации его эмоционального состояния.

Директор ООО «БИОТЕХПРОДАКТС»,  
кандидат технических наук



М.С. Ашاپкина

20 25 г.

«УТВЕРЖДАЮ»

Проректор по  
образовательной деятельности  
ФГБОУ ВО «Рязанский  
государственный радиотехнический  
университет им. В.Ф. Уткина»,  
к.п.н., доцент



Е.А. Соколова

2025 г.

АКТ

о внедрении результатов диссертационной работы  
Черненко Анны Дмитриевны на тему «Алгоритмы построения пространственно-  
временных дескрипторов признаков для обнаружения микролицевых движений»  
в учебный процесс ФГБОУ ВО «Рязанский  
государственный радиотехнический университет им. В.Ф. Уткина»

Настоящий акт составлен в том, что результаты диссертационной работы Черненко А.Д. на тему «Алгоритмы построения пространственно-временных дескрипторов признаков для обнаружения микролицевых движений» используются в учебном процессе федерального государственного бюджетного образовательного учреждения высшего образования «Рязанский государственный радиотехнический университет им. В.Ф. Уткина» у студентов, обучающихся по направлению подготовки бакалавров 12.03.04 «Биотехнические системы и технологии» на кафедре информационно-измерительной и биомедицинской техники по дисциплине «Биотехнические системы медицинского назначения» при проведении практических и лабораторных занятий.

Основные результаты и выводы исследования Черненко А.Д. изложены в мультимедийном электронном учебном пособии «Методы оценки психоэмоционального состояния человека на основе анализа видеоизображений» (номер регистрации экземпляра 7741, дата регистрации 10.09.2024) для студентов высших учебных заведений. Учебное пособие доступно по электронному адресу: <https://elib.rsreu.ru/ebs/download/4019>.

Декан факультета автоматки и  
информационных технологий  
в управлении,  
к.т.н., доцент

М.В. Ленков

Заведующий кафедрой  
информационно-измерительной  
и биомедицинской техники,  
д.т.н., профессор

В.И. Жулев

Доцент кафедры  
информационно-измерительной  
и биомедицинской техники,  
к.т.н., доцент

А.М. Абрамов