

На правах рукописи



**ДВОЙНОЙ Илья Ростиславович**

**МЕТОДЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЯ ЛИЦА ЧЕЛОВЕКА ПО  
ЦВЕТОВЫМ ПРИЗНАКАМ И ИДЕНТИФИКАЦИИ ЛИЧНОСТИ  
НА ОСНОВЕ СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ  
В СИСТЕМАХ ВИДЕОНАБЛЮДЕНИЯ**

Специальность 05.13.17 – теоретические основы информатики

**Автореферат**  
диссертации на соискание ученой степени  
кандидата технических наук

Пенза – 2013

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего профессионального образования «Пензенский государственный технологический университет» на кафедре «Вычислительные машины и системы».

Научный руководитель – доктор технических наук, профессор  
**Сальников Игорь Иванович.**

Официальные оппоненты: **Рыжаков Виктор Васильевич,**  
доктор технических наук, профессор,  
Заслуженный деятель науки РФ,  
ФГБОУ ВПО «Пензенский государственный  
технологический университет» заведующий  
кафедрой «Техническое управление качеством»;  
**Масленников Алексей Анатольевич,**  
кандидат технических наук, доцент,  
ФГБОУ ВПО «Пензенский государственный  
университет», доцент кафедры  
«Информационно-вычислительные системы».

Ведущая организация – Научно-исследовательский и конструкторский  
институт радиоэлектронной техники – филиал  
федерального государственного унитарного  
предприятия федерального научно-  
производственного центра «Производственное  
объединение «Старт» имени М.В. Проценко»,  
г. Заречный Пензенской обл.

Защита состоится 26 декабря 2013 года в 16 часов на заседании диссертационного совета Д 212.337.01 на базе ФГБОУ ВПО «Пензенский государственный технологический университет» по адресу: 440039, г. Пенза, пр. Байдукова / ул. Гагарина, д. 1а / 11, корпус 1, конференц-зал.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВПО «Пензенский государственный технологический университет».

Автореферат разослан 25 ноября 2013 года.

Ученый секретарь  
диссертационного совета



**Чулков Валерий Александрович**

## ВВЕДЕНИЕ

**Актуальность темы.** Возросший за последнее время уровень социальной напряженности в связи с непрекращающейся угрозой террористических атак усиливает роль технических средств видеоконтроля, устанавливаемых в местах повышенной концентрации людей (стадионы, станции метро, вокзалы и т.п.). В настоящее время системы видеонаблюдения применяются преимущественно для фиксации происходящих событий с целью обеспечения проведения возможных следственных действий. Актуальным является использование систем видеонаблюдения также для осуществления превентивной функции, заключающейся в своевременном автоматическом выявлении находящихся в розыске лиц, ранее причастных к организации беспорядков, порче общественного имущества и т.п. Особую роль для реализации подобных систем играют способы распознавания личности по изображению лица человека.

Существующие подходы к решению задачи установления личности человека по изображению лица во многом были сформированы в 80-х годах прошлого века. Несмотря на большое количество ежегодно выходящих посвященных данной теме публикаций, решение указанной задачи по-прежнему требует совершенствования аппаратных и программных средств с целью улучшения временных и вероятностных характеристик процесса идентификации.

Наиболее полный обзор существующих алгоритмов анализа и распознавания изображений в системах технического зрения приводится в работах таких ученых, как Дуда Р., Харт П., Шапиро Л., Стокман Д., Форсайт Д., Понс Ж., Прэтт У., Розенфельд А., Федотов Н.Г., Сальников И.И. и др.

Отличительной особенностью сложившегося подхода к распознаванию изображений является использование единого признака как для сегментации изображения на составляющие объекты, так и для последующей классификации объектов, следствием чего, как правило, оказываются высокая вычислительная сложность алгоритмов распознавания и наличие жестких ограничений на условия формирования изображений. По этой причине для решения практических задач требуется создание новых методов, основанных на анализе целого ряда дополняющих друг друга признаков, таких как цвет, форма, взаимное расположение выделенных объектов и их составных частей и т.п.

Кроме того, следует принимать во внимание непрерывное совершенствование методов анализа изображений, повышение быстродействия используемых для этой цели технических средств, а также возможности увеличения числа процессоров вычислительной системы и распараллеливания вычислительных процессов.

Современные видеокамеры наружного наблюдения позволяют получить цветное изображение размером 1920×1080 пикселей и более с частотой порядка 25–30 кадров в секунду. Столь высокие характеристики дают основание пересмотреть ранее сформировавшиеся подходы к задаче распознавания лиц в системах, основанных на применении видеокамер наблюдения.

Цвет является мощным средством распознавания и различения объектов на изображении. Несмотря на это, в существующих методах и алгоритмах рас-

познавания лица человека ему, как правило, отводится второстепенная роль, либо его значение вовсе не учитывается. Причиной этого является то, что до недавнего времени видеокамеры наблюдения, работающие в цветном режиме, были мало распространены из-за невысокой светочувствительности и низкой контрастности получаемого изображения. Современный уровень развития техники снимает подобные ограничения.

Как показывают эксперименты, цвет кожи разных людей локализуется в рамках небольшой ограниченной области цветового пространства. Даже при рассмотрении цветов кожи людей различной расовой принадлежности не наблюдается существенного увеличения цветового разброса. Исследования в данной области свидетельствуют о том, что основное отличие заключается не в оттенке цвета, а в его яркости, что позволяет использовать характерный цвет кожи как признак для распознавания лиц.

Метод распознавания лиц, основанный на анализе цветовых признаков, позволяет всего за один проход по изображению выделить области, содержащие лица людей. Однако он не может ответить на вопрос, кому эти лица принадлежат. Для решения задачи распознавания необходимо использовать методы, в большей степени учитывающие персонифицированные признаки, нежели общеклассовые. Для этих целей хорошо подходят алгоритмы, основанные на применении скрытых марковских моделей (СММ).

Как устойчивый распознаватель с большим количеством настраиваемых параметров метод на основе СММ позволяет в полной мере учесть особенности изображения лица, уменьшить влияние таких факторов, как освещение, искажения перспективы, искажения, связанные с использованием широкоформатных объективов и пр. Помимо этого существует возможность учесть двумерную структуру изображения, тем самым исключая возможность ошибочной классификации области глаз как области рта и т.п.

На сегодняшний день теория СММ составляет основу индустриального стандарта для приложений, осуществляющих распознавание речи. СММ успешно применяются для решения задач распознавания рукописного текста и биоинформатики. Возможность применения СММ в задачах идентификации личности человека проработана в значительно меньшей степени. Отчасти это объясняется наличием уже устоявшихся, проверенных временем алгоритмов, отчасти – сложностью выбора оптимальных с точки зрения идентификации личности параметров СММ, требующего проведения теоретических изысканий, разработок и экспериментальных исследований.

В этой связи, тема диссертационного исследования, посвященная распознаванию изображения лица человека по цветовым признакам с целью идентификации его личности для применения в системах видеонаблюдения, является **актуальной**.

**Цель** настоящей работы состоит в совершенствовании методов автоматизированной идентификации личности при анализе изображений в системах видеонаблюдения и повышении на этой основе уровня безопасности людей в местах их массового скопления.

Для достижения поставленной цели сформулированы следующие **задачи**:

1) разработка метода и алгоритма кодирования цветного изображения для сегментации по цветовым признакам с целью выделения областей, содержащих информацию о лице человека;

2) обоснование и разработка математической модели описания отличительных особенностей изображения лица на основе скрытой марковской модели, а также выбор признаков, формирующих алфавит наблюдаемой последовательности признаков;

3) разработка метода и алгоритма выделения ключевых признаков, характеризующих представленное на изображении лицо;

4) разработка алгоритма поиска в базе данных соответствий выделенных признаков анализируемого изображения эталонным признакам.

**Объектом исследования** в диссертационной работе являются системы автоматического установления личности по изображению лица, поступающему с видеокамер наблюдения.

**Предметом исследования** являются методы и алгоритмы для систем технического зрения, позволяющие выделять характерные признаки изображения лица человека, сохранять их в базе данных, а также производить поиск соответствий признаков анализируемого изображения эталонам.

**Методы и средства исследования.** При решении поставленных задач использовались положения теории распознавания образов, теория систем технического зрения, методы математического анализа, теория цветового кодирования изображений, методы цифровой обработки сигналов и интегрированные среды разработки программ.

**Научная новизна** работы. Новыми являются следующие научные результаты.

1. Разработан новый метод сегментации изображения, основанный на применении специализированного цветового пространства, реализующий алгоритмы выделения областей по цветовым признакам объектов, позволяющие выявить элементы изображения кожи человека с высокой степенью точности.

2. Математическая модель описания изображения лица человека с выявлением отличительных признаков на основе скрытой марковской модели с псевдодвумерной структурой, позволяющая распознать человека по изображению лица.

3. Метод выделения ключевых признаков, характеризующих изображение лица, на основе алгоритмов динамического программирования.

4. Алгоритм поиска соответствий выделенных признаков изображения эталонным признакам в базе данных, основанный на определении вероятности генерирования наблюдаемой кодовой последовательности каждой из имеющихся скрытых марковских моделей с помощью алгоритма прямого хода.

**Практическая значимость работы** заключается в том, что применение разработанных методов и алгоритмов позволяет увеличить вероятность правильного распознавания изображения лица человека по цветовым признакам и идентификации личности человека по его изображению в системах видеонаб-

людения и на этой основе повысить уровень безопасности людей в местах их массового скопления.

**Реализация и внедрение результатов работы.** Разработанные методы и алгоритмы выделения лица на изображении человека в системах видеонаблюдения использованы в разработках центра по охране объектов органов государственной власти и правительственных учреждений г. Москвы ФГКУ УВО ГУ МВД России по г. Москве, а также применяются в учебном процессе при обучении студентов по направлению подготовки 230100 – «Информатика и вычислительная техника» в Пензенском государственном технологическом университете в рамках дисциплин «Системы искусственного интеллекта» и «Системы технического зрения».

**Достоверность и обоснованность** научных положений и выводов подтверждается корректностью допущений и ограничений при выводе математических выражений, согласованностью основных теоретических положений диссертации с практической реализацией предложенных решений, а также с результатами вычислительных экспериментов.

**Положения, выносимые на защиту:**

1) метод и алгоритм сегментации изображения, основанный на применении специализированного цветового пространства, позволяющий выделить область лица на изображении человека по цветовым признакам;

2) математическая модель процесса выделения отличительных признаков изображения лица, основанная на теории скрытых марковских моделей, позволяющая распознать изображение лица человека;

3) метод выделения ключевых признаков, характеризующих изображение лица человека;

4) алгоритм поиска соответствия признаков лиц эталонным признакам в базе данных, основанный на определении вероятности генерации наблюдаемой последовательности каждой из имеющихся скрытых марковских моделей.

**Апробация работы.** Основные положения и результаты диссертационного исследования докладывались и обсуждались на ежегодной Всероссийской научно-технической конференции «Современные методы и средства обработки пространственно-временных сигналов» (г. Пенза, 2010, 2012, 2013 гг.), на III научной конференции аспирантов и молодых ученых (г. Пенза, 2010 г.), на IV Международной электронной научной конференции студентов и аспирантов «Студенческий научный форум» (2012 г.).

**Публикации.** По результатам исследований опубликовано 9 печатных работ, из них 3 – в изданиях, рекомендованных ВАК.

**Личный вклад автора.** Основные изложенные в диссертации новые результаты получены лично автором. В работах, опубликованных в соавторстве, автору принадлежат: в [3] – вывод уравнений и проведение экспериментальной части исследования; в [4] – разработка алгоритмов; в [5] – выбор группы цветовых пространств и проведение их сравнительного анализа; в [7] – вывод формул для представления кодов цвета в полярной системе координат и разработка алгоритма сегментации.

**Структура и объем работы.** Диссертация состоит из введения, четырех разделов, основных результатов и выводов по работе, списка литературы из 95 наименований и приложения. Работа изложена на 145 страницах основного текста, включающего 66 рисунков и 6 таблиц.

## ОСНОВНОЕ СОДЕЖАНИЕ РАБОТЫ

**Во введении** обоснована актуальность работы, сформулированы ее цель и задачи, научная новизна, описаны методы исследований, указана теоретическая и практическая значимость, а также выносимые на защиту научные положения, дается общая характеристика работы.

**В разделе 1** дан обзор признаков, используемых при выделении лиц на изображении человека, а также существующих методов обработки и анализа изображений, используемых для решения задач технического зрения.

Задача распознавания образов заключается в классификации изображений на основе определенных требований. Построение описания изображения на основе его представления с использованием признаков является едва ли не самой сложной задачей в процессе построения системы распознавания образов. Процедура выбора признаков до сих пор является эвристической и во многом зависит от предметной области.

Анализ особенностей, присущих изображениям лица, получаемым от видеокамер наблюдения, позволяет выявить геометрические искажения, имеющие различную природу. Дополнительные сложности вызывает необходимость адаптации к постоянно меняющемуся освещению. В результате анализа определены основные требования к системе идентификации: адаптация к изменяющимся условиям освещения, устойчивость к перспективному сокращению формы объекта и к геометрическим искажениям формы объекта, сохранение классифицирующих свойств при наличии факторов, усложняющих распознавание личности (головной убор, очки, борода и т.п.).

Для решения задачи выделения лица на изображении человека произведено исследование основных групп признаков, к которым отнесены:

- геометрические признаки, такие как ширина, высота, площадь, периметр объекта и пр.;
- топологические признаки, к которым относятся число связных элементов объекта, число отверстий в объекте, число Эйлера и т.д.;
- вероятностные признаки, в частности яркостные характеристики, текстурные характеристики, признаки стохастической геометрии и др.;
- спектральные признаки, получаемые применением разложения Карунена-Лоэва, преобразования Фурье, косинусного преобразования, полиномиальных моментов;
- цветовые признаки – яркость, насыщенность, тон и пр.

Проведенный анализ показал, что ни одна группа признаков в отдельности не может в полной мере одновременно удовлетворить требованиям точности и скорости выделения лица на изображении, получаемом с видеокамеры наблюдения. Решением этой проблемы является использование сочетания не-

скольких признаков для поэтапного уточнения области лица на изображении человека. Наибольший интерес в данном случае представляют цветковые признаки, так как их анализ занимает меньшее время. Эксперименты показали, что цвет кожи разных людей занимает достаточно небольшую ограниченную область цветового пространства даже при рассмотрении цветов кожи людей различных рас.

Так как лицо человека имеет определенную форму и структуру, для удаления ложно выделенных по цветовому признаку участков можно использовать геометрические и топологические признаки.

Для обоснованного выбора методов распознавания личности проведен обзор существующих методов анализа изображений в системах технического зрения. В частности были рассмотрены методы: главных компонент, линейного дискриминантного анализа, гибких контурных моделей, сравнения эластичных графов; методы, основанные на анализе геометрических признаков лица; метод скрытых марковских моделей. Для каждого метода были выявлены основные достоинства и недостатки, имеющие значение с точки зрения использования в системе распознавания личности.

В результате проведенного обзора обоснован выбор метода скрытых марковских моделей в качестве основы системы распознавания личности.

На сегодняшний день теория скрытых марковских моделей составляет основу индустриального стандарта для приложений, осуществляющих распознавание речи. СММ успешно применяются для решения таких задач, как распознавание жестов и распознавание рукописного текста. Несмотря на это, возможность применения СММ для распознавания лиц остается недостаточно исследованной. Ключевым вопросом является возможность применения СММ для обработки больших массивов данных в режиме реального времени.

**Раздел 2** посвящен разработке метода сегментации изображения, учитывающего цветковые, геометрические и топологические признаки лица.

Распознавание изображения по признаку цвета кожи человека является одним из наиболее эффективных подходов к поиску и выделению лица человека на изображении. Изменение состояния человека (эмоционального и физического) также слабо влияет на цвет его кожи.

Для отнесения пикселя к тому или иному классу по цветовому признаку использована операция пороговой бинаризации по диапазону. Проведенные исследования показали, что независимо от используемого метода цветового кодирования существует область перекрытия цвета фоновых объектов и цвета лица. Классификация пикселей, находящихся в области перекрытия, осуществляется путем анализа значений соседних элементов в соответствии с выражениями:

$$S_F = \sum_{i=-1}^1 \sum_{j=-1}^1 S_{ij(F)}; \quad S_B = \sum_{i=-1}^1 \sum_{j=-1}^1 S_{ij(B)}; \quad (1)$$

$$\begin{cases} S_{ij(N)} \in F, \text{ если } S_F \geq S_B \\ S_{ij(N)} \in B, \text{ в других случаях } \end{cases}$$



где  $S_F$  – количество соседних пикселей отнесенных к области лица;  $S_B$  – количество соседних пикселей, отнесенных к области фона;  $F$  – множество пикселей, отнесенных к области лица;  $B$  – множество пикселей, отнесенных к области фона.

После получения множества выделенных объектов выполняется уточнение их границ. Необходимость этого обусловлена тем, что в выборку по цветовому признаку могут не попасть затемненные или засвеченные области лица, также в одну область может попасть несколько расположенных вплотную лиц, в таком случае необходимо произвести их разделение. Далее происходит удаление областей, содержащих руки или другие открытые участки человеческого тела, по геометрическим признакам, таким как площадь  $Q$ , периметр  $L$  и отношение геометрических размеров  $C$ :

$$Q_{\text{лица}} \in [0,25Q_{\text{ср}}; 1,5Q_{\text{ср}}]; \quad (2)$$

$$C = \frac{4\pi Q}{L^2}; \quad (3)$$

$$0,8 \leq C_{\text{лица}} \leq 1,6, \quad (4)$$

где  $Q_{\text{лица}}$  – площадь, которая может соответствовать изображению лица;  $Q_{\text{ср}}$  – средняя площадь выделенных объектов;  $C_{\text{лица}}$  – коэффициент формы, свойственный изображению лица.

Для оценки количества неверно классифицированных пикселей изображения области лица, была получена формула

$$E_f = \frac{|(F_s \cap (F - F_f))|}{|F_f|} \cdot 100\%, \quad (5)$$

где  $F_f$  – множество пикселей, относящихся к области лица;  $F_s$  – множество пикселей, отнесенных к области лица;  $F$  – общее количество пикселей.

Оценка количества неверно классифицированных пикселей, относящихся к области фона, выполняется по выражению

$$E_b = \frac{|(F_f \cap (F - F_s))|}{|F - F_f|} \cdot 100\%. \quad (6)$$

На основании выражений (5) и (6) определено общее количество ошибок

$$E = \frac{(E_f \cdot |F - F_f| + E_b \cdot |F_f|)}{|F|}. \quad (7)$$

Как показали эксперименты, качество выделения области изображения по цветовому признаку во многом определяется степенью локализации цвета кожи лица человека внутри цветового пространства (рисунок 1).

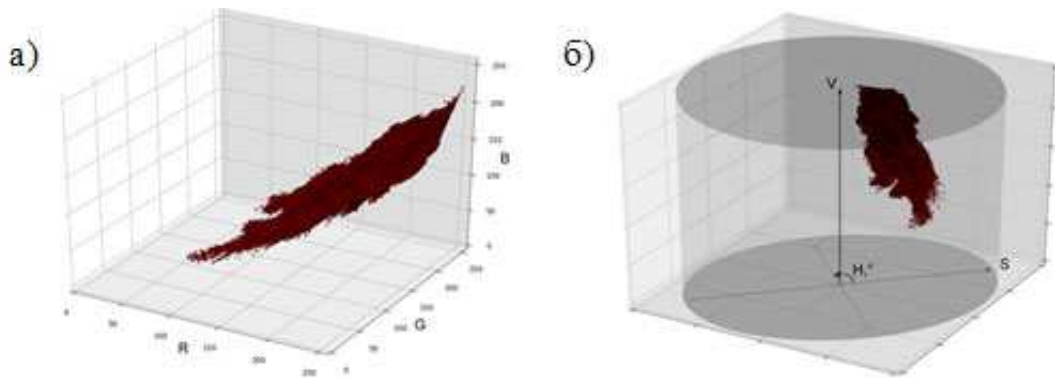


Рисунок 1 – Распределение цветов, характерных для кожи лица человека, в цветовом пространстве: а) *RGB*; б) *HSV*

Анализ существующих методов цветового кодирования показал возможность повышения плотности локализации цвета, характерного для кожи лица человека. Для этого был разработан алгоритм преобразования 3-х мерного *RGB* цветового пространства в 2-х мерное пространство, описывающее цвет с использованием полярных координат. На первом шаге выполняется преобразование *RGB*-куба в *RGB*-треугольник:

$$\begin{aligned} \begin{bmatrix} \varphi_R \\ r_R \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 90^\circ \\ 0 \end{bmatrix} = \begin{bmatrix} 90^\circ \\ R \end{bmatrix}; \\ \begin{bmatrix} \varphi_G \\ r_G \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 210^\circ \\ 0 \end{bmatrix} = \begin{bmatrix} 210^\circ \\ G \end{bmatrix}; \\ \begin{bmatrix} \varphi_B \\ r_B \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 330^\circ \\ 0 \end{bmatrix} = \begin{bmatrix} 330^\circ \\ B \end{bmatrix}. \end{aligned} \quad (8)$$

Направления векторов *R*, *G* и *B* являются неизменными, а их значения лежат в пределах от 0 до 255. При этом различные комбинации *R*, *G* и *B* задают различные цвета, различаются также формы треугольников. Треугольники различной формы имеют различный центр массы, называемый центроидом, при этом множество центроидов всех возможных *RGB*-треугольников образует двумерное цветовое пространство. Координаты центроида треугольника определяются в соответствии с формулами:

$$\begin{aligned} x &= \frac{r_R \cos(\varphi_R) + r_G \cos(\varphi_G) + r_B \cos(\varphi_B)}{3}; \\ y &= \frac{r_R \sin(\varphi_R) + r_G \sin(\varphi_G) + r_B \sin(\varphi_B)}{3}. \end{aligned} \quad (9)$$

Подставив в (9) значения углов поворота векторов  $\varphi_R$ ,  $\varphi_G$ ,  $\varphi_B$  и осуществив масштабное преобразование для сокращения количества вычислений, получим

$$\begin{aligned} x &= -\frac{\sqrt{3}}{2}(r_G - r_B); \\ y &= r_R - 0,5(r_G + r_B). \end{aligned} \quad (10)$$

Переходя от декартовой к полярной системе координат, получим

$$r = \sqrt{x^2 + y^2} = \sqrt{r_R^2 + r_G^2 + r_B^2 - r_R r_G - r_R r_B};$$

$$\varphi = \arctg\left(\frac{y}{x}\right) = \arctg\left(-\frac{2r_R - (r_G + r_B)}{\sqrt{3}(r_G - r_B)}\right).$$
(11)

На рисунке 2 представлено распределение характерных для кожи лица человека цветов, которое демонстрирует высокую плотность их расположения внутри разработанного специализированного цветового пространства (СЦП).

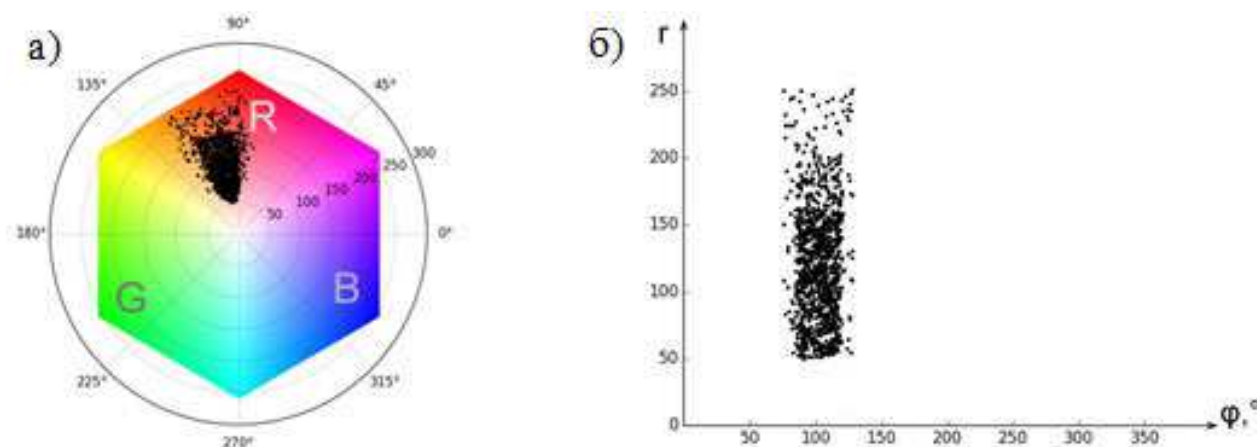


Рисунок 2 – Распределение цветов, характерных для кожи лица человека, в разработанном цветовом пространстве: а) в полярной системе координат; б) развернутое на плоскость

В таблице 1 приведены результаты оценки качества сегментации изображения, полученные при использовании различных методов цветового кодирования. Наилучшие результаты были получены при использовании разработанного специализированного цветового пространства. Диаграмма разработанного алгоритма выделения лица человека представлена на рисунке 3.

Таблица 1 – Результаты исследования цветовых пространств

Цветовое пространство	Ошибки классификации, %		Суммарная ошибка, %	Ошибки классификации пикселей, %
	1 рода	2 рода		
<i>RGB</i>	7,3	12,4	19,7	30,1
<i>HSV</i>	5,5	7,1	12,6	18,2
<i>HSI</i>	5,4	7,4	12,8	17,9
<i>HLS</i>	4,3	7,2	11,5	17,4
<i>YCbCr</i>	1,8	3,6	5,4	8,8
<i>CIE Lab</i>	5,4	6,1	11,5	19,1
СЦП	1,4	3,6	5,0	8,9

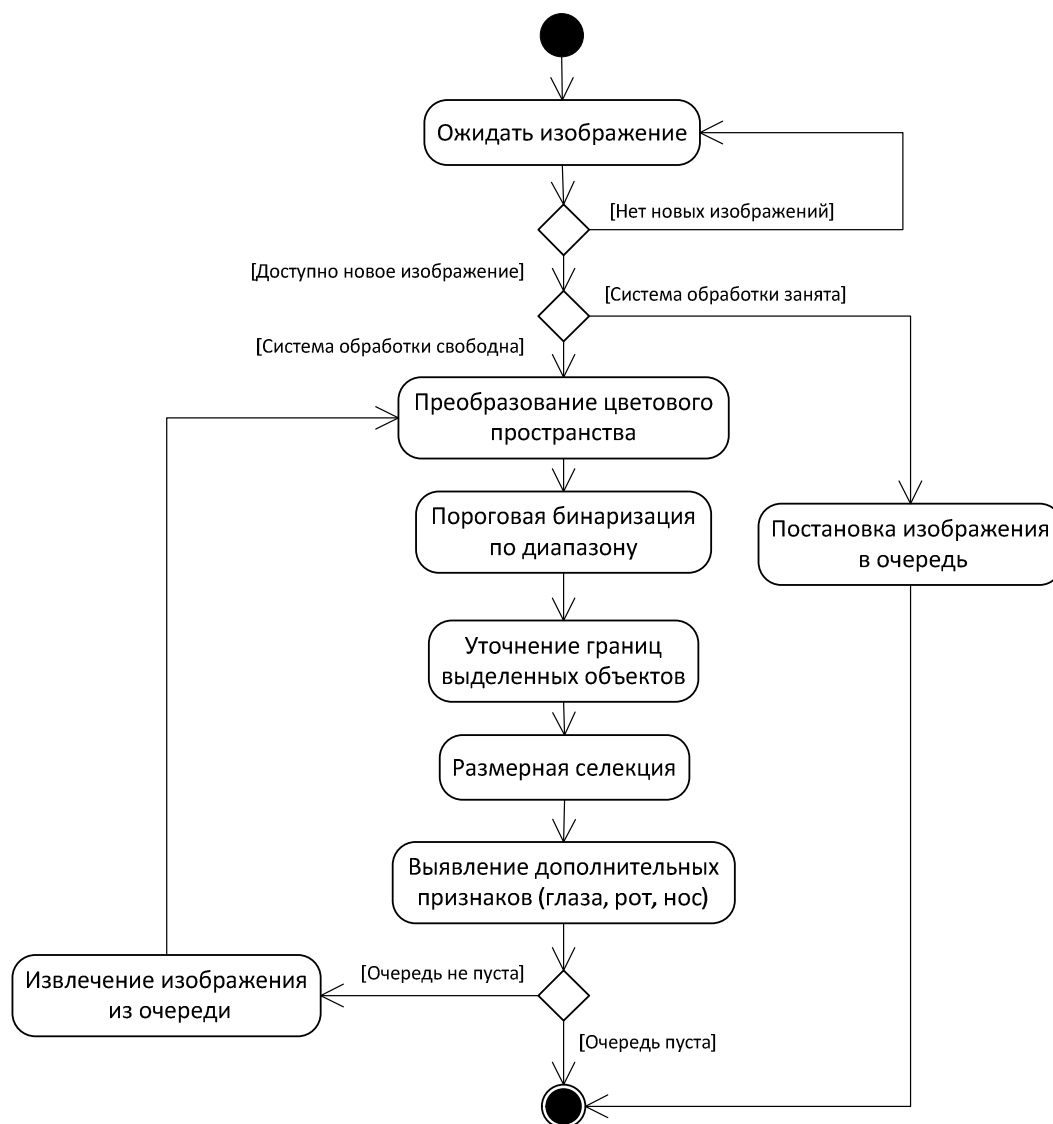


Рисунок 3 – Activity-диаграмма разработанного алгоритма выделения лица на изображении

**Раздел 3** посвящен разработке группы алгоритмов для выделения ключевых признаков, присущих конкретному изображению лица, а также поиска их соответствий характеристикам изображения лица человека в базе данных.

В общем случае процесс распознавания личности состоит из нескольких этапов: получение изображения лица, выделение ключевых признаков, поиск соответствий в базе данных, формирование результата распознавания. Рассматриваемая модель распознавания лица состоит из двух частей.

Первая часть (тренировка) служит для обучения модели, для чего на вход модели подаются известные изображения людей. В случае использования скрытой марковской модели процесс обучения включает следующие пять этапов: задание начальных параметров СММ; разбиение входной последовательности наблюдений на состояния с помощью алгоритма Витерби; основанная на алгоритме  $k$ -средних приближенная оценка вероятности появления наблюдаемых признаков; переоценка параметров системы с помощью алгоритма Баума-Велша; оценка вероятности генерирования наблюдаемой последовательности полученной СММ (рисунок 4, а).

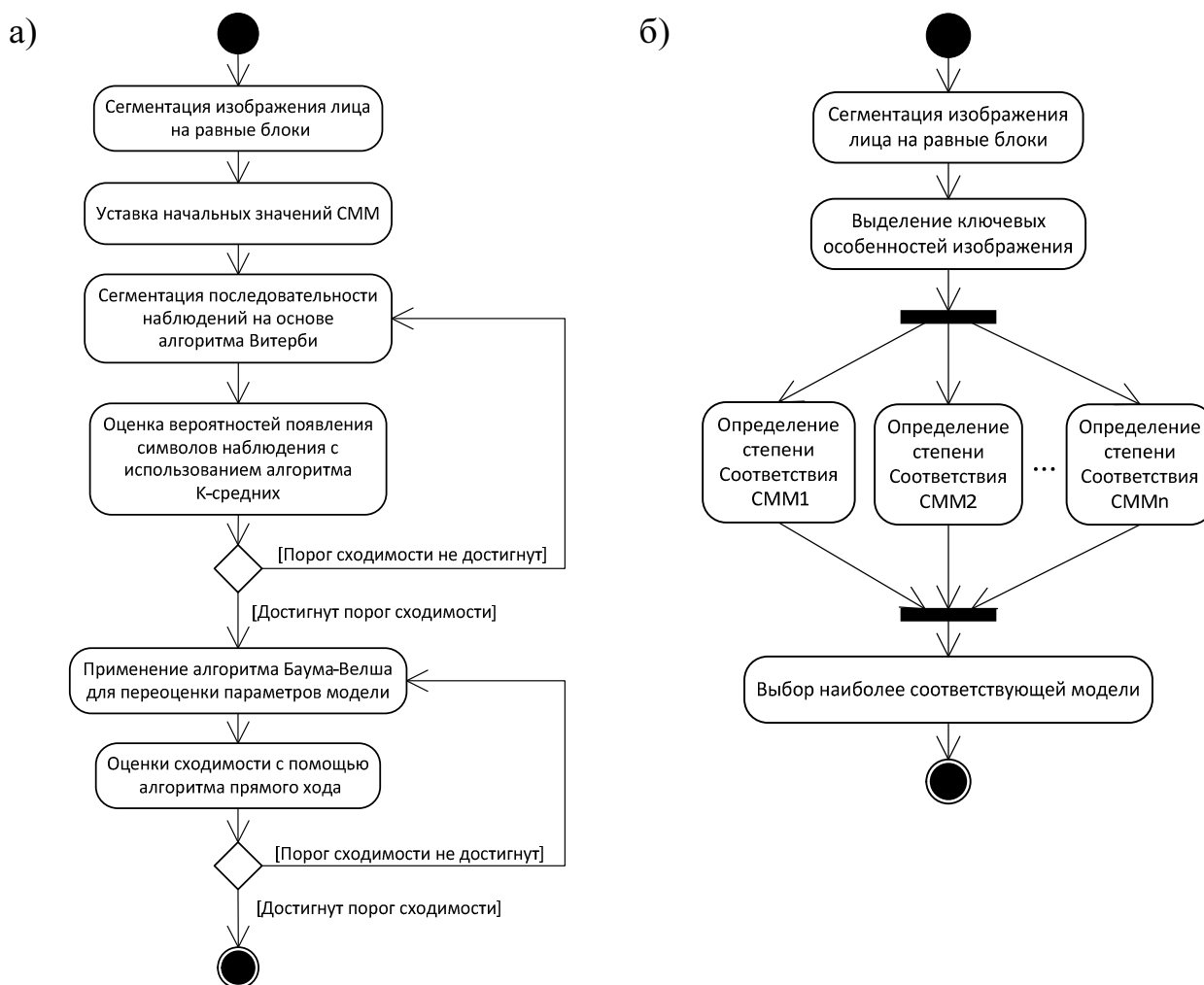


Рисунок 4 – Activity-диаграммы: а) алгоритма описания особенностей изображения лица в терминах СММ; б) алгоритма поиска соответствий

Вторая часть модели необходима для тестирования. При тестировании на вход модели поступает изображение неизвестного человека. Модель определяет значение вероятности соответствия данным по каждому содержащемуся в обучающем наборе человеку, то есть о степени схожести распознаваемого изображения с изображениями, имеющимися в базе. В случае использования СММ суть алгоритма поиска заключается в определении вероятности порождения наблюдаемой последовательности одной из последовательностей, хранимых в базе данных. Для этого так же, как на этапе обучения, происходит сегментация изображения и выделение ключевых признаков для каждой области. После этого с помощью алгоритма прямого хода оценивается вероятность  $P(O|\lambda)$  генерации выделенной последовательности признаков каждой из имеющихся в базе данных СММ (рисунок 4, б).

Для любой фиксированной последовательности состояний  $Q = \{q_1, q_2, \dots, q_T\}$  вероятность ее возникновения в данной модели  $\lambda$  равна

$$P(Q|\lambda) = \pi_{q_1} \cdot a_{q_1, q_2} \cdot a_{q_2, q_3} \cdot \dots \cdot a_{q_{T-1}, q_T}. \quad (12)$$

Вероятность же появления заданной последовательности наблюдения  $O = \{o_1, o_2, \dots, o_T\}$  для этой фиксированной последовательности  $Q$  составляет

$$P(O|Q, \lambda) = b_{o_1, q_1} \cdot b_{o_2, q_2} \cdot \dots \cdot b_{o_T, q_T}. \quad (13)$$

Поскольку для СММ возникновение некоторой конкретной последовательности состояний и появление последовательности наблюдений являются независимыми событиями, то окончательная вероятность  $P(O|\lambda)$  есть сумма совместных вероятностей  $P(O, Q|\lambda)$  для всех возможных последовательностей состояний  $Q = \{q_1, q_2, \dots, q_T\}$  длиной  $T$ :

$$P(Q|\lambda) = \sum_Q P(O, Q|\lambda), \quad (14)$$

а совместная вероятность этих событий в силу их независимости будет

$$P(O, Q|\lambda) = P(O|Q, \lambda)P(Q|\lambda), \quad (15)$$

из чего следует, что

$$P(O|\lambda) = \sum_Q \pi_{q_1} b_{o_1, q_1} a_{q_1, q_2} \cdot \dots \cdot b_{o_T, q_T} a_{q_{T-1}, q_T}. \quad (16)$$

В общем случае каждое состояние СММ может быть достигнуто за один шаг из любого другого состояния. На практике моделируемый сигнал может не отвечать данному требованию. Если говорить об изображении лица в системах распознавания, принято, как правило, разделять его на ряд вертикальных областей. В большинстве случаев выделяются следующие области: «волосы», «лоб», «глаза», «нос», «рот». Кроме этого иногда выделяют область «брови» и «подбородок». Эксперименты показали, что преимущество модели с семью состояниями по сравнению с моделью, состоящей из пяти состояний, проявляется только для изображений, полученных в стационарных условиях. Кроме того, такая модель более чувствительна к поворотам изображения и перспективным искажениям, обусловленным условиями установки видеокамер в системах видеонаблюдения. Кроме того, работа модели с пятью состояниями занимает меньше времени, что объясняет выбор в ее пользу.

В рамках СММ представленная структура изображения может быть описана с помощью «лево-правой» модели (рисунок 5, а), когда каждое состояние СММ соответствует одной из областей изображения. Другим вариантом является использование псевдодвумерной модели (рисунок 5, б).

Псевдодвумерная СММ является обобщением СММ, где каждое состояние само по себе является одномерной СММ. Таким образом, псевдодвумерная СММ состоит из набора супер-состояний, каждое из которых содержит набор вложенных состояний. Подобная структура позволяет не только наилучшим образом отобразить вертикальную структуру изображения, но и ее горизонтальную структуру. Хотя псевдодвумерная модель СММ имеет более сложную структуру по сравнению с «лево-правой» моделью, она в большей степени отражает характерные особенности изображения лица, благодаря дополнительному учету горизонтальной структуры изображения.

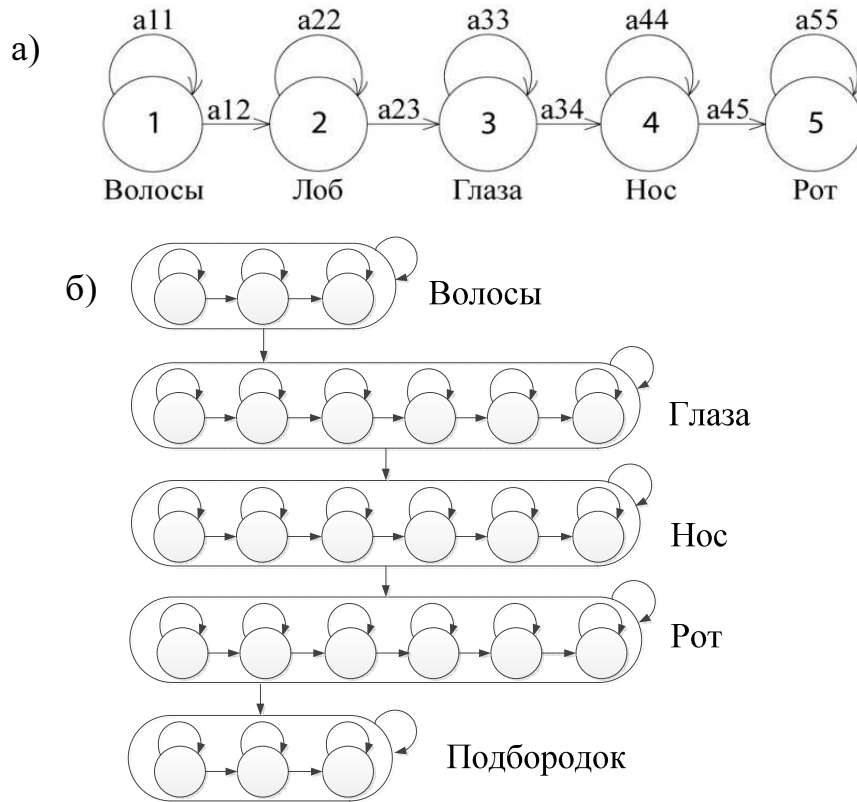


Рисунок 5 – Структура СММ, основанная на строении лица человека:  
 а) «лево-правая»; б) псевдодвумерная

Для обеспечения высокой точности распознавания необходимо не только выбрать наиболее подходящую структуру, но и оптимизировать набор признаков, формирующих алфавит скрытой марковской модели. С этой целью была произведена оценка использования следующих признаков: значения яркости пикселей; те же значения, получаемые после применения преобразования Хаара; значения коэффициентов дискретного косинусного преобразования; значения полиномиальных коэффициентов; значения элементов аппроксимирующей матрицы, полученной с помощью сингулярного разложения.

Лучший результат был получен при использовании полиномиальных коэффициентов  $\Delta$ . В технике распознавания речи данный метод применяется как основное средство устранения фонового шума. Однако ввиду того, что изображение по своей сути является двумерным сигналом, данный метод потребовал доработки.

Для описания изображения предложено определять  $n$  горизонтальных коэффициентов  $\Delta$ . Так для блока с координатами  $(b, a)$  полиномиальный коэффициент определяется как

$$\Delta_{c_n}^h(b,a) = \frac{\sum_{k=-K}^K k h_k c_n^{(b,a+k)}}{\sum_{k=-K}^K h_k k^2}. \quad (17)$$

Аналогично  $n$ -ый вертикальный коэффициент  $\Delta$  определяется выражением

$$\Delta_{c_n}^v = \frac{\sum_{k=-K}^K k h_k c_n^{(b+k,a)}}{\sum_{k=-K}^K h_k k^2}, \quad (18)$$

где  $h$  – вектор размерностью  $2K+1$ , а  $c_n$  –  $n$ -ый коэффициент дискретного косинусного преобразования (ДКП) в точке с координатами  $(b, a)$ . Двумерное ДКП для изображения  $F$  размером  $M \times N$  пикселей выглядит следующим образом:

$$C(u, v) = a(u)\beta(v) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} F(i, j) \cos \frac{(2i+1)u\pi}{2M} \cos \frac{(2j+1)v\pi}{2N}, \quad (19)$$

где

$$a(0) = \sqrt{\frac{1}{M}}, a(u) = \sqrt{\frac{2}{M}}; \quad (20)$$

$$\beta(0) = \sqrt{\frac{1}{N}}, \beta(v) = \sqrt{\frac{2}{N}}, \quad (21)$$

и

$$0 \leq u \leq M - 1, \quad 0 \leq v \leq N - 1. \quad (22)$$

Для распознавания изображения лиц были выбраны следующие значения:  $K=1$  и  $h=[1, 1, 1]^T$ , что соответствует прямоугольному окну. Чтобы продемонстрировать преимущества такого подхода, предположим, что имеются три следующих друг за другом блока:  $X$ ,  $Y$  и  $Z$ , при этом каждый блок содержит как информационную составляющую, так и шум:

$$\begin{aligned} X &= X_I + X_N; \\ Y &= Y_I + Y_N; \\ Z &= Z_I + Z_N. \end{aligned} \quad (23)$$

Если принять значения интенсивности шума во всех блоках равными:

$$X_N = Y_N = Z_N, \quad (24)$$

что логично, если блоки достаточно малы и расположены рядом, как в случае применения ДКП, то значения коэффициентов  $\Delta$  блока  $Y$  могут быть вычислены как:

$$\Delta^h Y = \frac{1}{2}(-X + Z) = \frac{1}{2}(-X_I - X_N + Z_I + Z_N) = \frac{1}{2}(Z_I + X_I); \quad (25)$$

$$\Delta^v Y = \frac{1}{2}(-X + Z) = \frac{1}{2}(-X_I - X_N + Z_I + Z_N) = \frac{1}{2}(Z_I + X_I). \quad (26)$$

В результате шумовая составляющая оказывается удалена. Далее необходимо изменить вектор наблюдений, состоящий из коэффициентов ДКП, путем замены первых трех элементов их горизонтальными и вертикальными коэффициентами  $\Delta$ .

Таким образом, для блока с координатами  $(b, a)$  получим новый вектор наблюдений, описываемый уравнением:

$$x = [\Delta_{c_0}^h, \Delta_{c_0}^v, \Delta_{c_1}^h, \Delta_{c_1}^v, \Delta_{c_2}^h, \Delta_{c_2}^v, c_3, c_4, \dots, c_{L-1}]^T. \quad (27)$$



Результаты работы системы распознавания, основанной на «лево-правой» и псевдодвумерной моделях СММ в случае использования различных признаков, представлены в таблице 2.

Таблица 2 – Результаты работы алгоритма распознавания личности

Наблюдаемый признак	«Лево-правая» модель, %	Псевдодвумерная модель, %
Яркость пикселей	62,37	63,31
Яркость пикселей после преобразования Хаара	68,46	83,72
Коэффициенты дискретного косинусного преобразования	70,93	87,80
Результат дискретного косинусного преобразования с коэффициентами $\Delta$	87,80	93,31
Значения элементов аппроксимирующей матрицы, полученной с помощью сингулярного разложения	85,42	91,25

**Раздел 4** посвящен разработке программного обеспечения и проверке путем моделирования разработанных методов и алгоритмов.

С этой целью было разработано программное средство для выделения признаков на изображении лица человека по цветовым критериям и распознавания личности на основе скрытых марковских моделей, представляющее собой расширение для системы *MATLAB*. Программное средство включает в себя следующий набор подпрограмм: предварительной обработки изображения, сегментации изображения, выделения признаков и поиска соответствий в базе данных.

Предобработка изображения выполняется в тех случаях, когда настройки видеокамеры аппаратно-программного комплекса не позволяют получить изображения надлежащего качества. После улучшения качества изображения выполняется выделение области лица на изображении с использованием цветовых, геометрических и топологических признаков, присущих изображению лица человека. Пример результата работы программы сегментации представлен на рисунке 6.



Рисунок 6 – Результат работы алгоритма выделения лица на изображении человека

Проведенные опыты показали, что при увеличении числа людей на изображении происходит существенный рост числа вычислительных операций, в

результате растет очередь на входе подсистемы распознавания, что в конечном итоге приводит к пропуску отдельных кадров, получаемых от системы видеонаблюдения. Наиболее очевидным решением данной проблемы является реализация разработанных алгоритмов на компилируемом языке программирования, поскольку система *MATLAB* ориентирована, прежде всего, на удобство и эффективность использования, а не на высокую производительность.

Для оценки прироста производительности разработанный алгоритм сегментации изображения был реализован на языке *C*. По сравнению с реализацией алгоритма в программном пакете *MATLAB* производительность вычислительных процедур в среднем возросла на 36 %.

В качестве дополнительного способа повышения быстродействия был разработан метод и алгоритм отслеживания перемещения. Применение данного метода связано с тем, что с того момента, как человек входит в поле зрения видеокамеры, и до момента, когда человек покидает его, может пройти неограниченное время. Все это время выполняется многократный процесс установления его личности.

Алгоритм отслеживания перемещения строится на предположении о том, что за время, проходящее между сменой двух кадров, объект не может существенным образом изменить свою форму и положение в пространстве (рисунок 7).

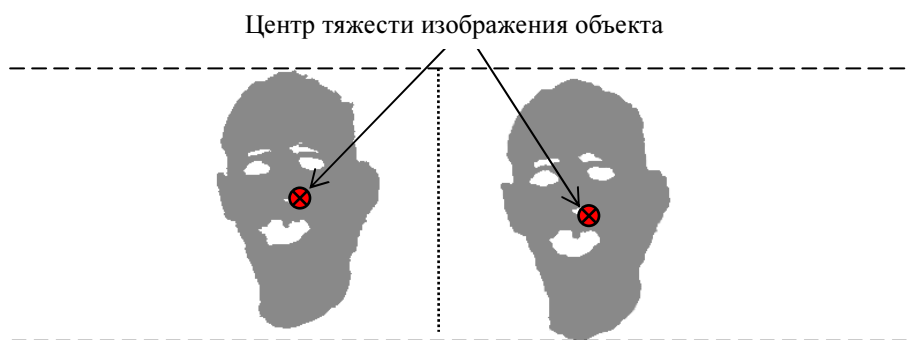


Рисунок 7 – Изменение координат центра тяжести объекта при смене кадров

Для отслеживания положения субъектов используются координаты центра тяжести, так как они статистически более устойчивы, чем координаты геометрического центра массы. Для их вычисления используются интегральные характеристики – сумма абсолютных координат по ортогональным направлениям и площадь объекта:

$$i_{\text{ц.т.}}^{(q)} = \frac{\sum_i \sum_j S_{i,j}^{(q)} \cdot i^{(q)}}{\sum_i \sum_j S_{i,j}^{(q)}}; \quad (28)$$

$$j_{\text{ц.т.}}^{(q)} = \frac{\sum_i \sum_j S_{i,j}^{(q)} \cdot j^{(q)}}{\sum_i \sum_j S_{i,j}^{(q)}}; \quad (29)$$

$$X_{\text{ц.т.}}^{(q)} = K_x \cdot i_{\text{ц.т.}}^{(q)}; \quad Y_{\text{ц.т.}}^{(q)} = K_y \cdot j_{\text{ц.т.}}^{(q)}. \quad (30)$$

Решение о совпадении объектов на соседних кадрах принимается на основе информации об их центре тяжести и геометрической форме. Диаграмма описанного алгоритма отслеживания перемещения представлена на рисунке 8.

Опыты по использованию алгоритма отслеживания перемещения в разработанном прототипе системы установления личности показали сокращение очереди к подпрограмме распознавания в среднем на 28 %.

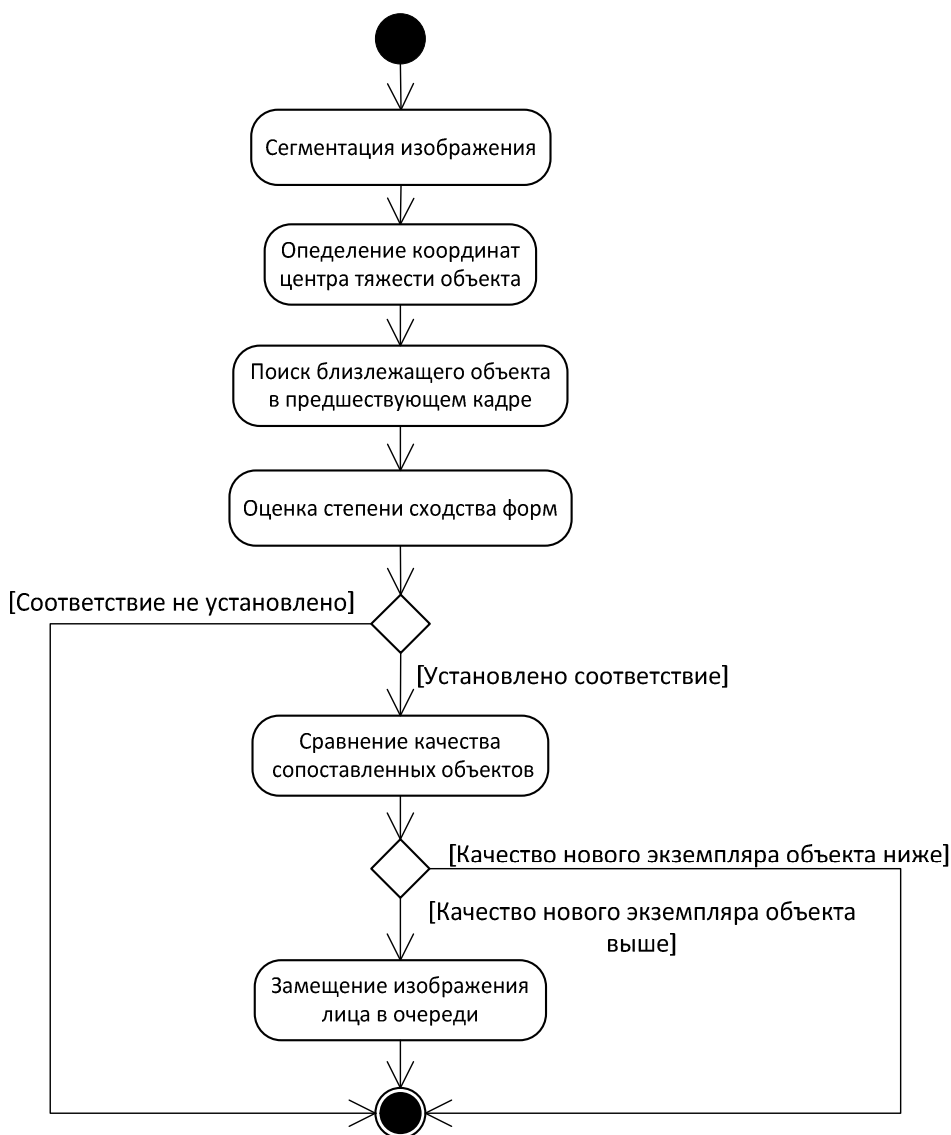


Рисунок 8 – *Activity*-диаграмма алгоритма отслеживания перемещения

Для оценки эффективности работы алгоритма выделения изображения лица и распознавания личности было использовано четыре независимых набора изображений, а также их объединение (таблица 3). Помимо этого проводились испытания с изображениями, по своему содержанию приближенными к изображениям, получаемым с камеры видеонаблюдения. Испытания подтвердили возможность применения разработанных алгоритмов для решения практических задач. На рисунке 9 представлен фрагмент обучающей последовательности, на рисунке 10 – результат распознавания изображения.

Таблица 3 – Обобщенные результаты проведенного эксперимента

Тестовая база	Процент правильно распознанных изображений	Количество изображений, обработанных за 1 секунду
<i>Cohn-Kanade AU-Coded Facial Expression Database</i>	96,9%	3,55
<i>MIT-CBCL Face Recognition Database</i>	95,65%	4,59
<i>Face Recognition Data, University of Essex, UK</i>	93,31%	4,18
Обобщенный набор изображений	93,42%	1,89



Рисунок 9 – Фрагмент обучающей последовательности

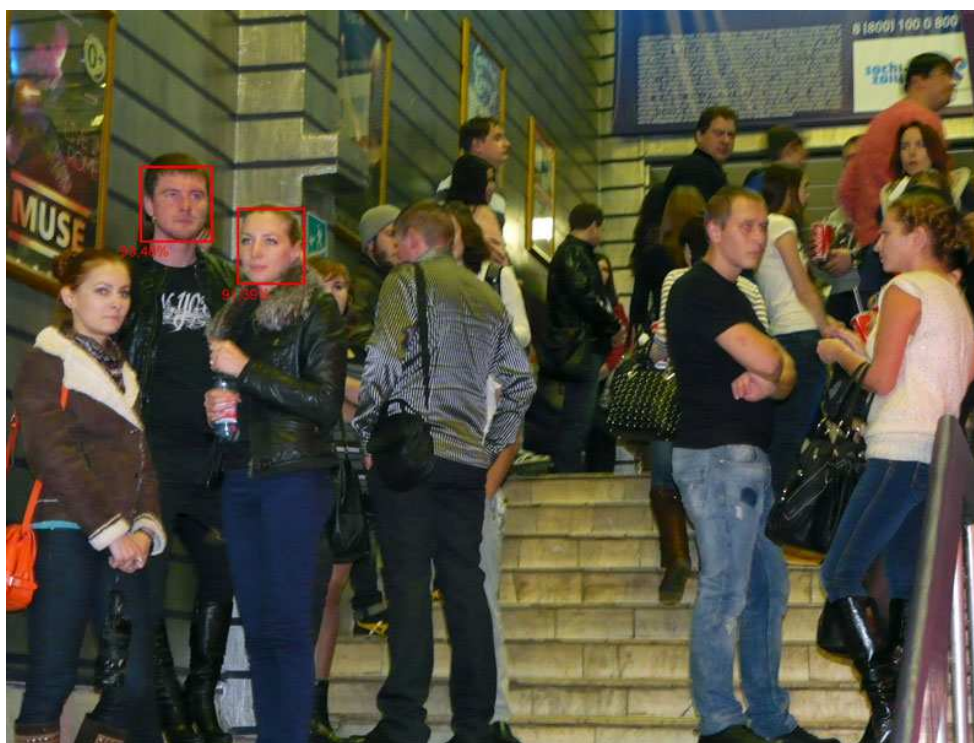


Рисунок 10 – Результат распознавания изображения

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. На основе данных, полученных в результате исследования областей локализации цветов, характерных для кожи лица человека в разных цветовых пространствах, разработаны метод и алгоритм цветового кодирования, обеспечивающие выявление характерных цветовых признаков на изображении лица человека.

2. Разработан алгоритм сегментации изображения лица человека, учитывающий цветовые, геометрические и топологические признаки лица на изображении, получаемом с видеокамеры наблюдения.

3. Обоснована математическая модель описания изображения лица на основе скрытой марковской модели, а также доказана эффективность применения псевдодвумерной структуры модели, содержащей пять супер-состояний, соответствующих областям волос, лба, глаз, носа и рта.

4. В результате сравнительного анализа доказана эффективность использования двумерного дискретного косинусного преобразования для сокращения объема анализируемой информации, подавления шумов на изображении, а также уменьшения влияния таких факторов, как перспективное сокращение и геометрическое искажение формы лица, незначительный поворот или наклон лица.

5. Разработаны метод и алгоритм распознавания изображенного лица, основанные на выявлении горизонтальной структуры лица, последующем выделении присущих каждой составляющей части изображения ключевых признаков и поиске соответствий кодов признаков их эталонам в базе данных. Оценка быстродействия алгоритма показала возможность его применения в приложениях, работающих в реальном времени.

6. Разработаны метод снижения вычислительной нагрузки за счет сокращения повторных процедур распознавания одного и того же субъекта и реализующий его алгоритм слежения за перемещением во времени области лица человека на изображении.

7. Разработаны программы для исследования методов выделения лица на изображении человека по цветовым признакам и распознавания личности на основе скрытых марковских моделей, позволяющие опытным путем определить основные требования к исходному изображению, получаемому от видеокамеры наблюдения.

8. На основе разработанного алгоритма сегментации изображения по цветовым признакам показано, что производительность вычислений в специализированной программе существенно повышается по сравнению с использованием универсальной вычислительной среды *MATLAB*.

## **ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ**

### **Публикации в изданиях, рекомендованных ВАК**

1. Двойной, И.Р. Выбор системы цветового кодирования для обнаружения изображения лица человека на основе его цвета [Текст] / И.Р. Двойной // XXI век: итоги прошлого и проблемы настоящего плюс. Серия: технические науки. Информационные технологии. – 2012. – № 05 (09). – С. 50–54.

2. Двойной, И.Р. Использование скрытых марковских моделей для установления личности человека по изображению его лица [Текст] / И.Р. Двойной // XXI век: итоги прошлого и проблемы настоящего плюс. Серия: технические науки. Информационные технологии. – 2013. – № 10 (14). – С. 33–38.

3. Двойной, И.Р. Сравнительный анализ структур скрытых марковских моделей, используемых в задаче установления личности человека по изображе-

нию лица [Электронный ресурс] / И.Р. Двойной, И.И. Сальников // Современные проблемы науки и образования. – 2013. – № 6. – URL: [www.science-education.ru/113-10751](http://www.science-education.ru/113-10751) (дата обращения: 13.11.2013).

#### **Публикации в других изданиях**

4. Двойной, И.Р. Программа анализа времени выполнения вычислительных операций в заданном алгоритме [Текст] / И.Р. Двойной, И.И. Сальников // Современные методы и средства обработки пространственно-временных сигналов: Сборник статей VIII Всероссийской научно-технической конференции. – Пенза: Приволжский дом знаний, 2010. – С. 76–80.

5. Двойной, И.Р. Анализ цветовых моделей, получивших широкое распространение, и причин их появления [Текст] / И.Р. Двойной, В.С. Редькин // Журнал научных публикаций аспирантов и докторантов. – 2012. – № 3. – С. 131-133.

6. Двойной, И.Р. Полярная система координат для представления цветовых характеристик в системе распознавания лиц [Текст] / И.Р. Двойной // Современные методы и средства обработки пространственно-временных сигналов: Сборник статей X Всероссийской научно-технической конференции. – Пенза: Приволжский дом знаний, 2012. – С. 60–63.

7. Двойной, И.Р. Использование полярной цветовой модели для выделения лица на изображении [Электронный ресурс] / И.Р. Двойной, И.И. Сальников // Студенческий научный форум: IV международная студенческая электронная научная конференция. – URL: <http://www.rae.ru/forum2012/218> (дата обращения: 07.08.2013).

8. Двойной, И.Р. Метод идентификации лиц на основе скрытых марковских моделей [Текст] / И.Р. Двойной // Современные методы и средства обработки пространственно-временных сигналов: Сборник статей XI Всероссийской научно-технической конференции. – Пенза: Приволжский дом знаний, 2013. – С. 61–64.

9. Двойной, И.Р. Отслеживание перемещения объектов в системах идентификации, для предотвращения повторного распознавания [Текст] / И.Р. Двойной // Журнал научных публикаций аспирантов и докторантов. – 2013. – № 11. – С. 119–123.

**ДВОЙНОЙ Илья Ростиславович**

**МЕТОДЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЯ ЛИЦА ЧЕЛОВЕКА ПО  
ЦВЕТОВЫМ ПРИЗНАКАМ И ИДЕНТИФИКАЦИИ ЛИЧНОСТИ НА  
ОСНОВЕ СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ В СИСТЕМАХ  
ВИДЕОНАБЛЮДЕНИЯ**

Специальность 05.13.17 – теоретические основы информатики

**Автореферат  
диссертации на соискание ученой степени  
кандидата технических наук**

Компьютерная верстка Т.А. Антиповой

Сдано в производство 12.11.13. Формат 60x84 <sup>1</sup>/<sub>16</sub>  
Бумага типогр. № 1. Печать трафаретная. Шрифт Times New Roman Cyr.  
Усл. печ. л. 1,39. Уч.-изд л. 1,29. Заказ № 2376. Тираж 100.

---

Пензенский государственный технологический университет.  
440039, Россия, г. Пенза, пр. Байдукова/ул. Гагарина, 1<sup>а</sup>/11

