

На правах рукописи

Нгуен Зуй Тхань

**МЕТОДЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ ЛИЦ НА ОСНОВЕ
ИНВАРИАНТОВ К АФФИННЫМ И ЯРКОСТНЫМ ПРЕОБРАЗОВАНИЯМ**

05.13.17 – Теоретические основы информатики

Автореферат

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2018

Работа выполнена на кафедре информационных технологий Российского университета дружбы народов

Научный руководитель: **Хачумов Вячеслав Михайлович**,
доктор технических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Официальные оппоненты: **Тулупьев Александр Львович**,
доктор физико-математических наук, доцент, главный научный сотрудник с возложенными обязанностями заведующего лабораторией теоретических и междисциплинарных проблем информатики Федерального государственного бюджетного учреждения науки «Санкт-Петербургский институт информатики и автоматизации Российской академии наук»

Конушин Антон Сергеевич,
кандидат физико-математических наук, доцент, заведующий лабораторией компьютерной графики и мультимедиа факультета вычислительной математики и кибернетики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В.Ломоносова»

Ведущая организация: **Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова Российской академии наук (ИПУ РАН)**

Защита диссертации состоится «21» декабря 2018 г. в 15 часов 30 минут на заседании диссертационного совета Д 212.203.28 на базе Российского университета дружбы народов по адресу г. Москва, ул. Орджоникидзе, дом 3, ауд. 110.

С диссертацией можно ознакомиться в научной библиотеке Российского университета дружбы народов по адресу: 117198, Москва, ул. Миклухо-Маклая, дом 6. или на официальном сайте диссоветов РУДН по адресу: <http://dissovet.rudn.ru/>

Отзывы на автореферат в двух экземплярах, заверенные печатью, направлять по адресу: 117198, Москва, ул. Миклухо-Маклая, дом 6., ученому секретарю диссертационного совета Д 212.203.28.

Автореферат разослан « ___ » _____ 2018г.

Ученый секретарь диссертационного совета,
кандидат физико-математических наук



Васильев С.А.

Общая характеристика работы

Актуальность исследования. В последнее время наблюдается повышение интереса к задаче распознавания лиц, сопровождаемое проведением российских и международных конкурсов. Так в 2015 году в России был проведен открытый конкурс Фонда перспективных исследований на лучший демонстрационный образец технологии распознавания лиц людей. В 2016 году прошло мировое соревнование по распознаванию лиц на данных «Megaface Challenge», организованное Вашингтонским университетом (University of Washington), а в 2017 году Американское агентство передовых исследований в сфере разведки (IARPA) совместно с Национальным институтом стандартов и технологий США (NIST) провело конкурс «Face Recognition Prize Challenge». На сегодняшний день имеется много различных методов и алгоритмов распознавания лиц, опирающихся на метод главных компонент, нейронные сети, Марковские цепи, эластичные графы, вейвлет-анализ, метод опорных векторов и другие инструментальные средства. Следует отметить подходы, связанные с выделением таких локальных особенностей лица, как губы, нос, овал или профиль, но ввиду множества факторов, усложняющих анализ изображений (зашумления, повороты, выражения различных эмоций, изменчивость освещения и др.), пока нет подхода, гарантирующего точное решение проблемы. В связи с этим все большее внимание уделяется 3D-моделям лиц, полученным в результате реконструкции с использованием камер высокого разрешения, применение которых в ряде случаев позволяет увеличить точность и полноту распознавания. Можно отметить рост числа работ по распознаванию лиц на основе различных инвариантов к яркостным и аффинным преобразованиям. Инварианты полезны тем, что способны при определенных условиях выступать в роли существенных признаков, сохраняющих и передающих характерные особенности изображений, подвергнутых преобразованиям, что увеличивает вероятность получения правильного решения. Здесь необходимо выделить работы ряда ученых (Путятин Е.П., Аверин С.И., Шмидт А.А., Hu M.K., Wong R.Y., Hall E.L., Zaeri N., Baker F., Dib R., Flusser J, Suk T, Zitová B.), в которых систематизируются знания о

применении различных инвариантов в задачах распознавания. Подход на основе инвариантов является одним из немногих подходов, допускающих точную математическую постановку задачи классификации. Однако, при ближайшем изучении оказывается, что описания инвариантов в различных источниках приведены без необходимых доказательств и содержат неточности, что требует фундаментальной проверки. Несмотря на создание новых алгоритмов, методов и подходов, проблема распознавания лиц до сих пор не решена в полном объеме.

В настоящей диссертационной работе основное внимание уделено задаче корректного сопоставления и распознавания изображений лиц с применением моментов как инвариантов к группе аффинных преобразований и линий положения как инвариантов к яркостным преобразованиям; а также анализу устойчивости 2D- и 3D-моделей лиц к аффинным и яркостным преобразованиям. Моменты не требуют совмещений и масштабирования объектов, могут служить интегральными дескрипторами изображений, но чувствительны к изменениям яркости. Линии положения, напротив, направлены на корректное совмещение изображений в условиях изменяющихся яркостей, что позволяет выполнять попиксельное сравнение.

Цель диссертационной работы

Целью диссертационной работы является исследование и разработка методов и алгоритмов распознавания лиц на основе инвариантов к аффинным и яркостным преобразованиям. Цель достигается решением следующих основных задач:

1. Анализ существующих методов распознавания 2D-и 3D-изображений лиц, выявление их преимуществ и недостатков.
2. Оценки точности восстановления (реконструкции) 3D-изображений по стереопаре в зависимости от выбранного критерия качества.
3. Исследование на инвариантность и отбор 2D-и 3D-моментов в качестве дескрипторов изображений, анализ их чувствительности к геометрическим и к яркостным искажениям.

4. Разработка и исследование наборов 2D-и 3D-инвариантных моментов совместно с различными классификаторами (метриками, нейронными сетями) для практического решения задач распознавания лиц.

5. Разработка и исследование методов сравнения графических объектов на основе построения геометрических моментов как инвариантов к аффинным преобразованиям и линий положения как яркостных инвариантов.

6. Разработка и исследование метода оптимального наведения 3D-моделей на 2D-изображения лиц с целью последующего распознавания.

7. Создание специализированного программного обеспечения для распознавания лиц человека на основе разработанных методов и алгоритмов.

Основные положения, выносимые на защиту:

1. Методы оценки качества трехмерной реконструкции изображений по стереоизображениям, формирование различных постановок оптимизационных задач.

2. Обоснование инвариантности 2D-и 3D-моментов в качестве дескрипторов изображений для решения задачи распознавания лиц.

3. Методы «грубого» (на основе линий положения) и «точного» (на основе наведения 3D-модели) совмещения изображений для решения задач распознавания.

4. Методы и алгоритмы решения задач распознавания лиц с применением 2D- и 3D-инвариантных моментов, линий положения, обобщенной метрикии нейросетевыхклассификаторов.

Научная новизна результатов диссертационной работы заключается в следующем:

1. Получены оценки устойчивости трехмернойреконструкции изображений лиц к флуктуациям при условии хорошей начальной обусловленности фундаментальной матрицы. Даны постановки оптимизационных задач трехмерной реконструкции для различных критериев качества.

2. Доказаны утверждения об инвариантности наборов 2D- и 3D-моментов идеальных изображений к группе аффинных преобразований и утверждения об

инвариантности линий положения к яркостным преобразованиям 2D- и 3D-изображений, что позволяет научно обоснованно использовать их для анализа и классификации лиц.

3. Получены оценки устойчивости 2D- и 3D-моментов к флуктуациям изображений, которые следует учитывать в задачах распознавания лиц.

4. Разработан и исследован алгоритм оптимизации наведения проекции 3D-модели изображения на 2D-изображение на основе среднеквадратичного критерия рассогласования моментов и пикселей.

Теоретическая значимость. Теоретическая значимость результатов диссертации состоит в разработке и исследовании интегрированных методов распознавания 2D- и 3D-изображений лиц с применением яркостных и геометрических инвариантов в сочетании с набором классификаторов на основе метрик и нейронных сетей.

Методы исследования. При выполнении диссертационной работы использовались методы вычислительной геометрии и компьютерной графики, алгебра матриц, методы теории кластеризации и распознавания образов, алгоритмы цифровой обработки изображений, технологии разработки программного обеспечения.

Обоснованность и достоверность результатов. Полученные в диссертации результаты обоснованы корректным использованием математических методов, постановкой и решением оптимизационных задач. Достоверность подтверждается проведенными математическими выкладками, доказательствами утверждений, результатами вычислительных экспериментов с использованием реальных баз изображений лиц.

Практическая значимость. На основе выполненного исследования разработано программное обеспечение, предназначенное для решения задач идентификации 2D- и 3D-изображений лиц. Экспериментально показано, что совместное применение инвариантных моментов и классификаторов в виде метрик Махаланобиса, Евклида-Махаланобиса и нейронных сетей позволяет решать практические задачи идентификации реальных изображений лиц,

содержащих гримасы и помехи, с достаточно высокой точностью. Проведенные серии экспериментов с фотографиями лиц показали, что доля ошибки распознавания составила от 6.67% до 11% в зависимости от условий эксперимента. Разработанные алгоритмы являются универсальными, их целесообразно использовать в составе поисковых систем на заключительных этапах после предварительного выделения лиц, что позволит повысить точность и полноту поиска. Работа выполнялась в рамках проекта «Разработка методов распознавания объектов на основе 3D-моделей и инвариантов к аффинным и яркостным преобразованиям с применением высокопроизводительных вычислений» по Программе ОНИТ-ИРАН «Интеллектуальные информационные технологии, системный анализ и автоматизация». Получено свидетельство о регистрации программы для ЭВМ.

Апробация работы. Результаты диссертационной работы обсуждались и докладывались на следующих конференциях и семинарах:

1. Всероссийская конференция (с международным участием) «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем» (ИТТММ-2015, 20–24 апреля 2015 г., г. Москва, РУДН); (ИТТММ-2016, 18–22 апреля 2016 г., г. Москва, РУДН);

2. IV Всероссийская научная конференция молодых ученых с международным участием «Информатика, управление и системный анализ» (ИУСА-2016, 8-11 июня 2016 г., г. Тверь, ТвГТУ);

3. 19-й Международная конференция «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь» (DCCN-2016, 21 – 25 ноября, 2016 г., г. Москва, РУДН);

4. Международная научно - практическая конференция «Концепции фундаментальных и прикладных научных исследований» (20 февраля 2017 г., г. Казань).

5. V Международной научно-практической конференции «Актуальные проблемы технических наук в России и за рубежом», (г. Новосибирск, февраль 2018).

6. XIX Международный научно-практический форум студентов, аспирантов и молодых ученых «Актуальные проблемы информатики и информационных технологий в образовании», Красноярск 22 Мая 2018 г.

7. Шестнадцатая Национальная конференция по искусственному интеллекту с международным участием КИИ-2018 (24–27 сентября 2018 г., г. Москва, Россия).

8. Семинары лаборатории 0-4 «Методы интеллектуального управления» ИСА ФИЦ ИУ РАН и кафедры «Информационные технологии» факультета ФМиЕН РУДН в (2015-2018г.г.).

Публикации. Основные результаты диссертационной работы изложены в 14 печатных работах, в том числе 4 статьи опубликованы в рецензируемых изданиях, рекомендованных ВАК РФ; 1 статья в журнале “Scientific and Technical Information Processing”; 5 статей в трудах международных и всероссийских научных конференций и 4 работы в виде тезисов докладов конференций. Получено 1 свидетельство о регистрации программы для ЭВМ.

Структура и объем работы. Диссертация содержит введение, четыре главы, заключение и список использованной литературы, содержащий 61 наименований. Общий объем диссертации составляет 95 страниц машинописного текста, иллюстрированного 16 рисунками и 20 таблицами. В Приложение вынесены доказательства инвариантности, коды программ и скан-копия свидетельства о регистрации программы для ЭВМ.

Содержание работы

Во введении обоснована актуальность темы диссертационной работы, приведены цель и методы исследования, сформулированы теоретическая и практическая значимость работы.

В первой главе представлен аналитический обзор методов распознавания 2D- и 3D-изображений лиц, выполнено их сравнение и обозначены преимущества и недостатки.

В разделах 1.1-1.3. дается анализ существующих методов распознавания лиц человека, выделены основные понятия и определения.

Приведем некоторые определения применительно к 3D-объектам, необходимые для дальнейших исследований.

Определение 1. Под 3D-изображением понимается кусочно-непрерывная функция $f(x, y, z)$ трех переменных, определенная на компактном носителе $D \subset R \times R \times R$ и имеющая конечный ненулевой интеграл.

Определение 2. Моменты 3D-изображения (*image moments*) – отображения («проекции») кусочно-непрерывной функции изображения f в полиномиальный базис, где P – множество многочленов, определенных на компактном носителе $D \subset R \times R \times R$.

Определение 3. Инвариант – функционал $I(F)$, определенный на изображении F таким образом, что для пространства всех допустимых преобразований D , в том числе операторов деградации, (ухудшения) изображения F , причем $I(F_1), I(F_2)$ должны быть «достаточно различны» для разных изображений F_1, F_2 .

В задачах распознавания образований информативно значимыми дескрипторами служат инварианты к яркостным и аффинным преобразованиям.

Определение 4. Яркостным инвариантом называется дескриптор изображения, который является устойчивым к аддитивным и/или мультипликативным изменениям яркости.

Определение 5. Инварианты моментов (*moment invariants*) или инвариантные моменты – это дескрипторы, построенные на основе степенных моментов и описывающие математическую (идеальную) модель непрерывного (аналогового) изображения F некоторого объекта ω и являющиеся инвариантными к его аффинным преобразованиям.

Определение 6. Субинвариантными моментами или (полуинвариантами моментов) цифрового изображения F назовем моменты $I(F)$, величины которых могут меняться при выполнении аффинных преобразований, не превышая при этом установленные допустимые величины отклонений Δ , приемлемые для решения задачи классификации.

В разделах 1.4-1.7 представлены различные виды инвариантных моментов. Задача распознавания лиц формулируется следующим образом: требуется построить распознающую функцию $F(\omega)$, выход которой $F_i(\omega)$ определяет класс Ω_i тестируемого изображения ω , представленного вектором признаков-инвариантов $(x_1(\omega), \dots, x_n(\omega))$, причем $F_k(\omega) = \begin{cases} 1, & \text{если } \omega \in \Omega_k \\ 0, & \text{если } \omega \notin \Omega_k \end{cases}$. Для решения задачи в диссертации применяются искусственные нейронные сети и обобщенная метрика. Сделаны основные выводы к Главе 1 и сформулированы задачи диссертационной работы.

В Главе 2 дается математическая постановка задачи построения (реконструкции) 3D-изображений лиц при различных критериях. Исследуется устойчивость построенных 3D-моделей к возможным флуктуациям данных.

В разделе 2.1 рассмотрены вопросы построения 3D-моделей путем трехмерной реконструкции изображения по стереопаре и проблема обусловленности фундаментальной матрицы, являющейся математической моделью стереопары. Трехмерная реконструкция сводится к решению задачи

$$AX = B, \quad (1)$$

$$\text{где } A = \begin{bmatrix} T_{11}^1 - T_{14}^1 x^{1*} & T_{21}^1 - T_{24}^1 x^{1*} & T_{31}^1 - T_{34}^1 x^{1*} \\ T_{12}^1 - T_{14}^1 y^{1*} & T_{22}^1 - T_{24}^1 y^{1*} & T_{32}^1 - T_{34}^1 y^{1*} \\ T_{11}^2 - T_{14}^2 x^{2*} & T_{21}^2 - T_{24}^2 x^{2*} & T_{31}^2 - T_{34}^2 x^{2*} \\ T_{12}^2 - T_{14}^2 y^{2*} & T_{22}^2 - T_{24}^2 y^{2*} & T_{32}^2 - T_{34}^2 y^{2*} \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad B = \begin{bmatrix} T_{44}^1 x^{1*} - T_{41}^1 \\ T_{44}^1 y^{1*} - T_{42}^1 \\ T_{44}^2 x^{2*} - T_{41}^2 \\ T_{44}^2 y^{2*} - T_{42}^2 \end{bmatrix}.$$

Здесь T_{ij}^k – коэффициент математической модели k -й камеры.

В разделе 2.2 дается оценка чувствительности модели к возмущениям. Пусть фундаментальная матрица A является хорошо обусловленной. Введем возмущающую матрицу A_1 , возмущающий вектор B_1 и потребуем, чтобы сингулярные числа матрицы A_1 , т.е. квадратные корни из собственных чисел матрицы $A_1^T A_1$ были ограничены сверху некоторой константой. Помимо системы (1) рассмотрим также систему (2)

$$(A + \varepsilon A_1)X_\varepsilon = B + \varepsilon B_1 \quad (2)$$

где ε – малый параметр.

Ее решение имеет вид $X_\varepsilon = ((A^T + \varepsilon A_1^T)(A + \varepsilon A_1))^{-1}(A^T + \varepsilon A_1^T)(B + \varepsilon B_1)$

Утверждение 1. Если коэффициенты матрицы A и столбец свободных членов B в уравнении (1) изменить на малые величины εA_1 и εB_1 , то и решение задачи реконструкции, полученное с помощью метода наименьших квадратов, изменится на малое значение

$$\Delta = O(\varepsilon) = \|(A + \varepsilon A_1)^T (A + \varepsilon A_1)\| - \|A^T A\|.$$

Доказательство утверждения 1 служит свидетельством устойчивости модели к малым возмущениям.

В разделе 2.3 рассматриваются варианты решения задачи трехмерной реконструкции, в зависимости от выбранного критерия.

Таблица 1 – Виды постановок и решения задач трехмерной реконструкции

Тип задачи	Решение
1. Минимизация среднеквадратичного отклонения восстановленной точки N от решения уравнений системы (1)	$X = (A^T A)^{-1} A^T B$
2. Минимизация самого большого отклонения точки N	$\begin{cases} a_{x1}x + a_{y1}y + a_{z1}z + b_1 = 0 \\ a_{x2}x + a_{y2}y + a_{z2}z + b_2 = 0 \\ a_{x3}x + a_{y3}y + a_{z3}z + b_3 = 0 \\ a_{x4}x + a_{y4}y + a_{z4}z + b_4 = 0 \end{cases}$
3. Минимизация среднеквадратичного отклонения N от граней тетраэдра	$\begin{cases} a_{x1}x + a_{y1}y + a_{z1}z + b_1 = \xi \\ a_{x2}x + a_{y2}y + a_{z2}z + b_2 = \xi \\ a_{x3}x + a_{y3}y + a_{z3}z + b_3 = \xi \\ a_{x4}x + a_{y4}y + a_{z4}z + b_4 = \xi \end{cases}$
4. Минимизация максимального отклонения N от грани тетраэдра	$\delta^2 = \sum_i \frac{(a_{xi}x + a_{yi}y + a_{zi}z + b_i)^2}{a_{xi}^2 + a_{yi}^2 + a_{zi}^2} \rightarrow \min_{x,y,z}$ $\begin{cases} x \sum_i \frac{a_{xi}^2}{\Delta_i^2} + y \sum_i \frac{a_{yi}a_{xi}}{\Delta_i^2} + z \sum_i \frac{a_{xi}a_{zi}}{\Delta_i^2} + \sum_i \frac{a_{xi}b_i}{\Delta_i^2} = 0 \\ x \sum_i \frac{a_{xi}a_{yi}}{\Delta_i^2} + y \sum_i \frac{a_{yi}^2}{\Delta_i^2} + z \sum_i \frac{a_{yi}a_{zi}}{\Delta_i^2} + \sum_i \frac{a_{yi}b_i}{\Delta_i^2} = 0 \\ x \sum_i \frac{a_{xi}a_{zi}}{\Delta_i^2} + y \sum_i \frac{a_{yi}a_{zi}}{\Delta_i^2} + z \sum_i \frac{a_{zi}^2}{\Delta_i^2} + \sum_i \frac{a_{zi}b_i}{\Delta_i^2} = 0 \end{cases}$ <p>где $\Delta_i^2 = \Delta_{xi}^2 + \Delta_{yi}^2 + \Delta_{zi}^2$, $i \in \{1, 2, 3, 4\}$.</p>

5. Минимизация
среднеквадратичного отклонения
одной из координат точки N

$$\begin{cases} x \sum_i \frac{a_{xi}^2}{a_{zi}^2} + y \sum_i \frac{a_{xi}a_{yi}}{a_{zi}^2} + z \sum_i \frac{a_{xi}}{a_{zi}} + \sum_i \frac{a_{xi}b_i}{a_{zi}^2} = 0 \\ x \sum_i \frac{a_{xi}a_{yi}}{a_{zi}^2} + y \sum_i \frac{a_{yi}^2}{a_{zi}^2} + z \sum_i \frac{a_{yi}}{a_{zi}} + \sum_i \frac{a_{yi}b_i}{a_{zi}^2} = 0 \\ x \sum_i \frac{a_{xi}}{a_{zi}} + y \sum_i \frac{a_{yi}}{a_{zi}} + 4z + \sum_i \frac{b_i}{a_{zi}} = 0 \end{cases}$$

Рассмотренные варианты трехмерной реконструкции соответствуют различным критериям восстановления точки $N = (x, y, z)$, лежащей внутри тетраэдра.

В Главе 3 выполнено исследование инвариантности 2D и 3D-моментов. Все доказательства утверждений, касающиеся инвариантности, достигаются подстановкой аддитивно-коммуникативных возмущений (яркостных и геометрических) в соответствующие формулы моментов и линий положения с последующими эквивалентными преобразованиями. Представлены методы корректного сопоставления 2D и 3D-изображений лиц на основе линий положения. Решается задача наведения 3D-модели изображения на 2D-изображение лица.

В разделе 3.1 приводится доказательство инвариантности моментов.

Утверждение 2. Моменты Ну (Ну М.К., 1962) (M_1, M_2, \dots, M_7) являются инвариантами к группе аффинных преобразований идеального изображения.

Покажем, без потери общности, что моменты M_1, M_2 являются инвариантами к операции поворота изображения вокруг начала координат. Для этого будем пользоваться выражением для центральных моментов бинарного изображения $m_{pq} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^p (y_i - \bar{y})^q$, где $p + q \leq 3$, (x_i, y_i) – искомая точка изображения, (\bar{x}, \bar{y}) – центр изображения, N – число точек изображения.

1. Проверка инвариантности момента M_1 :

$$M_1 = m_{20} + m_{02} = \frac{1}{N} \sum_{i=1}^N \left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right]$$

Подставив преобразование поворота точек изображения вокруг начала координат на угол α получим:

$$M_1 = \frac{1}{N} \sum_{i=1}^N \left[\left((x_i \ y_i) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 + \left((x_i \ y_i) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 \right].$$

Для доказательства инвариантности достаточно воспользоваться нижеприведенным тождеством:

$$\left((a \ b) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 + \left((a \ b) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 = a^2 + b^2.$$

2. Для проверки инварианта M_2 :

$$M_2 = (m_{20} - m_{02})^2 + 4m_{11}^2 = \left(\frac{1}{N} \sum_{i=1}^N [(x_i - \bar{x})^2 - (y_i - \bar{y})^2] \right)^2 + 4 \left(\frac{1}{N} \sum_{i=1}^N [(x_i - \bar{x})(y_i - \bar{y})] \right)^2$$

достаточно воспользоваться нижеприведенным тождеством:

$$\left((a \ b) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 - \left((a \ b) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 = (a^2 - b^2)(\cos^2 \alpha - \sin^2 \alpha) - 4ab \sin \alpha \cos \alpha.$$

Аналогично проводятся доказательства для других моментов, представленных в табл.2. Доказательства, в силу большего объема преобразований вынесены в Приложение.





В разделах 3.2-3.3 исследована чувствительность моментов H_u к геометрическим искажениям цифровых изображений и изменениям яркости. Положим, что $\delta x_i = \delta y_i$ – возможные изменения (флуктуации) измеряемых переменных x_i, y_i , вызванные ошибками измерений. Чувствительность моментов H_u отражена в таблице 2.

Таблица 2 –2D-инварианты H_u , чувствительность к масштабированию

Инвариантный момент	Чувствительность
$M_1 = m_{20} + m_{02}$	δ^2
$M_2 = (m_{20} - m_{02})^2 + 4m_{11}^2$	δ^4
$M_3 = (m_{30} - 3m_{12})^2 + (3m_{21} + m_{03})^2$	δ^6
$M_4 = (m_{30} + m_{12})^2 + (m_{21} + m_{03})^2$	δ^6
$M_5 = (m_{30} - 3m_{12})(m_{30} + m_{12})[(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})] +$ $+ (3m_{21} - m_{03})(m_{21} + m_{03})[3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2]$	δ^{12}
$M_6 = (m_{20} + m_{02})[(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2]$ $+ 4m_{11}(m_{30} + m_{12})(m_{21} + m_{03})$	δ^6
$M_7 = (3m_{21} - m_{03})(m_{30} + m_{12})[(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})] -$ $(m_{30} - 3m_{12})(m_{21} + m_{03})[3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2]$	δ^{12}

Полученные оценки свидетельствуют о том, что моменты имеют различную устойчивость к искажениям. Для распознавания (классификации) объектов целесообразно использовать моменты $M_1 - M_4, M_6$ как наиболее устойчивые к помехам и погрешностям измерений. Для реальных изображений моменты μ_i флуктуируют, при поворотах, т.е. являются субинвариантами (таблица 3).

Таблица 3– Проверка инвариантности моментов к поворотам изображения

	Варианты расположения 2D-изображений лиц			
				
M_1	0.33666	0.33631	0.30231	0.30245
M_2	0.03828	0.03805	0.01254	0.01261
M_3	5.68051	6.60609	21.43457	24.89683

В таблице 4 представлена чувствительность 2D-моментов к изменениям яркости в k раз для инвариантных моментов полутоновых изображений, где:

$$\mu_{lm} = \sum_x \sum_y (x - \bar{x})^l (y - \bar{y})^m \rho(x, y).$$

Таблица 4– Чувствительность к мультипликативному изменению яркости

Инвариантный момент полутонового изображения	Чувствительность
$M_1 = \mu_{20} + \mu_{02} + \mu_{00}$	k
$M_2 = \mu_{20}\mu_{02} + \mu_{20}\mu_{00} + \mu_{02}\mu_{00} - \mu_{10}^2 - \mu_{11}^2 - \mu_{01}^2$	k^2
$M_3 = \mu_{20}\mu_{02}\mu_{00} - \mu_{00}\mu_{11}^2 - \mu_{02}\mu_{10}^2 - \mu_{20}\mu_{01}^2 + 2\mu_{11}\mu_{10}\mu_{01} - \mu_{01}^2 - \mu_{10}^2$	k^3
$M_4 = \mu_{00}^2 + 6\mu_{01}^2 + 6\mu_{02}^2 + 6\mu_{03}^2 + 6\mu_{10}^2 + 15\mu_{11}^2 - 3\mu_{10}^2\mu_{12} + 6\mu_{12}^2 - 3\mu_{02}\mu_{20} + 6\mu_{20}^2 - 3\mu_{00}(\mu_{02} + \mu_{20}) - 3\mu_{03}\mu_{21} + 6\mu_{21}^2 - 3\mu_{01}(\mu_{03} + \mu_{21}) - 3\mu_{10}\mu_{30} - 3\mu_{12}\mu_{30} + \mu_{30}^2$	k^2
$M_5 = \mu_{20}^2 + \mu_{02}^2 + \mu_{00}^2 + 2\mu_{11}^2 + 2\mu_{10}^2 + 2\mu_{01}^2$	k^2
$M_6 = \mu_{20}^3 + 3\mu_{20}\mu_{11}^2 + 3\mu_{20}\mu_{10}^2 + 3\mu_{11}^2 + 3\mu_{10}^2\mu_{02} + 3\mu_{10}^2\mu_{00} + \mu_{02}^3 + 3\mu_{02}\mu_{01}^2 + 3\mu_{01}^2\mu_{00} + \mu_{00}^3 + 6\mu_{11}\mu_{10}\mu_{01}$	k^3
$M_7 = \mu_{30}^2 + \mu_{03}^2 + \mu_{00}^2 + 3\mu_{21}^2 + 3\mu_{20}^2 + 3\mu_{12}^2 + 3\mu_{10}^2 + 3\mu_{02}^2 + 3\mu_{01}^2 + 6\mu_{11}^2$	k^2
$M_8 = \mu_{30}^2 + 2\mu_{30}\mu_{12} + 2\mu_{30}\mu_{10} + 2\mu_{21}\mu_{03} + 2\mu_{20}\mu_{00} + \mu_{03}^2 + 2\mu_{03}\mu_{01} + 2\mu_{02}\mu_{00} + \mu_{00}^2 + \mu_{21}^2 + 2\mu_{21}\mu_{01} + \mu_{20}^2 + 2\mu_{20}\mu_{02} + \mu_{12}^2 + 2\mu_{12}\mu_{10} + \mu_{10}^2 + \mu_{02}^2 + \mu_{01}^2$	k^2

Моменты имеют разную чувствительность к преобразованиям яркости $\rho(x, y)$, поэтому применять их можно только при одинаковой освещенности.

Рассмотрены методы корректного сопоставления изображений лиц на основе линий положения. Линия положения $y = a_0 + a_1x$ проводится через центр тяжести изображения (\bar{x}, \bar{y}) так, что

$$S = \sum_{i=1}^n s_i^2 f_i \rightarrow \min, \quad (3)$$

где s_i – расстояние от точки (x_i, y_i) с яркостью f_i до линии положения объекта, $s_i = (a_0 + a_1x_i - y_i) / \sqrt{a_1^2 + 1}$. Пару значений (a_0, a_1) , удовлетворяющих условию, можно

получить из уравнения $pa_1^2 + qa_1 - p = 0$, где $p = \sum_{i=1}^n f_i \sum_{i=1}^n f_i x_i y_i - \sum_{i=1}^n f_i y_i \sum_{i=1}^n f_i x_i$,

$q = \sum_{i=1}^n f_i \left(\sum_{i=1}^n f_i x_i^2 - \sum_{i=1}^n f_i y_i^2 \right) + \left(\sum_{i=1}^n f_i y_i \right)^2 - \left(\sum_{i=1}^n f_i x_i \right)^2$. Решая уравнение, получаем пару

значений $(a_1^{(1)}, a_1^{(2)})$. В соответствии с формулами $a_0 = (\sum_{i=1}^n f_i y_i - a_1 \sum_{i=1}^n f_i x_i) / \sum_{i=1}^n f_i$,

$\bar{x} = \sum_{i=1}^n f_i x_i / \sum_{i=1}^n f_i$, $\bar{y} = \sum_{i=1}^n f_i y_i / \sum_{i=1}^n f_i$ получаем два значения другого коэффициента

$(a_0^{(1)}, a_0^{(2)})$. Графическим решением являются две перпендикулярные линии, проходящие через центр тяжести (\bar{x}, \bar{y}) , из которых главная соответствует минимальному значению критерия S . Алгоритм является составной частью технологии распознавания и применяется для нормализации положения объектов с целью последующего сравнения. На рисунке 1 показан пример построения линий положения.

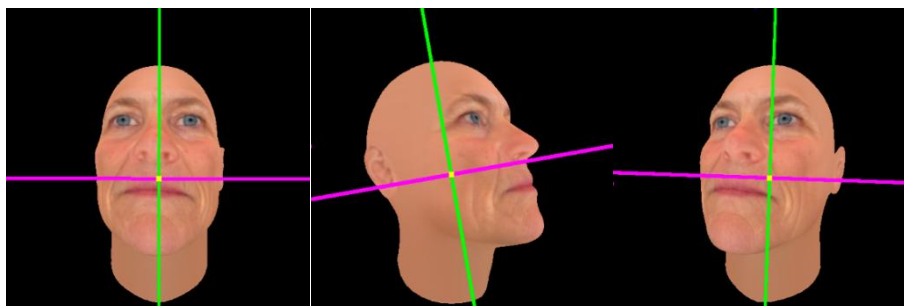


Рисунок 1 – Результат построения линий положения

Утверждение 3. Линии положения, построенные на основе критерия (3), являются инвариантами к мультипликативным яркостным преобразованиям.

Метод линий положения распространяется на трехмерный случай. Пусть трехмерное тело задано набором N точек с координатами (x_k, y_k, z_k) , $k=1, \dots, N$, которым приписаны веса f_k , соответствующие, например, яркости. Необходимо провести прямую пространственную линию так, чтобы выполнялось условие (3). Пусть линия положения проходит через «центр тяжести» объекта:

$$(\bar{x}, \bar{y}, \bar{z}) = \left(\frac{\sum_{i=1}^N f_i x_i}{\sum_{i=1}^N f_i}, \frac{\sum_{i=1}^N f_i y_i}{\sum_{i=1}^N f_i}, \frac{\sum_{i=1}^N f_i z_i}{\sum_{i=1}^N f_i} \right), \text{ где } (\bar{x}, \bar{y}, \bar{z}) \text{ – координаты центра тяжести объекта.}$$

Линия положения образует углы α, β, γ с осями координат OX, OY, OZ и ее уравнение можно записать в виде $\frac{x-\bar{x}}{l} = \frac{y-\bar{y}}{m} = \frac{z-\bar{z}}{n}$, где: l, m, n – угловые коэффициенты прямой в пространстве, которые должны быть определены в результате решения задачи; x, y, z – координаты произвольной точки объекта.

$$A = \sum_{i=1}^N (x_i - \bar{x})^2 f_i, \quad B = \sum_{i=1}^N (y_i - \bar{y})^2 f_i, \quad C = \sum_{i=1}^N (z_i - \bar{z})^2 f_i, \quad D = \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) f_i, \\ E = \sum_{i=1}^N (x_i - \bar{x})(z_i - \bar{z}) f_i, \quad F = \sum_{i=1}^N (y_i - \bar{y})(z_i - \bar{z}) f_i.$$

Величины A, \dots, F назовем моментами инерции и в дальнейшем будем использовать как постоянные коэффициенты. Для решения задачи следует найти частные производные функции S по переменным l, m, n и приравнять их нулю. После необходимых преобразований приходим к системе уравнений:

$$\begin{cases} g_1 = (l^2 + m^2 + n^2)((B+C)l - Dm - Cn) - l((B+C)l^2 + (C+A)m^2 + \\ + (A+B)n^2 - 2Dlm - 2El n - 2Fmn) = 0 \\ g_2 = (l^2 + m^2 + n^2)((C+A)m - Dl - Fn) - m((B+C)l^2 + (C+A)m^2 + \\ + (A+B)n^2 - 2Dlm - 2El n - 2Fmn) = 0 \\ g_3 = (l^2 + m^2 + n^2)((A+B)n - El - Fm) - n((B+C)l^2 + (C+A)m^2 + \\ + (A+B)n^2 - 2Dlm - 2El n - 2Fmn) = 0 \end{cases}$$

Положение тела в пространстве определяется эллипсоидом. Метод «линий положения» позволяет достаточно точно вычислять помимо линейного сдвига

еще и относительный разворот двух тел (неподвижного объекта и наблюдателя), что может быть важным для решения задачи наведения и выполнения операции сравнения с эталоном в задачах распознавания.

В разделе 3.4 решается задача формирования набора 3D-инвариантных моментов. Пусть $\rho(x, y, z)$ есть непрерывная функция, описывающая значение яркости точек с координатами (x, y, z) в трехмерном пространстве. Для корректного сопоставления изображений требуется построить моменты, инвариантные к группе аффинных преобразований. Для дискретного случая (цифрового изображения), который имеет место в нашем случае:

$$\mu_{lmn} = \sum_x \sum_y \sum_z (x - \bar{x})^l (y - \bar{y})^m (z - \bar{z})^n \rho(x, y, z),$$

где X, Y, Z – область определения координат пикселей изображения; $(\bar{x}, \bar{y}, \bar{z})$ – центр (центроид) 3D-объекта. Требуется построить моменты, инвариантные к операциям поворота, сдвига и масштабирования. Из различных источников выбраны следующие 3D-моменты:

$I_1 = \mu_{200} + \mu_{020} + \mu_{002};$ $I_2 = \mu_{200}\mu_{020} + \mu_{200}\mu_{002} + \mu_{020}\mu_{002} - \mu_{101}^2 - \mu_{110}^2 - \mu_{011}^2;$ $I_3 = \mu_{200}\mu_{020}\mu_{002} - \mu_{002}\mu_{110}^2 - \mu_{020}\mu_{101}^2 - \mu_{200}\mu_{011}^2 + 2\mu_{110}\mu_{101}\mu_{011} - \mu_{011}^2 - \mu_{101}^2.$ $F_1 = \mu_{003}^2 + 6\mu_{012}^2 + 6\mu_{021}^2 + 6\mu_{030}^2 + 6\mu_{102}^2 + 15\mu_{111}^2 - 3\mu_{102}^2\mu_{120} + 6\mu_{120}^2 - 3\mu_{021}\mu_{201} + 6\mu_{201}^2 - 3\mu_{003}(\mu_{021} + \mu_{201}) - 3\mu_{030}\mu_{210} + 6\mu_{210}^2 - 3\mu_{012}(\mu_{030} + \mu_{210}) - 3\mu_{102}\mu_{300} - 3\mu_{120}\mu_{300} + \mu_{300}^2;$ $F_2 = \mu_{200}^2 + \mu_{020}^2 + \mu_{002}^2 + 2\mu_{110}^2 + 2\mu_{101}^2 + 2\mu_{011}^2;$ $F_3 = \mu_{200}^3 + 3\mu_{200}\mu_{110}^2 + 3\mu_{200}\mu_{101}^2 + 3\mu_{110}^2 + 3\mu_{101}^2\mu_{020} + 3\mu_{101}^2\mu_{002} + \mu_{020}^3 + 3\mu_{020}\mu_{011}^2 + 3\mu_{011}^2\mu_{002} + \mu_{002}^3 + 6\mu_{110}\mu_{101}\mu_{011};$ $F_4 = \mu_{300}^2 + \mu_{030}^2 + \mu_{003}^2 + 3\mu_{210}^2 + 3\mu_{201}^2 + 3\mu_{120}^2 + 3\mu_{102}^2 + 3\mu_{021}^2 + 3\mu_{012}^2 + 6\mu_{111}^2;$ $F_5 = \mu_{300}^2 + 2\mu_{300}\mu_{120} + 2\mu_{300}\mu_{102} + 2\mu_{210}\mu_{030} + 2\mu_{201}\mu_{003} + \mu_{030}^2 + 2\mu_{030}\mu_{012} + 2\mu_{021}\mu_{003} + \mu_{003}^2 + \mu_{210}^2 + 2\mu_{210}\mu_{012} + \mu_{201}^2 + 2\mu_{201}\mu_{021} + \mu_{120}^2 + 2\mu_{120}\mu_{102} + \mu_{102}^2 + \mu_{021}^2 + \mu_{012}^2.$	(4)
---	-----

Установленный ограниченный набор 3D-инвариантов обеспечивает определенные возможности для распознавания лиц.

Утверждение 4. Моменты (4) являются 3D-инвариантами к операциям поворота и сдвига.

В таблице 4 в качестве примера приведены расчетные значения выбранной группы инвариантов для трех положений выбранной математической модели 3D-поверхности лица. Оценка теоретической чувствительности 3D-инвариантов к геометрическим искажениям представлена в таблице 5. Полученные оценки можно применить для получения инвариантов к масштабированию. Для этого достаточно положить $\delta = \sqrt{\mu_{200} + \mu_{020} + \mu_{002}}$ и выполнить нормирование путем деления моментов I_1, \dots, F_5 на соответствующие значения чувствительности (таблица 6).

Таблица 5– Пример применения 3D-инвариантов

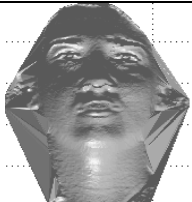
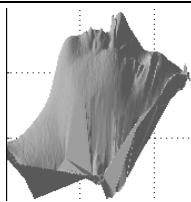
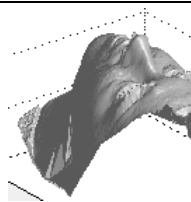
Момент	Варианты расположения 3D-изображения лица		
			
I_1	1	1	1
I_2	0.222449	0.222449	0.222449
I_3	0.0128318	0.0128318	0.0128318
F_1	255e-06	6.255e-06	6.255e-06
F_2	0.555102	0.555102	0.555102
F_3	0.371149	0.371149	0.371149
F_4	5.52475e-06	5.52475e-06	5.52475e-06
F_5	5.03792e-06	5.03792e-06	5.03792e-06

Таблица 6– Чувствительность 3D-инвариантов

Момент	I_1	I_2	I_3	F_1	F_2	F_3	F_4	F_5
Чувствительность	δ^2	δ^4	δ^6	δ^6	δ^4	δ^6	δ^6	δ^6

Утверждение 5. Нормированные моменты (4) являются 3D-инвариантами к аффинным преобразованиям.

В разделе 3.6 решается оптимизационная задача наведения 3D-изображений на 2D-изображения лиц. Пусть $f_m = f_m(x, y)$ – проекция модели трехмерного объекта на плоскость ХОУ, $f = f(x, y)$ – заданное полутонное изображение. В качестве расстояния между предъявленным изображением f и проекцией f_m

будем использовать величину $W(P) = \sqrt{\sum_{i=2}^8 (M_i^f - M_i^{f_m})^2}$, а в качестве параметров

управления углы поворота 3D-модели вокруг осей $P = (p_1, p_2, p_3)$. Задачу

оптимального наведения целесообразно решать методом градиентного спуска

$P^{[k]} = \arg \min(W^{[k]}(P))$, где k – номер итерации. Для этого полагаем известными

начальные значения углов положения $(p_1^{[0]}, p_2^{[0]}, p_3^{[0]})$. Значения параметров

приближения вычисляют по формуле $p_i^{[k+1]} = p_i^{[k]} + \lambda \Delta p_i^{[k]}$, $\Delta p_i^{[k]} = -W^{[k]}(P)$, где:

$W^{[k]}(P) = \frac{dW^{[k]}(P)}{dp_1}$, $k (k = 0, 1, \dots)$ – номер итерации; $\Delta p_i^{[k]}$ – шаг k -ой итерации; λ –

нормирующий коэффициент. В основе решения задачи лежит гипотеза о

монотонности и унимодальности функции $W(P)$ по каждому из параметров

управления (p_1, p_2, p_3) , выполнение которой упрощает реализацию метода. Вид

одной из функций, полученной экспериментально, отражен на рисунке 2 и

соответствует выдвинутой гипотезе.

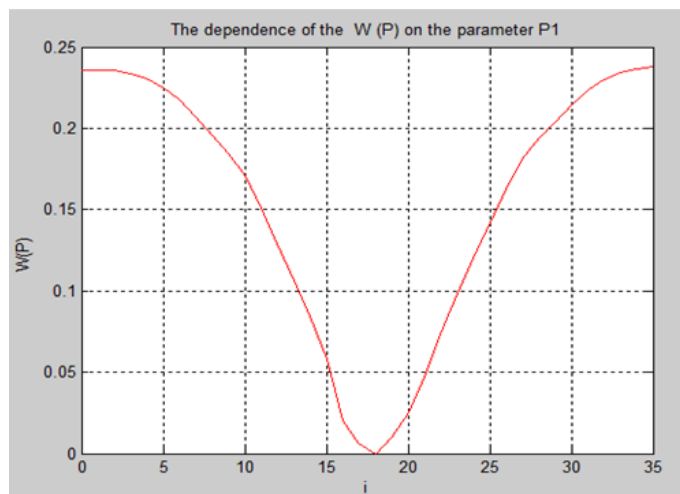


Рисунок 2 – Зависимость критерия $W(P)$ от параметра p_1

Для получения оптимума в этом случае достаточно последовательно решить задачи оптимизации по каждому из параметров. Производные вычисляются с помощью их конечно-разностных представлений $W^{[k]}(P) \approx \frac{(W^{[k+1]}(P) - W^{[k]}(P))}{(p_1^{[k+1]} - p_1^{[k]})}$. Вычисления останавливаются, когда значение функции $W(P)$ по абсолютной величине становятся меньше наперед заданного малого числа ε .

В Главе 4 описываются моделирующая среда и программные модули для проведения экспериментальных исследований с распознаванием изображений лиц, дается анализ полученных результатов.

В разделе 4.1. представлена архитектура и среда программирования системы распознавания образов.

Состав реализованных алгоритмов (модулей) представлен в таблице 7.

Таблица 7 – Состав модулей программного обеспечения

№	Наименование программного модуля
1	Модуль визуализации и управления 3D- модели лица
2	Модуль расчета 2D и 3D - инвариантов и их чувствительности
3	Модуль наведения 3D-модели на фотографию лица
4	Модули распознавания лиц на основе нейронных сетей и метрики

Для решения задач распознавания на основе 3D-моделей разработана программная среда. Среда предоставляет возможности манипулирования объектами, вычисления дескрипторов и выполнения сравнений изображений с целью распознавания. 3D-изображение представлено картой глубины, которая может быть визуализирована цветовыми каналами. В качестве классификаторов используются метрики (Евклида, Махаланобиса, Евклида-Махаланобиса) и искусственные нейронные сети.

Программная реализация построена на основе пакета открытого кода Irrlicht 1.8.3, который необходим для сборки исходных кодов. Для выполнения программы требуется видеокарта с поддержкой DirectX 9.0. Данные, полученные

в результате построения 3D-модели, а именно – RGB - изображение и карта глубины, готовы для дальнейшей обработки.

В разделе 4.2. приведены результаты экспериментов. Для первой серии экспериментов с применением метода инвариантных моментов были использованы фотографии из открытой базы данных 30 человек на белом фоне с разрешением 380x420 пикселей. При создании классов на каждого человека приходилось по три фотографии с различными эмоциями и поворотами головы. В таблице 8 приведены примеры фотографий нескольких классов и фотографий, которые подаются для распознавания.

Время обучения зависит от объема выборки и применяемого классификатора.

В таблице 9 представлены расчетные характеристики моментов.

Таблица 8–Результаты экспериментов по распознаванию лиц











Примеры фотографий трех классов из обучающих выборок				Примеры образцов для распознавания	
Класс 1				Пример 1	
Класс 2				Пример 2	
Класс 3				Пример 3	

Таблица 9– Математические ожидания и дисперсии моментов

Класс	M_1	D_1	M_2	D_2	M_3	D_3	M_4	D_4	M_5	D_5	M_6	D_6	M_7	D_7
1	9.7	0.1	0.4	0.1	0.2	0.1	0.2	0.1	7.9	0.0	3.5	0.0	1.1	0.0
2	5.8	0.5	4.6	16.8	0.5	0.1	1.0	0.4	7.8	0.0	4.1	0.3	1.1	0.0
3	6.7	0.0	4.3	0.8	0.7	0.0	0.4	0.0	7.8	0.0	3.7	0.0	1.1	0.0
4	5.3	0.3	5.6	13.8	1.0	0.0	2.9	0.2	7.7	0.0	4.6	0.4	1.2	0.1
5	7.4	0.2	6.1	3.3	0.8	0.0	0.9	0.2	7.8	0.0	4.0	0.1	1.1	0.0

В таблице 10 показаны результаты экспериментального исследования качества распознавания лиц различными методами.

Таблица 10 – Сводная таблица результатов экспериментов

Количество допущенных ошибок распознавания			
Метрика Евклида	Метрика Евклида-Махаланобиса	ИНС прямого распространения	Вероятностная ИНС
6	2	4	2

Из результатов эксперимента следует, что приоритет по качеству классификации принадлежит метрике Евклида-Махаланобиса и вероятностной нейронной сети. При использовании этой метрики система справляется с поворотами и/или наклонами головы, а также с яркостными изменениями изображения. Вероятностная нейронная сеть лучше справляется с такими «сложными» для системы факторами как закрытые глаза, измененная мимика лица (улыбка, гримаса и т.п.). Недостатком нейронных сетей являются существенные временные затраты на обучение при больших объемах данных.

Во второй серии экспериментов были использованы фотографии 168 человек с разрешением 180x200 пикселей из другой открытой базы данных. При этом 150 человек были представлены 3000 вариантами лиц с незначительной мимикой (группа 1), и 18 человек – с 360 изображениями с сильными эмоциями (группа 2). Результаты распознавания всех изображений в двух группах на основе инвариантов с помощью метрики Евклида-Махаланобиса, выполненного на персональном компьютере Dell intel core i5, отражены в таблице 11.

Таблица 11 – Сводная таблица результатов экспериментов

Группа	Время обучения	Время распознавания	Точность распознавания
1	21,5 мин.	1 мин.	93,33%
2	3 мин.	10 сек.	89 %

Из эксперимента можно сделать вывод, что метод с инвариантными моментами и метрикой Евклида-Махаланобиса работает хорошо, когда в изображениях лиц изменяются эмоции.

В заключении сформулированы основные результаты и выводы по диссертационной работе:

1. Получены оценки устойчивости трехмерной реконструкции изображений на основе стереопары к флуктуациям при условии хорошей начальной обусловленности фундаментальной матрицы, свидетельствующие о необходимости повышения разрешающей способности камер. Даны постановки различных оптимизационных задач трехмерной реконструкции, связанных с выбранным критерием качества.

2. Доказаны утверждения о том, что 2D- и 3D-моменты, отобранные для решения задачи классификации, являются инвариантами к группе аффинных преобразований в случае применения идеальных (математических) моделей изображений лиц.

3. Получены оценки устойчивости 2D- и 3D-моментов к флуктуациям в виде геометрических искажений исходных изображений (снимков), которые свидетельствуют о различной чувствительности моментов и служат для их аргументированного выбора в различных задачах.

4. Показано, что совместное применение инвариантных моментов и классификаторов в виде метрики Евклида-Махаланобиса и нейронных сетей позволяет решать задачи идентификации реальных изображений лиц, содержащих гримасы и помехи, с ошибкой от 6.67% до 11% в зависимости от условий эксперимента.

5. Разработаны и исследованы алгоритмы построения 2D- и 3D-линий положения объектов как инвариантов к изменениям освещенности, что позволяет осуществлять корректное совмещение изображений, служащее первым шагом процедуры распознавания лиц.

6. Предложен алгоритм высокоточного (оптимизационного) наведения проекции 3D-модели объекта на 2D-изображение, что позволяет решить задачу распознавания изображений лиц в различных ракурсах.

7. Разработано программное обеспечение для практического решения задач идентификации реальных изображений лиц, содержащих гримасы и помехи, с применением инвариантов, метрик и искусственных нейронных сетей.

8. Разработанные алгоритмы целесообразно использовать в составе различных поисковых систем на заключительных этапах процесса распознавания после предварительного выделения лиц из фона.

В **Приложении** приведены результаты математических преобразований, связанных с доказательством инвариантности моментов; программное обеспечение, разработанное и использованное в диссертации; скан-копия свидетельства о регистрации программы для ЭВМ.

Основные научные результаты диссертации опубликованы в следующих работах:

работы из списка, рекомендованного ВАК РФ:

1. Хачумов М.В., Нгуен Зуй Тхань. Распознавание лиц по фотографиям на основе инвариантных моментов. – Современные проблемы науки и образования, № 2-2, 2015. URL: <http://www.science-education.ru/129-23235>.

2. Нгуен Зуй Тхань. Инварианты в задачах распознавания графических образов. – Вестник Российского университета дружбы народов. Серия: «Математика, физика и информатика», № 1, 2016, с 76-85.

3. Нгуен Зуй Тхань, Хачумов В.М. Модели и методы сопоставления изображений в задаче распознавания лиц. – Искусственный интеллект и принятие решений, 2016, №4, с.5-14.

4. Нгуен Зуй Тхань, Хачумов М.В. Метод наведения 3D-модели объекта на 2D-изображение на основе инвариантных моментов. – Программные системы: теория и приложения, № 4, 2017, с.209-220. DOI: <https://doi.org/10.25209/2079-3316-2017-8-4-209-220>.

другие работы:

5. Nguen Duy Thanh, Khachumov V.M. Models and methods for matching images in the problem of face recognition //Scientific and Technical Information Processing, 2018, Vol 45, №5, pp 1-8.

6. Нгуен Зуй Тхань. Обзор и анализ методов распознавания лиц. – Тезисы докладов Всероссийской конференции (с международным участием) «Информационно-телекоммуникационные технологии математическое моделирование высокотехнологичных систем» (ИТТММ-2015, РУДН). – М.:РУДН, 2015, с. 167-169.

7. Нгуен Зуй Тхань. Анализ инвариантных моментов в задачах масштабирования и вращения изображений. – Тезисы докладов Всероссийской конференции (с международным участием) «Информационно-телекоммуникационные технологии математическое моделирование высокотехнологичных систем» (ИТТММ-2016, РУДН). – М.:РУДН, 2016, с. 156-159.

8. Нгуен Зуй Тхань. О задаче корректного сопоставления 3D-модели объекта с 2D-изображением. – Труды IV Всероссийской научной конференции молодых ученых с международным участием «Информатика, управление и системный анализ» (ИУСА-2016) (Тверской государственной технической университет, 8-11 июня 2016 г.). – Тверь: ТвГТУ, 2016, с.51-54.

9. Нгуен Зуй Тхань, Фомин М.Б., Хачумов В.М. Оценка точности восстановления координат при моделировании трехмерных объектов с использованием стереоизображений.–Материалы девятнадцатой международной научной конференции «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь (DCCN-2016)». Том 2 Математическое моделирование и задачи управления (21-25 ноября 2016 г., Москва), 2016, с.233-239.

10. Нгуен Зуй Тхань. 3D –Инвариантные моменты в задаче распознавания лиц. – Концепции фундаментальных и прикладных научных исследований: сборник статей Международной научно-практической конференции (20 февраля 2017 г., г. Казань). В 4 ч. Ч.2.– Уфа: АЭТЕРНА, 2017, с.39-44.

11. Морозова Е.Ф., Нгуен Зуй Тхань. Технология распознавания лиц на основе инвариантов к аффинным преобразованиям. – В сб. научных трудов по итогам V Международной научно-практической конференции «Актуальные проблемы технических наук в России и за рубежом», (г. Новосибирск, февраль 2018). – Новосибирск: Инновационный центр развития образования и науки, 2018. – 8 с.

12. З.Т Нгуен, Е.Ф Морозова. Исследование точности реконструкции и определение ориентации 3D- моделей лиц.- – В сб. научных трудов по итогам XIX Международный научно-практический форум студентов, аспирантов и

молодых ученых «Актуальные проблемы информатики и информационных технологий в образовании», Красноярск 22 Мая 2018 г. С.10-14.

13. Нгуен З. Т., Морозова Е. Ф., Анохина О. А., Метод оптимального сопоставления проекции 3Dобъекта с 2Дизображением. – Вестник науки и образования №6(42) 2018. Том 1 – с.18-21.

14. Нгуен Зуй Тхань, Морозова Е.Ф. Построение и управление 3D-моделями лиц человека. – Шестнадцатая Национальная конференция по искусственному интеллекту с международным участием КИИ-2018 (24–27 сентября 2018 г., г. Москва, Россия). Труды конференции. В 2-х томах. Т 2. – М.: Изд. дом Высшей школы экономики, 2018,С. 35-43.

15. Нгуен Зуй Тхань, Фраленко В.П., Хачумов М.В. Распознавание лиц на основе инвариантов к аффинным и яркостным преобразованиям. Свидетельство о государственной регистрации программы для ЭВМ №2017617799, дата регистрации: 12.07.2017.

Личный вклад автора в работах, опубликованных в соавторстве, заключается в следующем: в публикациях [1,14,15] соискателем выполнены экспериментальные исследования по распознаванию лиц; в [3,4,5,11,12,13] решены задачи выбора 3D-моментов и выполнены доказательства их инвариантности, реализованы методы наведения и сопоставления изображений; в [9] даны постановки оптимизационных задач трехмерной реконструкции. Основные теоретические и практические результаты, представленные в диссертации, получены лично автором.

Нгуен Зуй Тхань (Вьетнам)
Методы распознавания изображений лиц
на основе инвариантов к аффинным и яркостным преобразованиям
Аннотация

Получены оценки устойчивости к флуктуациям и даны постановки различных оптимизационных задач трехмерной реконструкции изображений лиц на основе стереопары при условии хорошей начальной обусловленности фундаментальной матрицы. Доказаны утверждения о том, что 2D- и 3D- моменты, отобранные для решения задач классификации лиц, являются инвариантами к группе аффинных преобразований для математических моделей изображений. Получены оценки устойчивости 2D и 3D-моментов к флуктуациям в виде геометрических искажений исходных изображений и изменений яркостных характеристик. Разработаны и исследованы методы и алгоритмы сопоставления 2D- и 3D- изображений лиц на основе линий положения и высокоточного наведения 3D- модели изображения на 2D- изображение, что предшествует и способствует решению задачи классификации. Показано, что предложенные комплексные методы, опирающиеся на совместное применение инвариантных моментов и классификаторов в виде метрик Махаланобиса или нейронных сетей, позволяют решать практические задачи идентификации реальных изображений лиц, содержащих гримасы и помехи, с достаточно высокой точностью.

Nguyen Duy Thanh (Vietnam)
Methods for recognizing face images
based on invariants to affine and luminance transformations
Annotation

The estimations of resistance to fluctuations were obtained and the statements of various optimization problems of three-dimensional reconstruction of face images based on a stereo pair under the condition of a good initial conditionality of the fundamental matrix were given. The statements that the 2D and 3D moments selected for solving the problems of face classification are invariants to the group of affine transformations for mathematical models of images were proved. The estimates of the stability of 2D and 3D moments to fluctuations in the form of geometric distortions of the original images and changes in the brightness characteristics were obtained. The methods and algorithms for comparing 2D- and 3D-images of persons based on lines position and highly accurate guidance of a 3D model of an image to a 2D image, which precede and contribute to solving the classification problem, have been developed and investigated. It was shown that the proposed complex methods, based on the combined use of invariant moments and classifiers in the form of Mahalanobis metrics or neural networks, allow us to solve practical problems of identifying real face images containing grimaces and interference with a sufficiently high accuracy.