

На правах рукописи



СЁМОВ Алексей Александрович

**МЕТОДЫ АНАЛИЗА И РАСПОЗНАВАНИЯ
ТРЕХМЕРНЫХ ИЗОБРАЖЕНИЙ
НА ОСНОВЕ СТОХАСТИЧЕСКОЙ ГЕОМЕТРИИ**

Специальность 05.13.17 – Теоретические основы информатики

А в т о р е ф е р а т
диссертации на соискание ученой степени
кандидата технических наук

ПЕНЗА 2015

Работа выполнена в ФГБОУ ВПО «Пензенский государственный университет» на кафедре «Экономическая кибернетика».

Научный руководитель: доктор технических наук, профессор
Федотов Николай Гаврилович

Официальные оппоненты: **Садыков Султан Сидыкович**,
доктор технических наук, профессор,
Муромский институт (филиал) ФГБОУ
ВПО «Владимирский государственный
университет им. Александра Григорьевича
и Николая Григорьевича Столетовых»,
профессор кафедры «Информационные
системы»;

Бурмистров Александр Владимирович,
кандидат технических наук, ПФ ФГУП
«Научно-технический центр «Атлас»,
ведущий инженер

Ведущая организация – АО «Пензенский научно-исследовательский
электротехнический институт» (г. Пенза)

Защита диссертации состоится 24 декабря 2015 г. в ____ часов, на заседании диссертационного совета Д 212.186.01 при ФГБОУ ВПО «Пензенский государственный университет» по адресу: 440026, г. Пенза, ул. Красная, 40.

С диссертацией и авторефератом можно ознакомиться в библиотеке ФГБОУ ВПО «Пензенский государственный университет». Диссертация размещена на сайте <http://dissov.pnzgu.ru/ecspertiza/syemov>

Автореферат разослан «____» _____ 2015 г.

Ученый секретарь
диссертационного совета



Гурин Евгений Иванович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. Одной из центральных проблем современной информатики является анализ и распознавание трехмерных (далее 3D) изображений. По сравнению с двумерными (далее 2D) изображениями они точнее описывают форму и свойства объектов, полнее отражают информацию об изменениях объектов с течением времени.

Все подходы к анализу и распознаванию 3D изображений можно разделить на две большие условные группы: методы, которые требуют предварительной нормализации положения 3D объекта, и методы, которые дают инвариантное описание 3D объекта вне зависимости от его пространственной ориентации, положения и масштаба. Вторая группа методов является предпочтительней, так как позволяет получать описание объекта более устойчивое к помехам и шумам. В большинстве публикаций по данной группе методов рассматриваются теоретические подходы, обладающие теми или иными ограничениями, которые не позволяют достичь инвариантности распознавания к группе движений и масштабированию 3D изображений с одновременной возможностью извлечения их различных характеристик, параметров движения и изменений масштаба.

В ряде технических задач вопросы извлечения параметров движений и масштабирования 3D изображений являются ключевыми, например, в области технической и медицинской диагностики, при создании систем организации видеонаблюдения и проектировании зрительных систем робототехники. Например, без определения параметров ориентации и движения объекта в пространстве невозможно реализовать функции перемещения робота в пространстве и позиционирование инструмента.

Важные результаты в области распознавания образов и анализа изображений получены отечественными научными школами Ю. И. Журавлева, К. В. Рудакова, В. А. Сойфера, Н. Г. Федотова, В. В. Сергеева, С. С. Садыкова; зарубежными научными школами R. Szeliski (Microsoft Research's Interactive Visual Media Group), I. H. Witten (University of Waikato New Zealand).

Целью диссертационной работы является разработка новых методов анализа и распознавания 3D изображений на основе стохастической геометрии. Для достижения поставленной цели необходимо решить следующие основные задачи:

1. Провести анализ современного состояния работ в области анализа и распознавания 3D изображений. Выявить сильные и слабые стороны методов, применяемых в данной области.

2. Разработать метод сканирования 3D изображений для достижения инвариантности их распознавания к группе движений и масштабированию.

3. Разработать математический инструмент для анализа 3D изображений с возможностью извлечения параметров их пространственной ориентации, положения и масштаба.

4. Построить новый тип признаков 3D изображений, который давал бы возможность описывать различные метрические характеристики пространственных объектов.

5. Провести эксперименты по оценке эффективности предлагаемого метода и сравнить полученные результаты с основными методами распознавания 3D изображений.

Объектом исследования являются 3D изображения.

Предметом исследования являются методы анализа и распознавания 3D изображений на основе положений стохастической геометрии.

Методы исследования основаны на теоретических положениях стохастической геометрии, функционального анализа, теории вероятностей, математической статистики и других областей.

Соответствие паспорту специальности. Диссертация выполнена в соответствии с требованием специальности 05.13.17 – Теоретические основы информатики. Области исследования: п. 5 – «Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях, разработка и исследование методов и алгоритмов анализа текста, устной речи и изображений», п. 7 – «Разработка методов распознавания образов, фильтрации, распознавания и синтеза изображений, решающих правил. Моделирование формирования эмпирического знания».

Научная новизна работы:

1. Впервые для анализа и распознавания 3D изображений предложен подход на основе стохастической геометрии, который благодаря построению признаков, инвариантных к группе движений и масштабированию, позволяет повысить надежность и универсальность распознавания. Сканирование со случайными параметрами улучшает соотношение «надежность – быстроедействие» распознавания 3D изображений по сравнению с детерминированным сканированием, что было не только обосновано теоретически, но и показано практически.

2. Разработан метод сканирования трехмерных изображений плоскостями. Данный метод в отличие от существующих позволяет анализировать 3D изображения без предварительного их упрощения или построения проекций на плоскости, анализируя непосредственно их трехмерную форму.

3. Создан математический инструмент для анализа 3D изображений – гипертрейс-матрица, позволяющий распознавать пространственные 3D объекты сложной формы и структуры благодаря построению единой математической модели. В отличие от математического аппарата других методов данный инструмент позволяет параллельно с распознаванием объекта извлекать параметры его пространственной ориентации, положения и масштаба, не требуя для этого дополнительного сканирования.

4. Построен новый тип признаков 3D изображений, имеющих аналитическую структуру, – гипертриплетные признаки. Благодаря их аналитической структуре возможна автоматическая генерация большого количества признаков с заранее заданными свойствами, в частности, инвариант-

ности и чувствительности по отношению к группе движений и масштабным преобразованиям.

Практическая значимость работы. Результаты исследования могут быть использованы при разработке интегрированных систем безопасности для сканирования и обнаружения пространственных объектов в системах видеонаблюдения, а также в сигнализационных комплексах и системах при распознавании 3D изображений в технических средствах обнаружения. Инвариантное описание, получаемое разработанным методом, позволяет увеличить скорость обработки результатов обнаружения искомого объекта и тем самым сократить время на включения сигнализации и системы оповещения. Данные положения подтверждаются актом о внедрении.

На основе результатов исследования разработаны программные комплексы по сканированию и распознаванию 3D изображений (свидетельства об официальной регистрации программ для ЭВМ в Роспатенте № 2015612257 от 16.02.15 и № 2015612814 от 26.02.15), которые позволяют повысить качество распознавания пространственных объектов в системах машинного зрения.

Основные положения, выносимые на защиту:

1. Метод сканирования 3D изображений плоскостями, обеспечивающий возможность достижения инвариантного распознавания объекта при его произвольной ориентации в пространстве.

2. Математический инструмент для анализа 3D изображений – гипертрейс-матрица, позволяющая извлекать параметры масштаба, положения и ориентации объекта в пространстве.

3. Процедура построения признаков нового типа, имеющих аналитическую структуру, – гипертриплетных признаков, которые способны описывать форму и структуру объекта и вычислять его метрические характеристики.

4. Процедура сокращения размерности признакового пространства для определения набора информативных признаков 3D изображений с указанием их различающей силы.

Реализация и внедрение результатов работы. Работа выполнялась по гранту РФФИ (проект № 12-07-00501). Результаты исследований используются на предприятии «НИКИРЭТ» (филиал ФГУП ФНЦП «ПО "СТАРТ" им. М. В. Проценко», г. Заречный), что подтверждается актом о внедрении.

Достоверность и обоснованность научных и практических результатов диссертации обеспечивается корректным применением математического аппарата; апробацией на научно-технических конференциях; сходимостью результатов и выводов теоретических и компьютерных исследований; проведением экспериментов по классификации базы данных 3D изображений The Princeton Shape Benchmark; практическим применением полученных результатов, подтвержденных актом о внедрении.

Апробация работы. Основные результаты и положения диссертации докладывались и обсуждались на конференциях: «Intelligent Information Processing» (2014), «Pattern Recognition and Image Analysis: New Information technologies» (2013), «Надежность и качество» (2013–2015), «Проблемы информатики в образовании, управлении, экономике и технике» (2013–2014), «Новые информационные технологии и системы» (2012, 2014), «Модели, системы, сети в экономике, технике, природе и обществе» (2014), «Современные методы и средства обработки пространственно-временных сигналов» (2012–2015).

Публикации. По теме диссертационной работы опубликованы 23 печатные работы. Из них 7 статей в журналах, входящих в перечень изданий, рекомендованных ВАК РФ, а также 2 свидетельства о государственной регистрации программ для ЭВМ.

Личный вклад автора. Основные научные результаты, приведенные в диссертации и выносимые на защиту, получены автором лично. Автором был выполнен следующий объем исследований: получение теоретических результатов (метод сканирования 3D изображений, математический инструмент для анализа 3D изображений, процедура сокращения признакового пространства), разработка и тестирование программных комплексов для распознавания 3D изображений, проведение эксперимента по классификации базы изображений The Princeton Shape Benchmark.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка литературы из 88 наименований и двух приложений. Общий объем работы – 140 страниц, в том числе: 121 страница основного текста (включая 10 таблиц, 23 рисунка), 10 страниц списка литературы.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность выбранной темы, сформулированы цель и задачи исследования, показана научная новизна и практическая значимость полученных результатов.

В первой главе описано современное состояние работ в области распознавания изображений. Рассмотрены основные методы анализа и распознавания 3D изображений, указаны их преимущества и недостатки. Проанализирована проблема достижения инвариантного распознавания объекта при его произвольной ориентации в пространстве. Для ее решения введено понятие «опорная сетка на сфере» и обоснована необходимость построения равномерного распределения узлов этой сетки на поверхности сферы.

Распознавание образов – это процесс отнесения объекта, представленного вектором признаков, к одному из заданных классов по определенному правилу в соответствии с поставленной целью.

Создание инвариантного описания 3D изображения осложнено наличием произвольной его пространственной ориентации. В отличие от двумерного случая, проблема поворота 3D изображения в трехмерном пространстве связана с невозможностью его поворота вокруг одной оси, не за-

трагивая вращения по другим осям. Поэтому необходимо разработать такую схему сканирования 3D изображения, чтобы ее результаты не зависели бы от пространственной ориентации объекта.

Сканирование 3D изображения будет осуществляться сетками параллельных плоскостей. Пусть F – исходное 3D изображение. Определим плоскость $B(\eta, r) = \{x \mid x^T \cdot \eta = r\}$ как касательную к сфере с центром в начале координат и с радиусом r , проходящую через заданную точку X и на расстоянии r от начала координат с заданными углами ω и φ , где $\eta = [\cos \omega \cdot \sin \varphi, \sin \omega \cdot \sin \varphi, \cos \varphi]$ – единичный вектор в R^3 , r , ω , φ – сферические координаты.

Стандартный перебор всех пар углов ω и φ , которыми определяется каждая сканирующая сетка параллельных плоскостей, в топологическом смысле для непрерывного случая дает модель концентрических сфер с центром в начале координат. Каждой сканирующей сетке плоскостей на единичной сфере сопоставим точку, которая будет являться точкой касания со сферой плоскости, параллельной плоскостям сетки. Множество точек на сфере образует опорную сетку (рисунок 1).

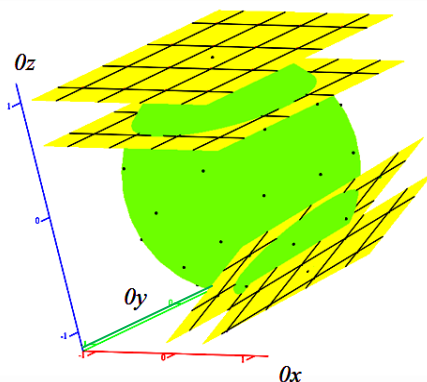


Рисунок 1 – Опорная сетка на сфере и соответствующие ей сетки сканирующих параллельных плоскостей

Стоит отметить, что пара углов (ω, φ) однозначно определяет узел опорной сетки, соответствующий единственной касательной плоскости и сетке сканирующих параллельных плоскостей. Если при повороте сферы вокруг своего центра опорная сетка перейдет сама в себя, то и все сетки сканирующих плоскостей полностью совпадут друг с другом и получаемые сечения не изменят свою форму. Поэтому вычисляемое значение признака не изменится.

Таким образом, необходимо построить опорную сетку, обладающую равномерным распределением точек на сфере для достижения меньшей ошибки совмещения узлов опорной сетки при ее повороте (из-за дискретного шага сканирования). Равномерное распределение точек опорной сетки обеспечит отсутствие плотных скоплений узлов на сфере, определяющих преимущественно сечения под теми или иными углами обзора исходного 3D изображения.

Во второй главе разработан метод сканирования 3D изображений. Описана его математическая модель. Теоретически обосновано преимущество сканирования 3D изображений со случайными параметрами. Создан и описан математический инструмент для анализа 3D изображений – гипертрейс-матрица. Описаны основные свойства матрицы и влияние движения и масштабирования изображения на ее значение. Приведены примеры анализа 3D изображений с извлечением параметров их движения и масштабирования на основе гипертрейс-матриц.

Сканирование 3D изображения производится сеткой параллельных плоскостей с расстоянием Δr между плоскостями и заданными углами ω и φ (рисунок 2,а). Взаимное положение изображения F и каждой сканирующей плоскости $B(\eta(\omega, \varphi), r)$ характеризуется числом G : $G = \text{HyperT}(F \cap B(\eta(\omega, \varphi), r))$. В качестве правила *HyperT* можно использовать, например, вычисление площади сечения.

Сканирование сеткой параллельных плоскостей повторяется для каждого нового значения угла обзора, определяемого выражениями $\omega + \Delta\omega$ и $\varphi + \Delta\varphi$, с тем же шагом Δr между сканирующими плоскостями. Углы ω и φ меняются согласно узлам опорной сетки.

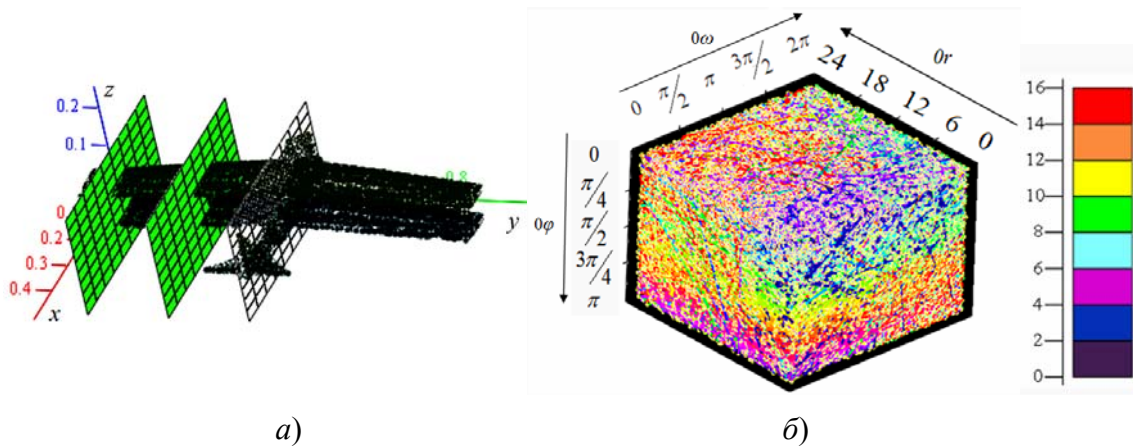


Рисунок 2 – Сканирования 3D объекта сеткой плоскостей слева (а) и его гипертрейс-матрица 3ТМ справа (б)

Результат вычислений *HyperT* функционала зависит от трех параметров плоскости (r , ω , φ). Поэтому результат сканирования удобно представить в виде трехмерной матрицы, у которой ось 0φ направлена вертикально, ось 0ω – горизонтально, ось $0r$ – вглубь (рисунок 2,б). На рисунке 2,б элемент матрицы показывает значение периметра сечения.

Таким образом, тройке $(\omega_i, \varphi_j, r_k)$ соответствует элемент матрицы 3ТМ с номером (i, j, k) и значением $\Pi(F_{\text{sect}})$, которое характеризует информативный признак сечения: $F_{\text{sect}} = F \cap B(\eta(\omega_i, \varphi_j), r_k)$.

После заполнения гипертрейс-матрицы последовательно обрабатываются ее глубинные, вертикальные и горизонтальные строки с помощью функционалов *HyperP*, *Hyper Ω* и *Hyper Θ* , соответственно. Например,

$HyperP = \int G(\omega, \varphi, r) dr$, $Hyper\Omega = \max_{\varphi} G(\omega, \varphi)$ и $Hyper\Theta = \min_{\omega} G(\omega)$. В результате получается число $Res(F)$ – гипертриpletный признак 3D изображения F в виде композиции четырех функционалов:

$$Res(F) = Hyper\Theta \circ Hyper\Omega \circ HyperP \circ HyperT(F_{sect}).$$

Для нахождения признака сечения используется 2D трейс-преобразование, введенное и описанное профессором Н. Г. Федотовым. Так, сканирование получаемых сечений F_{sect} осуществляется решеткой параллельных прямых $l(\theta, \rho)$ с расстоянием $\Delta\rho$ между линиями, где ρ, θ – полярные координаты прямой в плоскости сечения (рисунок 3,а). Взаимное положение 2D изображения F_{sect} и каждой сканирующей линии $l(\theta, \rho)$ характеризуется числом: $g = T(F_{sect} \cap l(\theta, \rho))$. В качестве правила T можно использовать, например, количество пересечений прямой с изображением.

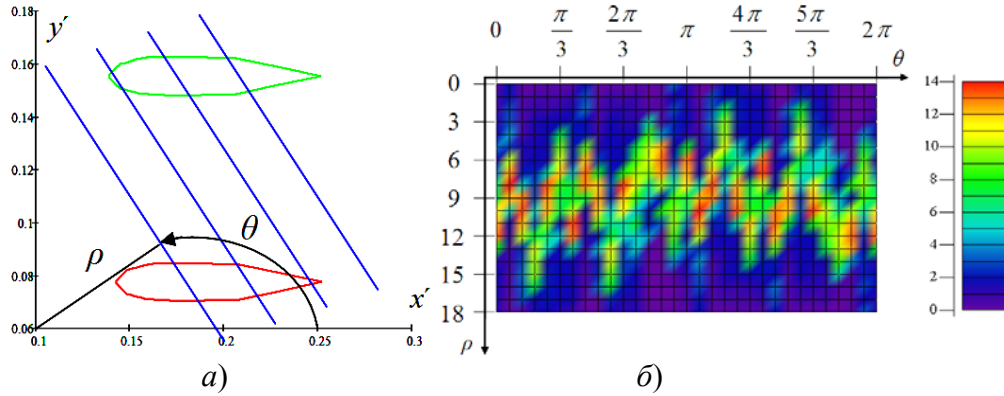


Рисунок 3 – Процесс сканирования 2D сечения сеткой параллельных прямых слева (а) и его трейс-матрица ТМ справа (б)

Сканирование сеткой параллельных прямых повторяется для каждого нового угла $\theta + \Delta\theta$ в той же плоскости сечения F_{sect} и с шагом $\Delta\rho$ до завершения оборота в 2π радиан. Результат вычислений трейс функционала T зависит от двух параметров прямой θ и ρ . Поэтому результат сканирования удобно представить в виде двумерной трейс-матрицы ТМ, у которой ось $\theta\theta$ направлена горизонтально, а ось $\theta\rho$ – вертикально (рисунок 3,б). Таким образом, паре (θ_i, ρ_j) соответствует элемент матрицы ТМ с номером (i, j) и значением $T(F_{sect} \cap l(\theta_i, \rho_j))$.

После заполнения матрицы ТМ последовательно обрабатываются ее вертикальные и горизонтальные строки с помощью функционалов P и Θ , соответственно. Например, $T = \max_{\rho} g(\theta, \rho)$ и $\Theta = \min_{\theta} g(\theta)$. В результате получается число $\Pi(F_{sect})$ – триpletный признак 2D изображения сечения F_{sect} в виде композиции трех функционалов:

$$HyperT(F_{sect}) = \Pi(F_{sect}) = \Theta \circ P \circ T(F_{sect} \cap l(\theta, \rho)).$$

Объединяя полученные формулы для $Res(F)$ и $\Pi(F_{\text{sect}})$, получаем окончательно аналитическую структуру признака 3D изображения:

$$Res(F) = Hyper\Theta \circ Hyper\Omega \circ HyperP \circ HyperT(\Theta \circ P \circ T(F_{\text{sect}} \cap l(\theta, \rho))).$$

Таким образом, благодаря композиционной структуре функционалов, входящих в структуру $\Pi(F_{\text{sect}})$ и $Res(F)$, возможно получение большого числа признаков, свойства которых можно регулировать.

Определить по гипертрейс-матрице углы поворота, совмещающие один 3D объект F' с другим F , можно следующим образом. Если для каждого объекта определить уникальную плоскость, которая жестко фиксирована по отношению к нему, то угол между этими плоскостями даст информацию об угле поворота, при котором объект F перейдет в объект F' . Выбор уникальной плоскости эквивалентен определению номера (i, j, k) элемента гипертрейс-матрицы со значением $HyperT(F \cap B(\eta(\omega_i, \varphi_j), r_k))$, которому соответствует тройка $(\omega_i, \varphi_j, r_k)$. Таким образом, зная углы поворота ω и φ , можно легко восстановить трехмерный угол поворота, переводящий объект F' в объект F и совмещающий их уникальные сканирующие плоскости.

Коэффициент масштабирования может быть получен из результатов сканирования. Для каждого сечения определяется F_{sect} n -мерный вектор S_n признаков сечения $\Pi(F_{\text{sect}})$. Для снижения уровня искажений для каждого среза сечений вычисляется k -й признак: $QS_{j,k} = h(S_k)_{i,j}$, где правило h может быть определено, например, как $h(x) = \max(x)$. Далее для k -го признака рассчитывается усредненная характеристика всех срезов: $P_k = \sum_j QS_{j,k} / m$, где

m – число узлов опорной сетки на сфере. Коэффициент масштабирования равен $\mu_k = \sqrt{\dim(P_k)} \sqrt{P_k / E_k}$, где $\dim(P_k)$ – размерность k -го признака (например, признак площади сечения имеет $\dim = 2$), E_k – значение k -го признака для единичного масштаба 3D изображения. Для получения более устойчивого значения найденные коэффициенты μ_k нужно усреднить.

В третьей главе исследованы свойства и математическая модель признаков нового типа – гипертриплетных признаков. Рассмотрены свойства и примеры функционалов, входящих в аналитическую структуру признака. Исследованы условия инвариантности конструируемых признаков к группе движений и масштабированию 3D изображения. Приведены примеры геометрического класса признаков, описывающих метрические характеристики объекта. Разработана процедура сокращения размерности признакового пространства.

Функционалы, входящие в аналитическую структуру признака, могут обладать разными свойствами. Варьируя различные их виды, можно получать большое количество признаков с заранее заданными свойствами исходного 3D изображения.

Ниже приведены признаки, которые являются инвариантными к переносу, повороту и масштабированию 3D изображения. Все они имеют структуру $Res(F) = Hyper\Theta \circ Hyper\Omega \circ HyperP \circ HyperT(\Theta \circ P \circ T)$:

$$1. T(F_{sect} \cap l(\theta, \rho)) = \max_t f(\theta, \rho, t), \quad P = \sum_{\rho} g(\theta, \rho), \quad \Theta = \max_{\theta} g(\theta) - \min_{\theta} g(\theta),$$

$$HyperT(F \cap B(\eta(\omega, \varphi), r)) = \Pi(F_{sect}) = G(\omega, \varphi, r), \quad HyperP = \max_r G(\omega, \varphi, r),$$

$$Hyper\Omega = \underset{\varphi}{\text{LocalMax}} G(\omega, \varphi), \quad Hyper\Theta = \underset{\omega}{\text{min}} G(\omega), \quad f(\theta, \rho, t) - \text{длина } t\text{-го от-}$$

резка, высекаемого ρ -й прямой под θ -м углом в плоскости сечения F_{sect} , $G(\omega, \varphi, r)$ – признак сечения, получаемого в r -й плоскости под парой углов (ω, φ) обзора 3D объекта, функция LocalMax вычисляет количество локальных максимумов по столбцам матрицы $G(\omega, \varphi)$.

$$2. T = \text{rows}_t f(\theta, \rho, t), \quad P = \max_i g(\theta, \rho_i), \quad \Theta = \max_{\theta} g(\theta), \quad HyperT = G(\omega, \varphi, r),$$

$$\arg \max_r G(\omega, \varphi, r) - \arg \min_r G(\omega, \varphi, r)$$

$$HyperP = \frac{r}{\text{Row3D}}, \quad Hyper\Omega = \max_{\varphi} G(\omega, \varphi),$$

$$Hyper\Theta = \max_{\omega} G(\omega), \quad \text{где Row3D} = \text{rows}_{r, G(\omega, \varphi, r) \neq 0} G(\omega, \varphi, r) - \text{количество нену-}$$

левых элементов в глубинных строках матрицы ЗТМ, функции $\arg\max$ и $\arg\min$ вычисляют номер аргумента (по переменной r), при котором достигается максимум и минимум глубинных строк матрицы ЗТМ.

$$3. T = \sum_t f(\theta, \rho, t), \quad P = \text{mean}_{\rho} g(\theta, \rho), \quad \Theta = \min_{\theta} g(\theta), \quad HyperT = G(\omega, \varphi, r),$$

$$HyperP = \text{mean}_r G(\omega, \varphi, r), \quad Hyper\Omega = \max_{\varphi} G(\omega, \varphi),$$

$$Hyper\Theta = \frac{\max_{\omega} G(\omega) - \min_{\omega} G(\omega)}{\max_{\omega} G(\omega) + \min_{\omega} G(\omega)}, \quad \text{где mean}(x) = \sum_{i=1}^n x_i / n.$$

Ниже приведены признаки с той же аналитической структурой, но которые вычисляют метрические характеристики 3D изображения:

$$1. \text{Объем 3D объекта: } T = \sum_t f(\theta, \rho, t), \quad P = \sum_{\rho} g(\theta, \rho) \cdot \Delta\rho, \quad \Theta = \max_{\theta} g(\theta),$$

$$HyperT = G(\omega, \varphi, r), \quad HyperP = \sum_r G(\omega, \varphi, r) \cdot \Delta r, \quad Hyper\Omega = \max_{\varphi} G(\omega, \varphi),$$

$$Hyper\Theta = \max_{\omega} G(\omega), \quad \text{где } \Delta r - \text{расстояние между параллельными плоско-} \\ \text{стями, } \Delta\rho - \text{расстояние между параллельными прямыми в сечении.}$$

$$2. \text{Максимальная площадь сечения объекта плоскостью:} \\ T = \sum_t f(\theta, \rho, t), \quad P = \sum_{\rho} g(\theta, \rho) \cdot \Delta\rho, \quad \Theta = \max_{\theta} g(\theta), \quad HyperT = G(\omega, \varphi, r),$$

$$HyperP = \max_r G(\omega, \varphi, r), \quad Hyper\Omega = \max_{\varphi} G(\omega, \varphi), \quad Hyper\Theta = \max_{\omega} G(\omega).$$

3. Площадь поверхности объекта: $T = \sum_t f(\theta, \rho, t)$, $\Theta = \text{mean}_{\theta} g(\theta)$,

$$P = (\text{Row2D} + 1) \cdot \Delta\rho + \sum_{i=1}^{\text{Row2D}-1} |g(\theta, \rho_{i+1}) - g(\theta, \rho_i)|, \quad \text{HyperT} = G(\omega, \varphi, r),$$

$$\text{HyperP} = \sum_r G(\omega, \varphi, r) \cdot \Delta r, \quad \text{Hyper}\Omega = \max_{\varphi} G(\omega, \varphi), \quad \text{Hyper}\Theta = \max_{\omega} G(\omega), \quad \text{где}$$

Row2D – количество ненулевых элементов в столбце трейс-матрицы ТМ.

Предлагаемый метод позволяет автоматически генерировать большое количество признаков 3D изображения. Сформированная таким образом система признаков, как правило, избыточна. Поэтому целесообразно разработать процедуру, которая минимизирует количество признаков и выделяет наиболее информативные среди них.

Обозначим через $\text{Res}_k^{A_i(s)}$ гипертриплетный признак k -го вида, вычисленный для s -го представителя i -го класса A_i . Среднее значение k -го признака для изображений класса A_i равно $\mu_k^{A_i} = \frac{2}{h_i} \sum_{s=1}^{h_i} \text{Res}_k^{A_i(s)}$, а его средне-

$$\text{квадратическое отклонение: } \sigma_k^{A_i} = \sqrt{\frac{2}{h_i} \left(\sum_{s=1}^{h_i} \left(\text{Res}_k^{A_i(s)} - \mu_k^{A_i} \right)^2 \right)}.$$

Для отбора потенциально эффективных признаков рассчитывается показатель $p(A_i, k)$, который определяет меру неподобия k -го признака для i -го

$$\text{класса } A_i: p(A_i, k) = \frac{q_1(A_i, k) + q_2(A_i, k)}{h_i}, \quad \text{где}$$

$$q_1(A_i(s), k) := q_1(A_i(s), k) + 1, \quad \text{если } \text{Res}_k^{A_j(s)} < \mu_k^{A_i} - \sigma_k^{A_i} \vee \mu_k^{A_i} + \sigma_k^{A_i} < \text{Res}_k^{A_j(s)}$$

$$\text{и } i = j, \quad q_2(A_i(s), k) := q_2(A_i(s), k) + 1, \quad \text{если } \mu_k^{A_i} - \sigma_k^{A_i} \leq \text{Res}_k^{A_j(s)} \leq \mu_k^{A_i} + \sigma_k^{A_i} \text{ и } i \neq j.$$

Коэффициент p представляет собой матрицу весов, рассчитанную для каждого типа признака k ($k = 1 \dots n$) в зависимости от класса объектов A_i ($i = 1 \dots m$). Показатель $q_1(A_i(s), k)$ показывает количество 3D изображений s i -го класса A_i , которые будут неправильно классифицированы, так как слишком далеко отстоят от среднего представителя своего класса. Показатель $q_2(A_i(s), k)$ показывает количество 3D изображений s i -го класса A_i , которые будут неправильно классифицированы, так как слишком близко находятся к среднему представителю другого класса.

Таким образом, чем выше коэффициент неподобия $p(A_i, k)$, тем меньшей различающей силой обладает k -й признак для i -го класса A_i . В связи с этим целесообразно задать некоторый порог δ , чтобы из всей совокупности признаков выделить потенциально эффективные, отсеяв заведомо неинформативные признаки: $\sum_{i=1}^m p(A_i, k) / m \leq \delta$.

Далее по полученной совокупности признаков нужно произвести расчет элементов матрицы парной корреляции средних значений признаков $\mu_k^{A_i}$ по классам. Выделяются те пары признаков, для которых значение коэффициента парной корреляции не ниже 0.7.

Для дальнейшего отбора информативных признаков производится сортировка их усредненных значений по классам: $\xi_k^{A_i}$ – среднее значение k -го признака по i -му классу A_i , отсортированное по возрастанию ($\xi_k^{A_1} \leq \xi_k^{A_2} \leq \dots \leq \xi_k^{A_m}$). Затем находится разница между значениями соседних элементов: $\Delta \xi_k^i = \xi_k^{A_i} - \xi_k^{A_{i-1}}$ ($i = 2 \dots m$).

Далее вычисляется нижняя граница интервала колебания среднего значения признаков $\Delta \xi_k^i$ ($i = 2 \dots m$): $\text{border}(k) = \text{mean}(\Delta \xi_k) - \text{stdev}(\Delta \xi_k)$, где

$$\text{mean}(\Delta \xi_k) = \frac{2}{m-1} \cdot \sum_{i=1}^{m-1} \Delta \xi_k^i, \text{ а } \text{stdev}(\Delta \xi_k) = \sqrt{\frac{2}{m-1} \left(\sum_{i=1}^{m-1} (\Delta \xi_k^i - \text{mean}(\Delta \xi_k))^2 \right)}.$$

Критерий отбора признаков следующий: чем выше значение $\text{border}(k)$, тем выше различающая сила и информативность k -го признака. Программа отбирает z лучших представителей по данному критерию, где порог z задает аналитик.

Обозначим через t тестовое 3D изображение, и пусть его k -й признак равен Res_k^t . Тогда расстояние между изображением t и классом A_i с учетом

весовых значений равно $d(t, A_i) = \sum_k p(A_i, k) \cdot \frac{|\text{Res}_k^t - \mu_k^{A_i}|}{\sigma_k^{A_i}}$. Система относит изображение t к классу A_j , если $d(t, A_j) = \min_i d(t, A_i)$.

В четвертой главе приведены результаты экспериментальной проверки свойств предлагаемого метода. Определены оптимальные параметры стохастического сканирования 3D изображения. Проведен анализ устойчивости вычисления признаков к группе движений и масштабированию. Оценена погрешность вычисления геометрических признаков. Экспериментально показано преимущество стохастического сканирования по сравнению с детерминированным вариантом, с точки зрения соотношения «надежность – быстродействие» распознавания. На примере тестовой базы изображений решена задача классификации разработанным и другими методами и полученные результаты сравнены между собой.

Для доказательства эффективности разработанного метода был осуществлен эксперимент классификации базы данных 3D изображений The Princeton Shape Benchmark. Кроме предлагаемого метода, данная база была проанализирована еще и другими известными методами. Из-за ограничения объема в настоящем автореферате представлен только один из них – *Light Field Descriptor* (далее *LFD*), имеющий наилучшие показатели по

сравнению с другими методами. Метод *LFD* представляет исходный 3D объект как совокупность 2D изображений, формируемых из равномерно выбранных углов сферы обзора, благодаря чему достигается инвариантность вычисляемых признаков к группе движений изображения.

Для оценки качества классификации использовались показатели:

1. Точность (*precision*), которая определяется как отношение найденных 3D объектов D_{true} , действительно принадлежащих к своему классу, к общему числу найденных объектов D_{find} : $precision = \frac{|D_{true}|}{|D_{find}|}$.

2. Полнота (*recall*), которая определяется как отношение найденных 3D объектов D_{true} , действительно принадлежащих к своему классу, к общему числу всех объектов этого класса D_{class} : $recall = \frac{|D_{true}|}{|D_{class}|}$.

3. *F*-мера – показатель, который дает оценку доли правильно распознанных объектов с учетом ложно пропущенных и ложно найденных объектов: $F = \frac{2}{1/recall + 1/precision} = \frac{2 \cdot recall \cdot precision}{recall + precision}$.

4. *DCG*-статистика – показатель, который придает больший вес правильно найденным объектам, расположенным в начале, чем в конце ранжированного по какому-то показателю списка R : $DCG = \frac{G_1 + \sum_{i=2}^N G_i / \log_2(i)}{1 + \sum_{i=2}^N 1 / \log_2(i)}$,

где $G_i = 1$, если R_i соответствует классу объекта, и $G_i = 0$ в противном случае.

Ниже приведены результаты эксперимента по классификации базы 3D изображений (таблица 1). Вычисление признаков разработанным методом производилось как с помощью матрицы весов, так и без нее.

Таблица 1 – Результаты эксперимента по классификации 3D объектов

Метод распознавания изображений	Результаты, %			
	<i>precision</i>	<i>recall</i>	<i>F</i> -мера	<i>DCG</i>
3D <i>TRACE</i> с матрицей весов	71.3	65.3	68.2	76.8
3D <i>TRACE</i> без матрицы весов	64.3	34.2	44.6	66.1
<i>LFD</i>	65.7	17.8	28.0	64.3

При использовании матрицы весов результаты классификации базы 3D изображений предлагаемым методом 3D *TRACE* заметно улучшаются. Так, коэффициент точности улучшил свое значение на 7 процентных пунктов (пп.), коэффициент полноты повысился на 30.9 пп., значение *F*-меры выросло на 23.6 пп., значение *DCG*-статистики повысилось на 10.7 пп.

Как видно из таблицы 1, коэффициенты точности, полноты, *F*-меры и *DCG*-статистики метода 3D *TRACE* больше аналогичных показателей метода *LFD* на 5.6 пп., 47.5 пп., 40.2 пп. и 12.5 пп., соответственно.

В заключении сформулированы основные результаты диссертационного исследования.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Предложен новый подход к анализу и распознаванию 3D изображений. Элементы теории стохастической геометрии, применявшиеся ранее только лишь для 2D изображений, впервые получили свое развитие на 3D изображения, что позволило заметно повысить точность и надежность распознавания, о чем свидетельствуют результаты экспериментальных исследований.

2. Разработан метод сканирования 3D изображений плоскостями. Данный метод позволяет достичь их полного инвариантного описания и решить на этой основе проблему распознавания 3D изображений при их произвольной ориентации в пространстве.

3. Создан и описан математический инструмент для анализа 3D изображений – гипертрейс-матрица, позволяющий извлекать параметры масштаба, положения и ориентации объекта в пространстве.

4. Разработана процедура построения признаков нового типа, имеющих аналитическую структуру, – гипертриплетных признаков, которые способны описывать форму и структуру объекта и вычислять его метрические характеристики.

5. Разработана процедура сокращения размерности признакового пространства, позволяющая получать набор информативных признаков с указанием их различающей силы для каждого класса 3D изображений, что позволяет повысить точность распознавания.

6. Проведено экспериментальное сравнение эффективности предлагаемого метода с основными методами распознавания 3D изображений при решении задачи классификации тестовой базы принстонского университета The Princeton Shape Benchmark. Значение коэффициента F -меры (среднегармоническое значение коэффициентов точности и полноты) при отнесении 907 пространственных объектов тестовой совокупности к одному из 92 классов для разработанного метода составило 68.2 %, тогда как наилучшее значение среди других рассматриваемых методов для данной базы равняется 28.0 %.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных ВАК РФ

1. Сёмов, А. А. 3D трейс-преобразование и его свойства / Н. Г. Федотов, А. А. Сёмов // XXI век: итоги прошлого и проблемы настоящего плюс : науч.-метод. журнал. – 2013. – № 10 (14). – С. 68–74. – (Технические науки. Информационные технологии).

2. Сёмов, А. А. Основные преимущества и дополнительные возможности 3D трейс-преобразования / Н. Г. Федотов, А. А. Сёмов // XXI век: итоги прошлого и проблемы настоящего плюс : науч.-метод. журнал. – 2014. – № 03 (19). – С. 77–83. – (Технические науки. Информационные технологии).

3. Сёмов, А. А. Экспериментальная проверка свойств 3D трейс-преобразования / А. А. Сёмов // XXI век: итоги прошлого и проблемы настоящего плюс : науч.-метод. журнал. – 2014. – № 03 (19). – С. 83–89. – (Технические науки. Информационные технологии).

4. Сёмов, А. А. 3D трейс-преобразование: режимы сканирования, особенности стохастической реализации, способы ускорения вычислений / Н. Г. Федотов, А. А. Сёмов, А. В. Моисеев // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2014. – № 3 (31). – С. 41–53.

5. Semov, A. A. Trace transform of three-dimensional objects: recognition, analysis and database search / N. G. Fedotov, S. V. Ryndina, A. A. Semov // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications. – 2014. – Vol. 24. – No. 4. – P. 566–574.

6. Сёмов, А. А. Гипертрейс-матрица как основной инструмент анализа 3D-объектов / Н. Г. Федотов, А. А. Сёмов // XXI век: итоги прошлого и проблемы настоящего плюс : науч.-метод. журнал. – 2015. – № 03 (25). – Т. 1. – С. 63–69. – (Технические науки. Информационные технологии).

7. Сёмов, А. А. Основные методы построения гипертрейс-матриц / А. А. Сёмов // XXI век: итоги прошлого и проблемы настоящего плюс : науч.-метод. журнал. – 2015. – № 03 (25). – Т. 1. – С. 69–76. – (Технические науки. Информационные технологии).

Публикации в других изданиях

8. Сёмов, А. А. Краткий обзор основных подходов к анализу 3D-моделей и разработка 3D трейс-преобразования / Н. Г. Федотов, А. А. Сёмов // Новые информационные технологии и системы (НИТИС-2012) : сб. тр. X Междунар. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2012. – С. 222–225.

9. Сёмов, А. А. Применение трейс-преобразования для распознавания 3D изображений / Н. Г. Федотов, А. А. Сёмов // Современные методы и средства обработки пространственно-временных сигналов : сб. ст. XI Всерос. науч.-техн. конф. – Пенза : Приволжский Дом знаний, 2013. – С. 73–75.

10. Сёмов, А. А. Идеи построения равномерной сетки на сфере и 3d трейс-преобразование / Н. Г. Федотов, А. А. Сёмов // Проблемы информатики в образовании, управлении, экономике и технике : сб. ст. XIII Междунар. науч.-техн. конф. – Пенза : Приволжский Дом знаний, 2013. – С. 23–26.

11. Syemov, A. A. Trace transform of spatial images / N. G. Fedotov, S. V. Ryndina, A. A. Syemov // Processings of the 11th International conference on Pattern Recognition and Image Analasis: New Information technologies (PRIA-11-2013). – Samara : IPSI RAS. – 2013. – Vol. I. – P. 186–189.

12. Сёмов, А. А. Об одном подходе к распознаванию 3D-изображений / А. А. Сёмов // Надежность и качество : тр. Междунар. симп. – Пенза : Изд-во ПГУ, 2013. – Т. 1. – С. 350–351.

13. Syemov, A. Intelligent capabilities hypertrace transform: constructing features with predetermined properties / N. Fedotov, A. Syemov, A. Moiseev // Processings of the International conference “Intelligent Information Processing” IIP-10: theses of reports of the 10th international conference, Greece, Crete, Hersonissos. – M. : Torus Press, 2014. – P. 111.

14. Сёмов, А. А. Роль гиперфункционалов в гипертрейс-преобразовании и повышение надежности распознавания 3D объектов / А. А. Сёмов // Надежность и качество : тр. Междунар. симп. – Пенза : Изд-во ПГУ, 2014. – Т. 1. – С. 393–396.

15. Сёмов, А. А. Особенности реализации 3D трейс-преобразования. Способы ускорения его вычислений гипертриплетных признаков / Н. Г. Федотов, А. А. Сёмов, Е. А. Крючкова // Надежность и качество : тр. Междунар. симп. – Пенза : Изд-во ПГУ, 2014. – Т. 1. – С. 396–399.

16. Сёмов, А. А. Об особенностях техники сканирования 3D изображений / Н. Г. Федотов, А. А. Сёмов, Е. А. Крючкова // Современные методы и средства обработки пространственно-временных сигналов : сб. ст. XII Всерос. науч.-техн. конф. – Пенза : Приволжский Дом знаний, 2014. – С. 138–144.

17. Сёмов, А. А. Гипертрейс-преобразование, инвариантное к группе движений 3D-объектов / Н. Г. Федотов, А. А. Сёмов // Современные методы и средства обработки пространственно-временных сигналов : сб. ст. XII Всерос. науч.-техн. конф. – Пенза : Приволжский Дом знаний, 2014. – С. 38–43.

18. Сёмов, А. А. Проблемы распознавания 3D изображений у машин и людей: сравнительная характеристика / Н. Г. Федотов, А. А. Сёмов, А. А. Курносов // Проблемы информатики в образовании, управлении, экономике и технике : сб. ст. XIV Междунар. науч.-техн. конф. – Пенза : Приволжский Дом знаний, 2014. – С. 185–193.

19. Сёмов, А. А. Интеллектуальные возможности гипертрейс-преобразования: конструирование признаков с заданными свойствами / Н. Г. Федотов, А. А. Сёмов, А. В. Моисеев // Машинное обучение и анализ данных. – 2014. – Т. 1, № 9. – С. 1200–1214.

20. Сёмов, А. А. К вопросу извлечения параметров 3D объектов из результатов сканирования / Н. Г. Федотов, А. А. Сёмов // Новые информационные технологии и системы (НИТИС-2014) : сб. тр. XI Междунар. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2014. – С. 146–149.

21. Сёмов, А. А. Актуальность разработки нового метода анализа и распознавания 3D объектов / А. А. Сёмов, М. А. Сёмов // Модели, системы, сети в экономике, технике, природе и обществе. – 2014. – № 2 (10). – С. 199–204.

Регистрация программного продукта

22. Свидетельство об официальной регистрации программ для ЭВМ № 2015612257 Роспатента от 16.02.15. Программный комплекс анализа и распознавания 3D изображений на основе пространственного трейс-преобразования со случайными параметрами сканирования / Н. Г. Федотов, А. А. Сёмов.

23. Свидетельство об официальной регистрации программ для ЭВМ № 2015612814 Роспатента от 26.02.15. Программный комплекс анализа и распознавания 3D изображений на основе пространственного трейс-преобразования с детерминированными параметрами сканирования / А. А. Сёмов.

Научное издание

СЁМОВ Алексей Александрович

**МЕТОДЫ АНАЛИЗА И РАСПОЗНАВАНИЯ
ТРЕХМЕРНЫХ ИЗОБРАЖЕНИЙ
НА ОСНОВЕ СТОХАСТИЧЕСКОЙ ГЕОМЕТРИИ**

Специальность 05.13.17 – Теоретические основы информатики

Редактор *Т. Н. Судовчихина*
Технический редактор *Р. Б. Бердникова*
Компьютерная верстка *Р. Б. Бердниковой*

Распоряжение № 23/78-2015 от 22.10.2015.
Подписано в печать 22.10.2015. Формат 60×84^{1/16}.
Усл. печ. л. 1,16. Заказ № 932. Тираж 100.

Издательство ПГУ.
440026, Пенза, Красная, 40.
Тел./факс: (8412) 56-47-33; e-mail: iic@pnzgu.ru