

ОБРАБОТКА ИЗОБРАЖЕНИЙ, РАСПОЗНАВАНИЕ ОБРАЗОВ

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ГИПЕРТРЕЙС-ПРЕОБРАЗОВАНИЯ: ТЕХНИКА СКАНИРОВАНИЯ, МАТЕМАТИЧЕСКИЙ АППАРАТ И ЭКСПЕРИМЕНТАЛЬНАЯ ПРОВЕРКА

Н.Г. Федотов¹, А.А. Сёмов², А.В. Моисеев¹

¹ ФГБОУ ВО «Пензенский государственный университет», Пенза, Россия,

² ООО «КОМХЭЛФ», Пенза, Россия

Аннотация

В статье последовательно описывается теоретическая основа нового геометрического метода анализа и распознавания трёхмерных (3D) изображений. Дается описание техники сканирования для формирования гипертрейс-преобразования и его математическая модель. Данный метод в отличие от существующих позволяет анализировать 3D-изображения без предварительного их упрощения или построения проекций на плоскости, анализируя непосредственно их трёхмерную форму. Обоснован выбор сканирующего инструмента и построение опорной сетки на сфере, необходимой для решения проблемы инвариантности распознавания 3D-изображения к повороту. Разработан математический аппарат стохастической реализации техники сканирования на основе стохастической геометрии и функционального анализа. Введён новый математический инструмент для анализа 3D-изображений – гипертрейс-матрица, позволяющий распознавать пространственные объекты сложной формы и структуры благодаря построению единой математической модели 3D-изображения. Представлено описание нового типа признаков 3D-изображений, имеющих аналитическую структуру, – гипертриплетные признаки, благодаря аналитической структуре которых возможна автоматическая генерация большого количества признаков с заранее заданными свойствами. Рассматриваются результаты экспериментальной проверки, демонстрирующие точность вычисления признаков для распознавания 3D-изображения и доказывающие адекватность разработанного математического аппарата.

Ключевые слова: распознавание 3D-изображений, геометрическое гипертрейс-преобразование, сетка параллельных плоскостей, стохастическое сканирование, аналитическая структура признака, гипертрейс-матрица, инвариантность распознавания.

Цитирование: Федотов, Н.Г. Теоретические основы гипертрейс-преобразования: техника сканирования, математический аппарат и экспериментальная проверка / Н.Г. Федотов, А.А. Сёмов, А.В. Моисеев // Компьютерная оптика. – 2018. – Т. 42, № 2. – С. 273-282. – DOI: 10.18287/2412-6179-2018-42-2-273-282.

Введение

В настоящее время в условиях интенсивного развития компьютерных технологий возрастает роль методов обработки данных, представленных в виде изображений. В качестве примера можно привести такие области, как аэрокосмические исследования, машинное зрение, трёхмерная компьютерная томография, исследования Земли из космоса, автоматизация процесса компьютерного распознавания в дефектоскопии, дактилоскопии и тому подобные [1–4].

Для современного этапа развития теории распознавания образов актуально расширение круга рассматриваемых задач распознавания на трёхмерные (далее 3D) изображения, в то время как ранее внимание исследователей было сосредоточено на решении задач анализа и распознавания двумерных (далее 2D) изображений. Эффективность машинного зрения при распознавании пространственных изображений значительно уступает возможностям человека, что объясняется не слабой мощностью вычислительных средств, а недостаточно развитыми теоретическими методами и подходами.

В данной статье предлагается подход на основе стохастической геометрии и функционального анализа, который благодаря построению признаков, инвари-

антных к группе движений и масштабированию, позволяет повысить надёжность и универсальность распознавания. Сканирование со случайными параметрами улучшает соотношение «надёжность–быстродействие» распознавания 3D-изображений по сравнению с детерминированным сканированием.

1. Инструмент сканирования

Пусть F – исходная трёхмерная модель. Авторы рассматривали полигональные модели, однако метод применим и к другим видам моделей. Определим плоскость $B(\eta, r) = \{x | x^T \cdot \eta = r\}$ как касательную к сфере с центром в начале координат и радиусом r , проходящую через заданную точку X на расстоянии r от начала координат с заданными углами ω и ϕ (рис. 1), где $\eta = [\cos \omega \cdot \sin \omega, \cos \phi]$ – единичный вектор в R^3 , ω – угол между осью Ox и проекцией отрезка OX на плоскость Oxy , ϕ – угол между осью Oz и отрезком OX .

Сканирование 3D-изображения будет осуществляться плоскостями. Данный выбор оправдан тем, что пересечение плоскости с любым другим пространственным геометрическим примитивом хорошо известно в высшей математике и имеет строгое аналитическое представление, которое нам пригодится при разработке аналитической структуры признака.

Наличие сетки параллельных плоскостей помогает решить проблему инвариантного описания объекта (даёт необходимое условие для конструирования признаков, инвариантных к переносу 3D-изображения). Если пространственный объект сканируется сеткой параллельных плоскостей, то перемещение исходного 3D-изображения не изменяет форму получаемых сканирующими плоскостями сечений (дискретный шаг сканирования игнорируется). Поэтому получаемые сечения и извлекаемые на их основе признаки не изменят своего значения при переносе 3D-изображения на любой вектор в пространстве. Другими словами, вычисляемые признаки не будут зависеть от пространственного положения объекта.

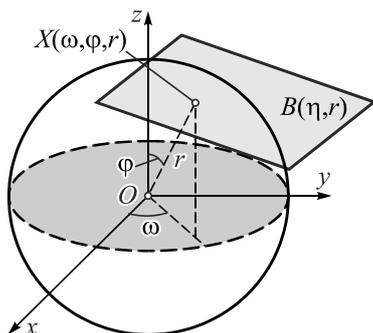


Рис. 1. Определение сферических координат плоскости

2. Проблема инвариантности распознавания 3D-изображения при его произвольной ориентации в пространстве

Для трёхмерного случая наличие произвольной пространственной ориентации 3D-изображения сильно осложняет создание его инвариантного описания. В отличие от двумерного случая, проблема поворота 3D-изображения в трёхмерном пространстве связана с невозможностью его поворота вокруг одной оси, независимо от совершённого вращения по другим осям.

Чтобы схема сканирования 3D-изображения не была привязана к пространственной ориентации объекта, необходимо и достаточно, чтобы сканирующие элементы, если все их одновременно зафиксировать в пространстве, давали одинаковые сечения трёхмерного объекта при любом его угле вращения. Другими словами, необходимо добиться, чтобы все сканирующие сетки параллельных плоскостей под разными углами ω и φ обзора распознаваемого 3D-изображения совпадали бы друг с другом при любом его пространственном повороте (дискретный шаг игнорируется).

Стандартный перебор всех углов ω и φ , которыми идентифицируется каждая сканирующая сетка параллельных плоскостей, в топологическом смысле для непрерывного случая даёт модель концентрических сфер с центром в начале координат. Каждой сканирующей сетке параллельных плоскостей на единичной сфере сопоставим точку, которая будет являться точкой касания со сферой плоскости, параллельной плоскостям данной сетки (отдельно для каждой пары (ω, φ) углов обзора). Множество точек на сфере образуют сетку, которую будем называть опорной (рис. 2).

Стоит отметить, что пара углов (ω, φ) однозначно определяет узел опорной сетки, соответствующий единственной касательной плоскости к сфере в этой точке, а значит, и единственной сетке сканирующих параллельных плоскостей, что подробно рассмотрено в [5].

Для дискретного случая на обычной карте глобуса вблизи полюса наблюдается более плотное скопление точек, чем у экватора. Поэтому если при повороте полюс совместить с точкой на экваторе, то будут заметны отклонения точек исходной и повернутой сеток. Так как каждая точка опорной сетки на сфере однозначно определяет угол наклона сетки параллельных плоскостей, то изменение угла наклона сетки плоскостей повлияет на форму получаемых сечений. Вследствие этого увеличится ошибка расчёта признака и снизится точность распознавания 3D-объекта.

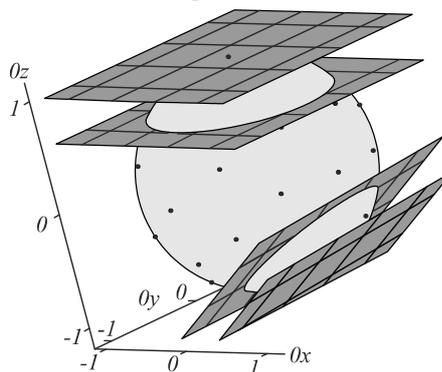


Рис. 2. Опорная сетка на сфере и соответствующие ей сетки сканирующих параллельных плоскостей

С другой стороны, если при повороте сферы вокруг своего центра опорная сетка перейдёт сама в себя, то соответствующие сетки секущих параллельных плоскостей полностью совпадут друг с другом и получаемые сечения будут одинаковыми (не изменят своей формы). Поэтому вычисляемое значение признака не изменится.

Таким образом, необходимо построить опорную сетку, обладающую равномерным распределением точек на сфере для достижения меньшей ошибки при совмещении узлов опорной сетки при её повороте из-за дискретного шага сканирования. Равномерное распределение точек опорной сетки на сфере обеспечит отсутствие более плотных скоплений узлов опорной сетки на поверхности сферы, определяющих преимущественно сечения под теми или иными углами обзора объекта. В связи с этим все результаты сканирования будут принимать равноправное участие при вычислении значения признака 3D-изображения без повышения влияния каких-либо определённых значений сечений, так как частота появления любого среза сечений будет приблизительно одинакова (равномерный обзор трёхмерного тела со всех сторон). Другими словами, значение вычисляемого признака не будет зависеть от ориентации 3D-изображения в пространстве.

3. Равномерная опорная сетка на сфере

Указанное свойство равномерного распределения точек опорной сетки на сфере является необходимым

условием инвариантности конструируемых признаков к повороту. Математическая формулировка данной проблемы имеет следующий вид.

Рассмотрим сферу, заданную параметрически: $x(\omega, \phi) = R \cdot \cos \omega \cdot \sin \phi$, $y(\omega, \phi) = R \cdot \sin \omega \cdot \sin \phi$, $z(\omega, \phi) = R \cdot \cos \phi$. Необходимо определить аналитически функцию $f(\omega, \phi)$ плотности совместного распределения параметров ω и ϕ , соответствующую равномерному распределению точек на поверхности сферы.

Рассмотрим небольшой участок dS поверхности сферы, ограниченный приращениями $d\phi$ и $d\omega$ (рис. 3).

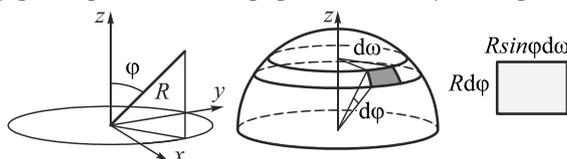


Рис. 3. Участок на поверхности сферы и значение его сторон

В случае, когда точки имеют равномерное распределение на поверхности сферы, вероятность попадания произвольной точки A на элемент поверхности dS с одной стороны равна: $P(A \in dS) = dS/S$.

При постоянном значении ϕ изменение угла $d\omega$ описывает дугу $R \cdot \sin \phi \cdot d\omega$. Поэтому площадь малого элемента поверхности сферы равна $dS = R^2 \cdot \sin \phi \cdot d\omega \cdot d\phi$. Таким образом, вероятность попадания произвольной точки A на элемент поверхности dS будет равна:

$$P(A \in dS) = \frac{R^2 \cdot \sin \phi \cdot d\omega \cdot d\phi}{\int_0^\pi \int_0^{2\pi} R^2 \cdot d\omega \cdot d\phi} = \frac{R^2 \cdot \sin \phi \cdot d\omega \cdot d\phi}{4 \cdot \pi \cdot R^2} = \frac{\sin \phi \cdot d\omega \cdot d\phi}{4\pi}.$$

С другой стороны, вероятность попадания точки A на данный элемент поверхности равна: $P(A \in dS) = f(\omega, \phi) \cdot d\omega \cdot d\phi$. Следовательно, совместная плотность распределения вероятности ω и ϕ равна:

$$\frac{\sin \phi \cdot d\omega \cdot d\phi}{4\pi} = f(\omega, \phi) \cdot d\omega \cdot d\phi \Rightarrow f(\omega, \phi) = \frac{\sin \phi}{4\pi}.$$

Генерируя значения параметров ω и ϕ с использованием функции плотности распределения $f(\omega, \phi)$, получаем равномерное распределение точек на поверхности сферы. Различные алгоритмы построения равномерного распределения конечного числа точек рассмотрены в [6, 7].

Данное условие является необходимым условием инвариантности конструируемых признаков к повороту 2D-изображения в плоскости сечения и 3D-изображения в целом.

4. Описание техники сканирования

Сканирование 3D-изображения производится сеткой параллельных плоскостей с расстоянием Δr между плоскостями и заданными углами ω и ϕ . Взаимное положение 3D-изображения F и каждой сканирующей плоскости $B(\eta(\omega, \phi), r)$ характеризуется числом G , которое определяется по некоторому правилу

HuperT: $G = \text{HuperT}(F \cap B(\eta(\omega, \phi), r))$. В качестве указанного правила можно использовать вычисление площади сечения, количество пересечений плоскости с исходным объектом, свойства окрестности полученного сечения и тому подобное. Другими словами, функционал HuperT характеризует свойство признака сечения.

Затем сканирование производится сеткой параллельных плоскостей для нового значения угла $\omega + \Delta\omega$ и $\phi + \Delta\phi$, получившего дискретные приращения $\Delta\omega$ и $\Delta\phi$ соответственно, с тем же шагом Δr между сканирующими плоскостями. К сечениям новой сетки из плоскостей $B(\eta(\omega + \Delta\omega, \phi + \Delta\phi), r_i)$ применяется такое же ранее выбранное правило HuperT.

Сканирование повторяется для каждого нового угла $\omega + \Delta\omega$ и $\phi + \Delta\phi$ до завершения оборота в 2π и π радиан соответственно. Важно отметить, что углы изменяются не произвольным образом, а согласно построению опорной сетки на сфере, обеспечивающей равномерную плотность плоскостей в пространстве (для выполнения необходимого условия инвариантности конструируемых признаков к повороту).

Результат вычислений функционала HuperT зависит от трёх параметров плоскости (r, ω, ϕ) . Поэтому, если каждому 2D-изображению, полученному при сечении исходной трёхмерной модели сканирующей плоскостью, сопоставить некоторый информативный признак $\Pi(F_{sect})$ по правилу HuperT, то при численном анализе результат гипертрейс-преобразования удобно представить в виде трёхмерной гипертрейс-матрицы ЗТМ, у которой ось 0ϕ направлена вертикально, ось 0ω – горизонтально, ось $0r$ – вглубь.

Каждое 2D-изображение, получившееся в сечении исходной 3D-модели сеткой параллельных плоскостей под разными углами обзора, сканируется, чтобы извлечь какие-нибудь значимые признаки (например, периметр контура фигуры сечения). Для нахождения признака двумерного изображения сечения используется трейс-преобразование, которое будет описано ниже.

Сканирование получаемых в сечении изображений F_{sect} осуществляется решеткой параллельных прямых $l(\theta, \rho)$ с расстоянием $\Delta\rho$ между линиями, где ρ , θ – полярные координаты прямой в плоскости сечения. Взаимное положение 2D-изображения F_{sect} и каждой сканирующей линии $l(\theta, \rho)$ характеризуется числом g , вычисляемым по некоторому правилу T: $g = T(F_{sect} \cap l(\theta, \rho))$.

В качестве указанного правила можно использовать вычисление длины части прямой, лежащей внутри изображения, свойства окрестности точки пересечения прямой с изображением и тому подобное.

Затем сканирование производится для нового значения угла $\theta + \Delta\theta$, получившего дискретное приращение $\Delta\theta$, сеткой параллельных прямых в той же плоскости сечения F_{sect} и с тем же шагом $\Delta\rho$. К пересечению новой прямой $l(\theta + \Delta\theta, \rho)$ и сечения F_{sect} применяется такое же ранее выбранное правило T. Сканирование повторяется для каждого нового угла $\theta + \Delta\theta$ до завершения оборота в 2π радиан.

Результат вычислений функционала T зависит от двух параметров прямой θ и ρ . При численном анализе результат трейс-преобразования удобно представить в виде двумерной трейс-матрицы TM , у которой ось 0θ направлена горизонтально, а ось 0ρ – вертикально. Подробное описание свойств трейс-преобразования можно найти в [8].

5. Режимы сканирования и стохастическая реализация

Данный метод поддерживает несколько различных режимов сканирования 3D-изображений, которые имеют свои преимущества и недостатки в различных практических ситуациях:

1) Детерминированный режим сканирования.

В этом случае опорная сетка на сфере имеет детерминированную структуру, и расстояние между сканирующими плоскостями одинаково для любой пары (ω, ϕ) углов обзора. Аналогичное свойство наблюдается в плоскости сечения для опорной сетки на окружности и для расстояния между сканирующими прямыми.

2) Стохастический режим сканирования.

В данном случае опорная сетка на сфере имеет стохастическую структуру, и расстояние между сканирующими плоскостями случайно, но обладает равномерной плотностью в пространстве в пределах размеров 3D-объекта отдельно для каждой пары (ω, ϕ) углов обзора. Аналогичное свойство наблюдается в плоскости сечения для опорной сетки на окружности и для расстояния между сканирующими прямыми.

Кроме того, заранее предопределённый алгоритм построения равномерной сетки на сфере по скорости выполнения будет заметно уступать генерации случайной сетки с равномерным распределением точек на её поверхности.

Стоит отметить, что стохастическое сканирование со случайными параметрами улучшает соотношение «надёжность – быстродействие» распознавания 2D-изображений по сравнению с детерминированным сканированием с фиксированными параметрами, что показано в [9]. Аналогичное свойство справедливо также и при анализе и распознавании 3D-изображений.

При стохастическом сканировании строки и столбцы гипертрейс-матрицы будут обозначать не номер, а диапазон варьирования каждой из переменных ω , ϕ и r . Поэтому перед формированием гипертрейс-матрицы строится трёхмерное вероятностное распределение переменных, аналогично стандартным методам стохастической геометрии.

Так, при проектировании распознающих систем используются дискретные представления переменных ω , ϕ и r . Соответственно, параметры сканирующих плоскостей образуют три множества:

$$M_1 = \{\omega_1, \omega_2, \dots, \omega_m\}, M_2 = \{\phi_1, \phi_2, \dots, \phi_n\},$$

$$M_3 = \{r_1, r_2, \dots, r_l\}.$$

Детерминированное сканирование позволяет однозначно определить каждый элемент матриц 3TM. Так, тройке (ω_i, ϕ_j, r_k) соответствует элемент матрицы с но-

мером (i, j, k) и значением $\text{HyperT}(F \cap B(\eta(\omega, \phi), r))$. При применении случайного сканирования на множествах M_1, M_2, M_3 задаются вероятностные распределения:

$$M_1 = \{p(\omega_1), p(\omega_2), \dots, p(\omega_m)\},$$

$$M_2 = \{q(\phi_1), q(\phi_2), \dots, q(\phi_n)\},$$

$$M_3 = \{s(r_1), s(r_2), \dots, s(r_l)\},$$

$$\sum_i p(\omega_i) = 1, \sum_j q(\phi_j) = 1, \sum_k s(r_k) = 1.$$

При стохастическом способе генерирования равномерного распределения точек на сфере используются генераторы псевдослучайных чисел, которые позволяют равномерные сетки на сфере для любого количества точек. Авторы решили модифицировать алгоритм Saff [6] для построения детерминированной спирали. Каждая точка данной спирали получила случайный равномерный прирост по каждой координате на величину не более $1/\sqrt{N}$. После чего расстояние от точки до начала координат приводится к единичному масштабу.

Выбирая параметры сканирования случайным образом на основе заданных распределений, получаем трёхмерное распределение случайных величин в виде матрицы 3TM, где $G_{i,j,k} = \text{HyperT}(F \cap B(\eta(\omega_i, \phi_j), r_k))$ с вероятностью $p(\omega_i) \cdot q(\phi_j) \cdot s(r_k)$ и $G_{i,j,k} = 0$ с вероятностью $1 - p(\omega_i) \cdot q(\phi_j) \cdot s(r_k)$. Аналогичное свойство имеет место и для трейс-матриц в плоскости сечения, которые являются двумерным распределением случайных величин θ и ρ .

Стохастическое сканирование 3D-изображений можно реализовать в следующем виде: случайная величина r имеет непрерывное равномерное распределение на отрезке $[a, b]$ с функцией плотности

$$f(r) = \begin{cases} \frac{1}{b-a}, & r \in [a, b], \\ 0, & r \notin [a, b], \end{cases}$$

случайные величины ω , ϕ имеют совместное равномерное распределение на поверхности сферы с функцией плотности

$$f(\omega, \phi) = \frac{\sin \phi}{4\pi}.$$

6. Математический инструмент для анализа 3D-изображения

Трёхмерная гипертрейс-матрица 3TM – математический инструмент для анализа 3D-изображения. Гипертрейс-матрица 3TM – 3D-трейс-образ исходного пространственного изображения, который содержит информацию о первоначальном объекте.

Каждая глубинная строка матрицы содержит элементы-признаки, которые вычисляются по 2D-изображениям, полученным в результате сечений исходного 3D-объекта сканирующими плоскостями при обходе всех значений переменной расстояния r с фиксированными значениями углов ω и ϕ . Соответственно, каж-

дый вертикальный столбец матрицы содержит элементы-признаки, которые вычисляются по 2D-изображениям, полученным в результате сечений исходного 3D-объекта сканирующими плоскостями при обходе всех углов φ (от 0 до π радиан) при одинаковых значениях угла ω и расстояния r . И, наконец, каждая глубинная строка матрицы содержит элементы-признаки, которые вычисляются по 2D-изображениям, полученным в результате сечений исходного 3D-объекта сканирующими плоскостями для всех значений расстояний r при фиксированных значениях углов ω и φ .

Если плоскость B не пересекает 3D-изображение, то есть $F \cap B(\eta(\omega, \varphi), r) = \emptyset$, то значение гипертрейс-функционала полагают равным нулю:

$$\text{HyperT}(F \cap B(\eta(\omega, \varphi), r)) = 0.$$

Графическое представление гипертрейс-матрицы 3ТМ называется гипертрейс-трансформантой, где полученное в результате сканирования множество чисел G образует точки $(\omega_i, \varphi_j, r_k)$ в системе координат с осями $O\omega$, $O\varphi$ и Or . Так как результат вычислений функционала $\text{HyperT}(F \cap B(\eta(\omega_j, \varphi_i), r_k))$ зависит от трёх параметров сканирующей плоскости r , ω и φ , имеющих дискретный шаг сканирования, то реальная гипертрейс-трансформанта имеет дискретную структуру.

Двумерная трейс-матрица ТМ – математический инструмент для анализа 2D-изображения. Трейс-матрица ТМ – 2D-трейс-образ плоского изображения 2D-фигуры сечения, полученного при пересечении плоскости $B(\eta(\omega, \varphi), r)$ исходного 3D-объекта F .

Каждый вертикальный столбец матрицы ТМ содержит значения, вычисляемые по всем прямым сканирующей сетки при одинаковом значении угла θ и различными значениями ρ внутри конкретного 2D-изображения сечения. Соответственно, каждая горизонтальная строка матрицы ТМ содержит значения, вычисляемые для всех прямых l , имеющих одинаковое расстояние ρ до начала координат и различных значениях угла θ внутри конкретного 2D-изображения сечения.

Если прямая l не пересекает изображение: $F_{\text{sect}} \cap l(\theta, \rho) = \emptyset$, то значение трейс-функционала полагают равным нулю $T(F_{\text{sect}} \cap l(\theta, \rho)) = 0$.

Графическое представление трейс-матрицы ТМ называется трейс-трансформантой, где полученное в результате сканирования множество чисел g образуют точки (θ_i, ρ_j) в системе координат с осями $O\theta$ и $O\rho$.

Таким образом, паре (θ_i, ρ_j) соответствует элемент матрицы ТМ с номером (i, j) и значением $T(F_{\text{sect}} \cap l(\theta_i, \rho_j))$.

Данный математический инструмент будет особо полезным в области медицинской томографии при анализе и извлечении метрических характеристик пространственных объектов (внутреннего органа человека). С помощью гипертрейс-матрицы возможен не только анализ движущихся 3D-объектов, но и извлечение параметров этих движений. Подробнее об этом можно найти в [5].

7. Особенности формирования и построения гипертрейс-матрицы

Если результаты сканирования 3D-изображения (признаки сечений) заносить в элементы матрицы от точки $(0, 0, 1)$ оси Oz в направлении по часовой стрелке от оси Ox , то при формировании гипертрейс-матрицы 3ТМ возникают определённые трудности сохранения целостности её структуры – нарушается порядок следования строк и столбцов друг за другом в матрице, так как равномерная сетка на сфере неизоморфна равномерной сетке на плоскости.

При повороте 3D-изображения глубинные строки сохраняют порядок следования элементов, так как при повороте не изменяется порядок следования сечений (дискретная форма пространственного объекта). Однако при повороте 3D-изображения порядок следования элементов в горизонтальных строках (ось $O\omega$) и вертикальных столбцах (ось $O\varphi$) матрицы 3ТМ в общем случае без наложения дополнительных условий не сохраняется из-за произвольной пространственной ориентации 3D-изображения.

Так, при обработке функционалами трёхмерная матрица сворачивается в число в строго заданных направлениях (сначала глубинные, далее вертикальные, а затем горизонтальные строки). Поэтому из-за произвольной неизвестной ориентации тела в пространстве возможен случайный поворот матрицы относительно оси Or в пространстве $O\omega\varphi r$. Произвольное нарушение порядка следования строк матрицы друг за другом в данных фиксированных направлениях приведёт к изменению значения вычисляемого признака, и, как следствие, инвариантность распознавания 3D-изображения будет нарушена.

В связи с этим правила нумерации узлов опорной сетки, по которым формируется трёхмерная гипертрейс-матрица и определяется порядок следования её элементов, должны определяться не относительно координатных осей, а относительно произвольно ориентированного пространственного объекта.

Для определения направления нумерации узлов опорной сетки достаточно идентифицировать некоторые опорные ключевые точки – узлы опорной сетки, которые однозначно определяются вне зависимости от пространственной ориентации 3D-объекта. Данные опорные точки определяют начало отсчёта (первую глубинную строку), от которого начинается заполняться гипертрейс-матрица. Например, построение гипертрейс-матрицы при нумерации узлов опорной сетки от первой ключевой точки по часовой стрелке в направлении второй ключевой точки для заполнения элементов матрицы.

Ключевые точки характеризуют уникальные свойства пространственного объекта, на основании которых могут быть построены гипертриплетные признаки, инвариантные к повороту 3D-изображения. Так, в качестве ключевой точки, например, можно взять узел на опорной сетке, который соответствует сетке секущих плоскостей, содержащей максимальное по площади сечение исходного 3D-объекта.

Стоит отметить, что при повороте объекта секущие плоскости (например, сечения, перпендикулярные главной оси объекта) будут соответствовать другим точкам опорной сетки, чем точкам сетки до его поворота. Поэтому необходимо задать правило, определяющее соотношение каждой точки опорной сетки смещённого объекта с соответствующей точкой опорной сетки исходного объекта. Другими словами, необходимо создать правило инвариантной нумерации узлов опорной сетки, которое позволяет получать однородную структуру матрицы, при обработке функционалами строк и столбцов которой получаются одинаковые значения признака вне зависимости от пространственной ориентации 3D-объекта (из-за дискретного шага сканирования будет присутствовать ограниченный сверху уровень колебания в значениях вычисляемого признака).

Порядок следования горизонтальных строк (ось 0ω) и вертикальных столбцов (ось 0ϕ) при произвольной пространственной ориентации трёхмерного тела имеет некоторые особенности в зависимости от метода построения гипертрейс-матрицы 3ТМ. Особенности формирования данных матриц и их разновидностей можно найти в [5].

8. Конструирование признаков 3D-изображения

После заполнения трёхмерной гипертрейс-матрицы обрабатываются её глубинные строки с помощью функционала $\text{Hyper } P$, который можно задать, например, как $\text{Hyper } P = \int G(\omega, \phi, r) dr$. В результате исходная трёхмерная гипертрейс-матрица 3ТМ становится двумерной матрицей 2ТМ. Далее применяется постолбцовая обработка матрицы 2ТМ посредством функционала $\text{Hyper } \Omega$, который можно задать, например, как

$$\text{Hyper } \Omega = \max_{\phi} G(\omega, \phi).$$

В результате получается горизонтальная строка 1ТМ – вектор значений, непрерывным аналогом которого в общем случае является 2π -периодическая кривая. К полученному набору чисел применяют функционал $\text{Hyper } \Theta$, что приводит к появлению некоторого числа – признака изображения $\text{Res}(F)$. Этот функционал можно задать, например, амплитудой второй гармоники при разложении вектора в ряд Фурье.

Таким образом, гипертриплетный признак 3D-изображения F обладает структурой в виде композиции четырёх функционалов, каждый из которых, кроме функционала $\text{Hyper } T$, при последовательном применении сокращает размерность матрицы 3ТМ на единицу:

$$\text{Res}(F) = \text{Hyper } \Theta \circ \text{Hyper } \Omega \circ \text{Hyper } P \circ \text{Hyper } T(F_{sect}).$$

После заполнения двумерной трейс-матрицы с помощью диаметрального функционала T обрабатываются столбцы матрицы ТМ. Его можно задать, например, как

$$T = \frac{\int g(\theta, \rho) d\rho}{\max_{\rho} g(\theta, \rho)}.$$

В результате исходная двумерная матрица ТМ становится одномерной матрицей – вектором чисел, непрерывным аналогом которого будет 2π -периодическая кривая. Затем к полученному набору чисел применяют функционал Θ , который можно задать как

$$\Theta = \min_{\theta} g(\theta).$$

В результате получается число $\Pi(F_{sect})$ – признак 2D-изображения сечения F_{sect} .

Таким образом, триплетный признак 2D-изображения F_{sect} обладает структурой в виде композиции трёх функционалов, каждый из которых, кроме функционала T , при последовательном применении сокращает размерность матрицы ТМ на единицу [7]:

$$\Pi(F_{sect}) = \text{Hyper } T(F_{sect}) = \Theta \circ P \circ T(F_{sect} \cap I(\theta, \rho)).$$

Объединяя полученные формулы для $\text{Res}(F)$ и $\Pi(F_{sect})$, окончательно получаем следующую аналитическую структуру признака 3D-изображения в виде композиции множества функционалов:

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

Благодаря композиционной структуре функционалов, входящих в структуру $\Pi(F_{sect})$ и $\text{Res}(F)$, возможно получение огромного числа признаков. Причем возможно конструирование признаков, описывающих те или иные геометрические характеристики 3D-объекта, что облегчает задачу анализа свойств 3D-изображений и построения информативных признаков [5].

Стоит отметить, что расположение системы координат в плоскости сечения и её ориентация относительно фигуры сечения совершенно несущественны, так как трейс-преобразование полностью инвариантно к группе движений и масштабированию 2D-изображения [6].

9. Планирование эксперимента

В отличие от большинства известных методов, рассмотренных ранее в работе, предлагаемый метод позволяет в режиме автоматической генерации формировать большое число признаков, которые будут являться как инвариантными, так и чувствительными к группе движений и масштабированию 3D-изображений [5]. Данные гипертриплетные признаки имеют аналитическую структуру, которая позволяет строить признаки с заданными свойствами, описывающие те или иные метрические характеристики пространственного объекта. В силу того, что реальное сканирование происходит с дискретным шагом, то возникает определённая погрешность вычисления данных признаков. Поэтому целесообразной становится задача оценки погрешности вычисления признаков при различных параметрах движения и масштабирования 3D-изображений.

Так как функционалы, входящие в состав признака, имеют композиционную структуру, то по мере уменьшения шага число сканирований 3D-изобра-

жения растёт комбинаторно, и время вычисления признака заметно возрастает. Поэтому целесообразно исследовать сканирование 3D-изображения на оптимальность соотношения «надёжность – быстродействие» распознающей системы при использовании различных параметров сканирования (количество узлов опорной сетки на сфере, количество узлов опорной сетки на окружности и другие).

Все рассматриваемые в данной статье 3D-изображения были взяты из известной базы данных принстонского университета The Princeton Shape Benchmark [10]. Каждый пространственный объект данной базы имеет полигональную структуру поверхности. Данная база включает наборы 3D-моделей различных классов: от предметов неживой природы (двери, машины, самолёты и другие) до объектов живой природы (муравьи, кошки, люди, птицы и другие). Количество полигонов в моделях объектов варьируется в диапазоне от 1000 до 300 000 штук.

Реализация и экспериментальная проверка предложенного метода 3D-трейс-преобразования производилась с использованием инженерного пакета MathCAD 15 M030. Визуализация трёхмерных объектов осуществлялась с помощью средств пакета MeshLab 1.3.

Прежде, чем приступить к анализу параметров стохастического сканирования и определения их оптимальной структуры, необходимо сделать несколько замечаний.

Во-первых, более общим параметром, влияющим на точность вычисления признака, является параметр $N\phi$ – количество узлов опорной сетки на сфере.

Во-вторых, для детерминированного сканирования используется 4 параметра: $\Delta 3$ – расстояние между параллельными плоскостями, $N\phi$ – количество различных углов сканирования плоскостями 3D-изображения (количество узлов равномерной опорной сетки на сфере), $\Delta 2$ – расстояние между параллельными прямыми в плоскости сечения, $N\theta$ – количество различных углов сканирования прямыми 2D-изображения сечения (количество узлов равномерной опорной сетки на окружности).

Для стохастического сканирования показатель $\Delta 3$ показывает усреднённое расстояние между параллельными плоскостями: $\Delta 3 = \text{dist}/\text{cross}$, где dist – расстояние между двумя граничными касательными плоскостями в сетке плоскостей под данным углом обзора, cross – количество пересечений пространственного объекта сканирующей плоскостью. Аналогичное замечание будет и для показателя $\Delta 2$. Параметры $N\phi$ и $N\theta$ для стохастического сканирования характеризуют равномерное случайное распределение точек опорной сетки на сфере и окружности соответственно.

В-третьих, полноценный анализ всех параметров сканирования затруднён, так как точность вычисления признака меняется при совместном изменении параметров сканирования, причём как в прямом, так

и обратном направлениях прироста значений данных параметров.

Чтобы при анализе конкретной переменной (например, Δr) нивелировать эффект влияния других параметров сканирования изображения, все остальные параметры сканирования ($N\phi$, $\Delta\rho$, $N\theta$) должны иметь значения, при которых точность вычисления признака является максимально высокой. Тем самым, точность сканирования 3D-изображения будет являться чувствительной к изменению только одного параметра (Δr) и не зависеть от уровня значений других параметров.

В связи с этим анализ времени вычисления признака не даст информативной картины, так как вычисление признака будет осуществляться достаточно долго. В этом случае необходимо руководствоваться общим принципом при отборе требуемого уровня точности вычисления признака: чем крупнее шаг сканирования и меньше число узлов на сетке, тем меньше времени требуется на его вычисление.

В-четвёртых, для стохастического сканирования эксперимент повторялся сто раз, а затем вычислялись усреднённые характеристики применяемых параметров.

10. Экспериментальная часть: определение оптимальных параметров сканирования

В качестве анализируемого пространственного объекта была взята 3D-модель храмового комплекса с кодовым номером m405 в указанной выше базе данных [10]. Вычисляемый признак Res данного объекта имеет следующую структуру:

$$\text{Res} = \text{Hyper } \Theta \circ \text{Hyper } \Omega \circ \text{Hyper } P \circ \text{Hyper } T(\Theta \circ P \circ T),$$

где

$$T = \sum_t f(\theta, \rho, t), \quad P = \max_{\rho} \left(\left| \frac{\Delta g(\theta, \rho)}{\Delta \rho} \right| \right),$$

$$\Theta = \min_{\theta} g(\theta), \quad \text{Hyper } T = \Pi(F_{\text{sect}}) = G(\phi, r),$$

$$\text{Hyper } P = \left| \text{Row3D} + 1 - 2 \cdot \text{Arg max}_r (G(\omega, \phi, r)) \right|,$$

$$\text{Hyper } \Omega = \max_{\phi} G(\omega, \phi), \quad \text{Hyper } \Theta = \max_{\omega} G(\omega),$$

где Row3D – количество ненулевых элементов в глубинных строках гипертрейс-матрицы 3ТМ.

Для оценки уровня колебания значений вычисляемого признака использовался коэффициент колебания его средней относительной погрешности, который рассчитывается по формуле:

$$\epsilon = \frac{\sum_{j=1}^n \left| \text{Res}_j - \frac{\sum_{i=1}^n |\text{Res}_i|}{n} \right|}{\sum_{i=1}^n |\text{Res}_i|},$$

где Res_i – значение признака в i -м эксперименте, n – количество экспериментов.

Зависимость погрешности вычислений признака Res от параметра Δr демонстрирует рис. 4.

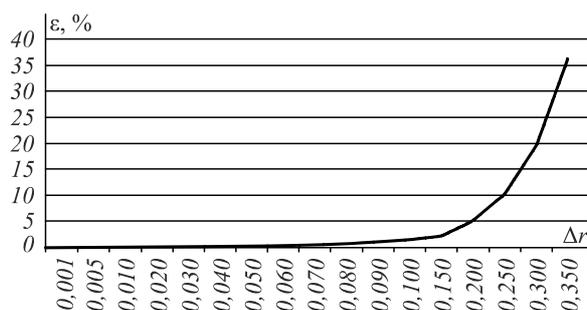


Рис. 4. Зависимость погрешности вычисления признака от параметра Δr

Из графика 6 видно, что при $\Delta r < 0,06$ точность вычисления признака стабильна. Так как шаг Δr и время работы распознающего алгоритма обратно пропорциональны, то оптимальное значение $\Delta r = 0,06$, что составляет менее 10% от параметров размера исходного 3D-изображения.

Зависимость погрешности вычислений признака от параметра сканирования $N\varphi$ показана на рис. 5.

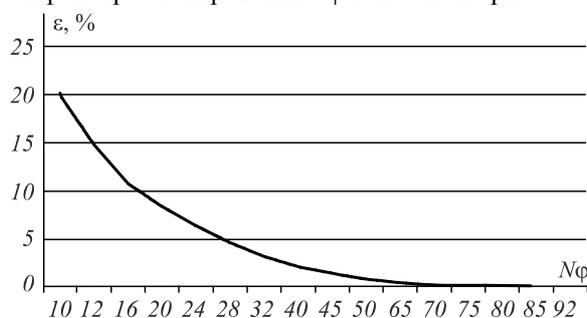


Рис. 5. Зависимость погрешности вычисления признака от параметра $N\varphi$

Как видно из представленного выше графика, достаточно уже 64 узла на опорной сетке на сфере, чтобы погрешность вычисления признака ε была низкой (менее 1%).

Зависимость погрешности вычислений признака Res от параметра сканирования Δr демонстрирует рис. 6.

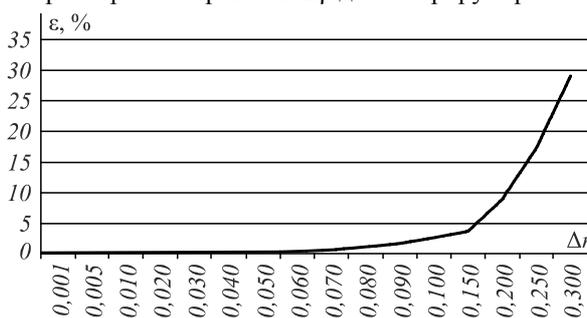


Рис. 6. Зависимость погрешности вычисления признака от параметра Δr

Согласно графику на рис. 6, приемлемым значением с точки зрения соотношения качества инвариантности признака и времени работы алгоритма является значение $\Delta r = 0,06$ (согласно аналогичным рас-суждениям, приведённым для Δr).

Зависимость погрешности вычислений признака Res от параметра сканирования $N\theta$ демонстрирует рис. 7.

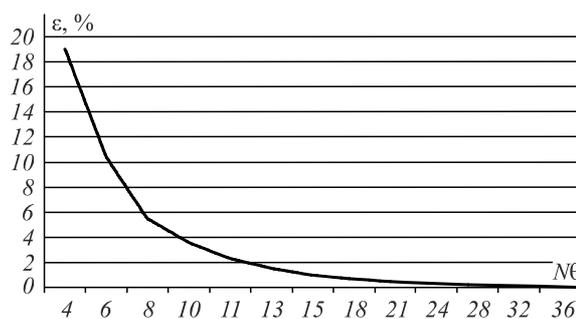


Рис. 7. Зависимость погрешности вычисления признака от параметра $N\theta$

Так как шаг $N\theta$ и время работы распознающего алгоритма прямо пропорциональны, то согласно представленному выше графику оптимальное значение $N\theta$ достигается при значении 18, то есть $N\theta = 18$.

Стоит отметить, что для разных видов признаков данные графики будут отличаться друг от друга, но основные направления роста погрешности вычисления признаков сохранятся. Цель данного эксперимента состояла в наглядной демонстрации основных свойств зависимости погрешности вычисления признака от различных параметров сканирования и проверке адекватности разработанного математического аппарата.

Благодарности

Работа выполнена при поддержке гранта РФФИ (проект №15-07-04484).

Литература

1. **Kiy, K.I.** Segmentation and detection of contrast objects and their application in robot navigation / K.I. Kiy // Pattern Recognition and Image Analysis. – 2015. – Vol. 25, Issue 2. – P. 338-346. – DOI: 10.1134/S1054661815020145.
2. **Wang, C.** VFM: visual feedback model for robust object recognition / C. Wang, K.-Q. Huang // Journal of Computer Science and Technology. – 2015. – Vol. 30, Issue 2. – P. 325-339. – DOI: 10.1007/s11390-015-1526-1.
3. **Гайдель А.В.** Исследование текстурных признаков для диагностики заболеваний костной ткани по рентгеновским изображениям / А.В. Гайдель, С.С. Первушкин // Компьютерная оптика. – 2013. – Т. 37, № 1. – С. 113-119.
4. **Гайдель, А.В.** Возможности текстурного анализа компьютерных томограмм в диагностике хронической обструктивной болезни / А.В. Гайдель, П.М. Зельтер, А.В. Капишников, А.Г. Храмов // Компьютерная оптика. – 2014. – Т. 38, № 4. – С. 843-850.
5. **Федотов, Н.Г.** 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. – С. 41-53.
6. **Rakhmanov, E.A.** Minimal discrete energy on the sphere / E.A. Rakhmanov, E.B. Saff, Y.M. Zhou // Mathematical Research Letters. – 1994. – Vol. 1. – No 6. – P. 647-662. – DOI: 10.4310/MRL.1994.v1.n6.a3.
7. **Lovisolo, L.** Uniform distribution of points on a hypersphere with applications to vector bit-plane encoding / L. Lovisolo, L.E.A.B. da Silva // IEE Proceedings - Vision, Image and Signal Processing. – 2001. – Vol. 148, Issue 3. – P. 187-193 - DOI: 10.1049/ip-vis:20010361.

8. **Федотов, Н.Г.** Теория признаков распознавания образов на основе стохастической геометрии и функционального анализа / Н.Г. Федотов. – М.: Физматлит, 2010. – 304 с. – ISBN: 978-5-9221-0996-3.
9. **Fedotov, N.G.** Random scanning for speedier systems of pattern recognition based on stochastic geometry methods / N.G. Fedotov, L.A. Shul'ga, A.V. Moiseev // Pattern Recognition and Image Analysis. – 2005. – Vol. 15, Issue 2. – P. 387-388.
10. Princeton Shape Benchmark [Electronical Resource]. - URL: <http://shape.cs.princeton.edu/benchmark/> (date request 10.10.2017).

Сведения об авторах

Федотов Николай Гаврилович, 1945 года рождения. В 1967 году окончил с отличием Пензенский политехнический институт (ныне Пензенский государственный университет) по специальности «Вычислительная техника». По окончании аспирантуры при том же институте в 1973 году в ученом совете Рязанского радиотехнического института защитил диссертацию на соискание ученой степени кандидата технических наук. В 1990 году в Московском институте электроники и математики (МИЭМ) защитил диссертацию на соискание ученой степени доктора технических наук. Имеет ученые звания доцента (1977) и профессора (1992). В Пензенском государственном университете работает с 1971 года. С 1996 года заведует кафедрой экономической кибернетики. Область научных интересов: распознавание образов, стохастическая геометрия, признаки распознавания, анализ изображений, функциональный анализ, искусственный интеллект, экономическая кибернетика. E-mail: fedotov@pnzgu.ru.

Сёмов Алексей Александрович, 1990 года рождения, в 2012 году окончил Пензенский государственный университет по специальности 080116 «Математические методы в экономике». В 2015 году защитил диссертацию на соискание ученой степени кандидата технических наук. Работает аналитиком в ООО «КомХэлф». Область научных интересов: математическое моделирование, математическая статистика, теория вероятностей, обработка изображений, распознавание образов. E-mail: mathematik_aleksey@mail.ru.

Моисеев Александр Владимирович, 1972 года рождения, в 1994 году окончил Пензенский государственный педагогический университет им. В.Г. Белинского по специальности «Математика». В 1997 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук. Имеет ученое звание доцент (2001). Работает доцентом кафедры экономической кибернетики Пензенского государственного университета. Область научных интересов: математический анализ, краевые задачи, теория вероятностей, математическая статистика, математическое моделирование экономических процессов, распознавание образов. E-mail: moigus@mail.ru.

ГРПТИ: 28.23.15

Поступила в редакцию 9 августа 2017 г. Окончательный вариант – 16 октября 2017 г.

THEORETICAL FOUNDATIONS OF HYPERTRACE-TRANSFORM: SCANNING TECHNIQUES, MATHEMATICAL APPARATUS AND EXPERIMENTAL VERIFICATION

N.G. Fedotov¹, A.A. Syemov², A.V. Moiseev¹

¹ Penza State University, Penza, Russia,

² LLC «KomHelf», Penza, Russia

Abstract

We consistently describe the theoretical basis of a new geometric method of analysis and recognition of three-dimensional (3D) images. The description of a scanning technique for forming a hypertrace transform and its mathematical model are given. This method, unlike the existing ones, enables 3D images to be analyzed directly from their 3D shape, without first simplifying them or constructing plane projections. We substantiate the selection of a particular scanning tool and the need to construct a reference spherical grid to address the problem of the rotational invariance of the 3D image recognition. A mathematical apparatus of the stochastic realization of the scanning technique based on stochastic geometry and functional analysis is developed. We introduce a new mathematical tool for 3D image analysis – a hypertrex matrix that allows spatial objects of complex shape and structure to be recognized by constructing a single mathematical model of the 3D image. We describe a new type of 3D image features that have an analytic structure – hypertryplet features, whose analytical structure makes possible an automatic generation of a large number of features with predetermined properties. Results of the experimental verification are presented, demonstrating the accurate calculation of features for 3D image recognition and proving the adequacy of the developed mathematical apparatus.

Keywords: recognition of 3D images, geometric hypertrace-transform, grid of parallel planes, stochastic scanning, analytical structure of the feature, hypertrace matrix, and invariant recognition.

Citation: Fedotov NG, Syemov AA, Moiseev AV. Theoretical foundations of hypertrace-transform: scanning techniques, mathematical apparatus and experimental verification. *Computer Optics* 2018; 42(2): 273-282. DOI: 10.18287/2412-6179-2018-42-2-273-282.

Acknowledgements: The work was partially funded by the Russian Foundation for Basic Research under grant No. 15-07-04484.

References

- | | |
|--|---|
| <p>[1] Kiy KI. Segmentation and detection of contrast objects and their application in robot navigation. <i>Pattern Recognition and Image Analysis</i> 2015; 25(2): 338-346. DOI: 10.1134/S1054661815020145.</p> <p>[2] Wang C, Huang K-Q. VFM: visual feedback model for robust object recognition. <i>Journal of Computer Science and Technology</i> 2015; 30(2): 325-339. DOI: 10.1007/s11390-015-1526-1.</p> <p>[3] Gaidel AV, Pervushkin SS. Research of the textural features for the bony tissue diseases diagnostics using the roentgenograms [In Russian]. <i>Computer Optics</i> 2013; 37(1): 113-119.</p> <p>[4] Gaidel AV, Zelter PM, Kapishnikov AV, Khramov AG. Computed tomography texture analysis capabilities in diagnosing a chronic obstructive pulmonary disease [In Russian]. <i>Computer Optics</i> 2014; 38(4): 843-850.</p> <p>[5] Fedotov NG, Semov AA, Moiseev AV. 3D-trace-conversion: scanning modes, stochastic implementation features,</p> | <p>methods of computational speedup [In Russian]. University proceedings. Volga region. Technical sciences 2014; 3: 41-53.</p> <p>[6] Rakhmanov EA, Saff EB, Zhou YM. Minimal discrete energy on the sphere. <i>Math Res Lett</i> 1994; 1(6): 647-662. DOI: 10.4310/MRL.1994.v1.n6.a3.</p> <p>[7] Lovisolo L, da Silva LEAB. Uniform distribution of points on a hyper-sphere with applications to vector bit-plane encoding. <i>IEE Proc – Vis Image Signal Process</i> 2001; 148(3): 187-193. DOI: 10.1049/ip-vis:20010361.</p> <p>[8] Fedotov NG. The theory of pattern recognition features based on stochastic geometry and functional analysis [In Russian]. Moscow: “Fizmatlit” Publisher; 2010. ISBN: 978-5-9221-0996-3.</p> <p>[9] Fedotov NG, Shul'ga LA, Moiseev AV. Random scanning for speedier systems of pattern recognition based on stochastic geometry methods. <i>Pattern Recognition and Image Analysis</i> 2005; 15(2): 387-388.</p> <p>[10] Princeton Shape Benchmark. Source: http://shape.cs.princeton.edu/benchmark/.</p> |
|--|---|

Author's information

Nikolay Gavrilovich Fedotov (b. 1945) graduated with honors from Penza Polytechnic Institute (now the Penza State University) in 1967 on a specialty “Computer Science”. After graduate school at the same institute in 1973, the Scientific Council of the Ryazan Radio Engineering Institute defended his thesis for the degree of candidate of technical sciences. In 1990, at the Moscow Institute of Electronics and Mathematics (MIEM) defended his thesis for the degree of Doctor of Technical Sciences. He has the rank of assistant professor (1977) and professor (1992). Since 1996, Head of Economic Cybernetics department. His research interests include pattern recognition, stochastic geometry, signs of recognition, image analysis, functional analysis, artificial intelligence, economic cybernetics. E-mail: fedotov@pnzgu.ru.

Aleksey Alexandrovich Syemov (b. 1990) graduated from the Penza State University in 2012, majoring in 080116 “Mathematical Methods in Economics”. In 2015 he defended his thesis for the degree of candidate of technical sciences. He worked as an analyst in the company “KomHelf”. Research interests: mathematical modeling, mathematical statistics, probability theory, image processing, pattern recognition. E-mail: mathematik_aleksey@mail.ru.

Alexander Vladimirovich Moiseev (b. 1972) graduated from Penza State Pedagogical Institute in 1994, specialty “Mathematics”. In 1997 he defended his thesis for the degree of candidate of physical and mathematical sciences. He holds the title of Associate Professor (2001). It works assistant professor of “Economic Cybernetics” Penza State University. His research interests include mathematical analysis, boundary value problems, probability theory, mathematical statistics, mathematical modeling of economic processes, pattern recognition. E-mail: moigus@mail.ru.

Received August 9, 2017. The final version – October 16, 2017.