

Исследование быстродействия распознавания трехмерного изображения методами стохастической геометрии в зависимости от количества опорных точек на сфере*

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

fedotov@pnzgu.ru; matematik_aleksey@mail.ru; moigus@mail.ru

¹Пензенский государственный университет, Россия, г. Пенза, ул. Красная, 40

²ООО «Комэрф», Россия, г. Пенза, ул. Гагарина, 11а

Рассматривается новый подход к распознаванию трехмерных (3D) изображений, дающий инвариантное описание объекта при любой его пространственной ориентации. Метод позволяет параллельно с распознаванием производить анализ исходного изображения, а также конструировать признаки с заданными свойствами. Предлагаемый подход основан на методах стохастической геометрии и функционального анализа. Разработанное авторами гипертрейс-преобразование создает 3D трейс-образ исходного пространственного объекта, который является удобным инструментом для анализа 3D изображения в отличие от математического аппарата других методов. Стохастическое сканирование со случайными параметрами более эффективно по сравнению с детерминированным сканированием с точки зрения соотношения «надежность–быстродействие» распознавания 3D изображений. Приведены результаты экспериментов, показывающие теоретическую и практическую значимость, а также эффективность предлагаемого подхода. Анализируется быстродействие распознавания 3D объектов в зависимости от количества опорных точек на сфере с использованием различных видов сканирования. Предлагаются возможные пути ускорения работы распознающей системы.

Ключевые слова: 3D распознавание образов; гипертрейс-преобразование; производительность распознающей системы; опорная сетка на сфере; инвариантное описание; стохастическое сканирование

DOI: 10.21469/22233792.3.3.01

1 Введение

Устойчивой тенденцией научно-технического прогресса является увеличение числа людей, занятых обработкой информации. Одной из важнейших проблем, возникающих при создании интеллектуальных информационных систем, является автоматизация процесса распознавания образов [1–3]. Успехи в ее реализации создают решающие предпосылки для построения интеллектуальных систем, важнейшими областями применения которых являются робототехника (машинное зрение), наблюдение Земли из космоса, аэрокосмические исследования.

На пути построения интеллектуальной системы распознавания образов необходимо решить множество сложных задач [4]. Важно научиться придавать системам способность понимать и анализировать всю 3D сцену (в том числе положения, ориентации, текстуры и освещения пространственного объекта), а также сегментировать, интерпретировать и классифицировать данное 3D изображение.

*Работа выполнена при финансовой поддержке РФФИ, проект № 15-07-04484.

Одним из основных подходов к анализу распознавания 3D изображений являются методы, требующие предварительной нормализации положения 3D объекта. При этом используются разнообразные техники сканирования. Например, в работах [5–7] можно найти техники сканирования, которые используют плавающее окно 3D детектора обнаружения объекта и его границ.

Для рассмотренных и других аналогичных методов данного подхода необходимо определение точного позиционирования 3D моделей и последующая нормализация их положения. Однако нормализация по повороту за счет пространственной ориентации во многих случаях не является устойчивой из-за невозможности однозначного определения множества главных осей. Наличие искажений может сильно повлиять на определение точного коэффициента масштабирования и вектора переноса, в результате чего эффективность распознавания может быть значительно снижена.

Другим подходом к анализу распознавания 3D изображений считаются методы инвариантного описания пространственного объекта, которые анализируют геометрию поверхности объекта, особенности его локальных участков и/или кривизны его отдельных частей. К данной группе методов можно отнести работу [8]. Собственные функции оператора Лапласа–Бельтрами дают набор вещественных функций, которые предоставляют информацию о структуре и морфологии формы. Анизотропный оператор Лапласа–Бельтрами вводит свойство изменчивости в направлениях главной кривизны, благодаря чему построенные признаки обладают одинаковыми изотропными свойствами вне зависимости от ориентации 3D модели. Данный метод позволяет обнаруживать повторяющиеся регионы на поверхности тела.

Существуют также и другие методы, аналогичные выше описанным, которые концентрируются на анализе поверхности 3D объекта, но при этом используются другие операторы при конструировании дескриптора признаков. Так, дескрипторы с использованием 3D дискретного преобразования косинуса, применяемые для поиска 3D объектов в базах данных (аналогично двумерному (2D) аналогу, используемому в алгоритме сжатия JPEG), описаны в работе [9]; дескрипторы с использованием анизотропной диффузии тензорных полей для анализа геометрии сгибов и деформаций анатомических органов и частей тела человека — в работе [10]; с использованием семейства параметрических спектральных дескрипторов Лапласа для анализа и распознавания 3D человеческих фигур — в работе [11].

К очевидным недостаткам методов данной группы можно отнести тот факт, что признаки не имеют явной геометрической интерпретации и указывают лишь обобщенные свойства поверхности 3D объекта. Поэтому данные методы не эффективны при решении задачи классификации базы 3D изображений, которая требует построения признаков, более чувствительных к различению формы. Кроме того, выделение однородных участков и определение границ на поверхности объекта является сложной задачей, неправильное решение которой будет приводить к смещению вычисляемого признака.

Третьим подходом к анализу распознаванию 3D изображений являются методы инвариантного описания пространственного объекта, которые анализируют непосредственно сразу объемную 3D форму объекта как одно целое. К данной группе методов можно отнести работу [12]. Описанный в ней метод заключается в извлечении инвариантного к группе движений дескриптора признаков с использованием многомерной регрессионной линейной модели, примененной к описанию пространственного объекта, заданного в виде облака точек.

Достоинством данного метода является возможность построения признаков, инвариантных не только к группе движений, но и к аффинным преобразованиям. Однако его

существенным минусом является тот факт, что инвариантные признаки описываются не на основе особенностей 3D формы объекта, а на основе определенных статистических связей между точками объектов, учитывающих общие свойства пространственной формы. Так, при отсутствии ярко выраженных таких связей и ависимостей признаки будут слабо информативными. Так как при оценке 3D формы объекта учитываются только его вершины, а грани игнорируются, то объекты разного класса могут быть идентифицированы как однородные. Кроме того, отсутствует свойство инвариантности конструируемых признаков к масштабированию 3D изображения.

В настоящей статье предлагается новый подход к конструированию признаков 3D изображения на основе стохастической геометрии, дающий инвариантное описание объекта при любой его пространственной ориентации. Благодаря структуре признака в виде композиции функционалов возможно построение большого числа признаков 3D изображений в режиме автоматической компьютерной генерации. Стохастическое сканирование со случайными параметрами более эффективно по сравнению с детерминированным сканированием с точки зрения соотношения «надежность–быстродействие» распознавания 3D изображений. При этом в статье анализируется задача оценки быстродействия распознавания 3D изображения методами стохастической геометрии в зависимости от количества опорных точек на сфере. В конце статьи предлагаются возможные дальнейшие пути по ускорению работы распознающей системы.

2 Математическая постановка проблемы

Пусть имеется множество объектов x , относительно которых нужно произвести классификацию. Объекты задаются значениями некоторых признаков f_i , $i = 1, \dots, n$, наборы которых одинаковы для всех объектов. Таким образом, описание Info каждого объекта будет иметь вид:

$$\text{Info}(x) = (f_1(x), f_2(x), \dots, f_n(x)).$$

Множество допускает разбиение на подмножества K_1, \dots, K_n , называемые классами: $X = \bigcup_{j=1}^m K_j$. Данное разбиение осуществляется на основе некоторой априорной информации. Имеется описание классов $\text{Info}(K_j)$, $j = 1, \dots, m$, например как усредненная характеристика признаков объектов внутри соответствующего класса. Задача распознавания состоит в том, чтобы для каждого объекта x , о котором неизвестно, к какому из классов K_1, \dots, K_n он принадлежит, по известной информации $\text{Info}(K)$ и вычисленной в ходе сканирования объекта $\text{Info}(x)$ установить значение предиката $P_j = (x \in K_j)$.

Решающая процедура определения класса изображения одинакова как для стохастического распознавания, так и для детерминированного. Поэтому эффективность распознавания 3D изображения в данном случае будет зависеть, в первую очередь, от способа сканирования, результатами которого является вектор дескриптор признаков, описывающий исходное 3D изображение. В связи с этим актуальным становится решение задачи анализа быстродействия формирования признаков 3D изображения в зависимости от заданной точности вычисления признака и метода сканирования.

3 Гипертрейс-преобразование

Прежде чем переходить к описанию положения сканирующих элементов и исследованию свойства быстродействия распознавания изображения, необходимо сначала ввести их описание. Пусть F — исходная 3D модель. Определим плоскость

$$B(\eta(\omega, \varphi), r) = \{x | x^T \cdot \eta(\omega, \varphi) = r\}$$

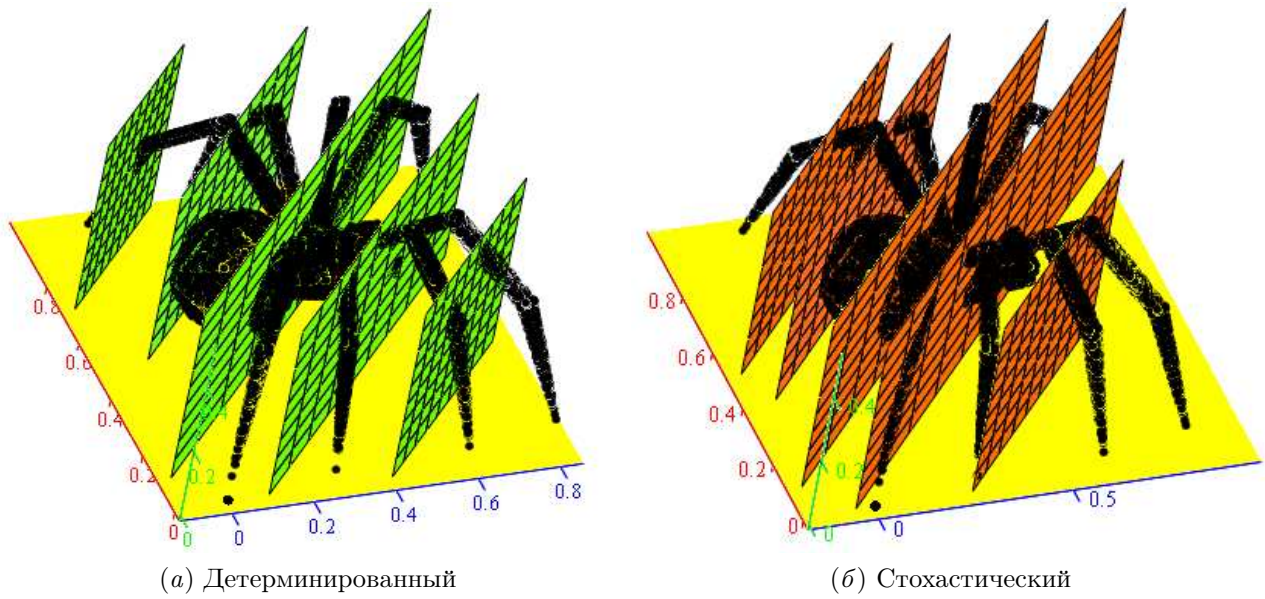


Рис. 1 Типы сканирования 3D изображения объекта

как касательную к сфере с центром в начале координат и радиусом r , где $\eta(\omega, \varphi)$ — единичный вектор в R^3 :

$$\eta(\omega, \varphi) = [\cos \omega \cdot \sin \varphi; \sin \omega \cdot \sin \varphi; \cos \varphi];$$

ω и φ — сферические углы.

Сканирование 3D изображения производится сеткой параллельных плоскостей с расстоянием Δr между плоскостями и заданными углами ω и φ (рис. 1, а) [13]. Взаимное положение изображения F и каждой сканирующей плоскости $B(\eta(\omega, \varphi), r)$ характеризуется числом $G = \text{НурегТ}(F \cap B(\eta(\omega, \varphi), r))$, описывающим любой заданный признак 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 для завершения полного обзора объекта со всех сторон.

Результат вычислений НурегТ функционала зависит от трех параметров плоскости: (r, ω, φ) , поэтому результат сканирования удобно представить в виде 3D матрицы ЗТМ, у которой ось 0φ направлена вертикально, ось 0ω — горизонтально, ось $0r$ — вглубь (рис. 2) [14]. При стохастическом сканировании задаются вероятностные распределения (см. рис. 1, б):

$$\begin{aligned} M_1 &= \{p_1(\omega), \dots, p_m(\omega)\}; \\ M_2 &= \{q_1(\varphi), \dots, q_n(\varphi)\}; \\ M_3 &= \{s_1(r), \dots, s_l(r)\}, \end{aligned}$$

где

$$\sum_i p_i(\omega) = 1; \quad \sum_j q_j(\varphi) = 1; \quad \sum_k s_k(r) = 1.$$

Элемент матрицы ЗТМ показывает значение выбранного признака сечения F_{sect} . Другими словами, тройке $(\omega_i, \varphi_j, r_k)$ соответствует элемент матрицы ЗТМ с номером (i, j, k)

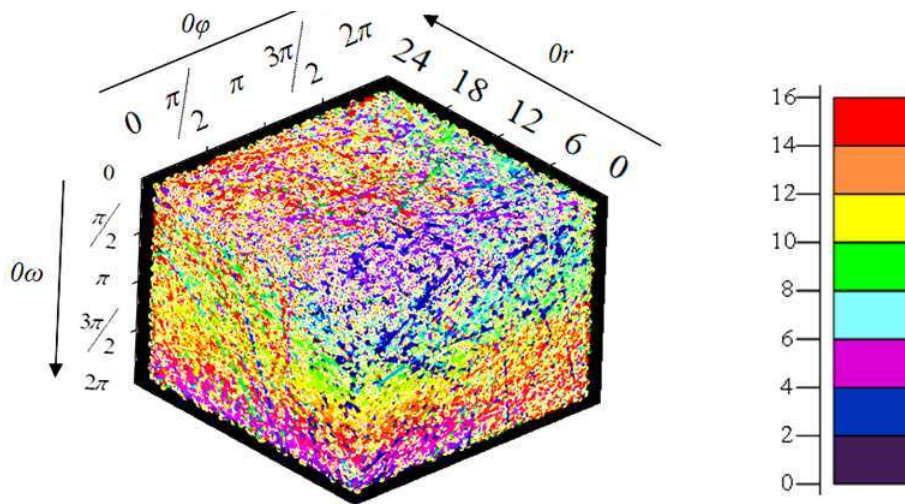


Рис. 2 Пример графического представления гипертрейс-матрицы ЗТМ

и значением $\Pi(F_{\text{sect}})$, которое характеризует информативный признак сечения 3D изображения F плоскостью $B(\eta(\omega_i, \varphi_j), r_k)$:

$$F_{\text{sect}} = F \cap B(\eta(\omega_i, \varphi_j), r_k).$$

Выбирая параметры сканирования случайным образом на основе заданных распределений в виде множеств M_1, M_2 и M_3 , получаем матрицу случайных величин ЗТМ, где $\text{ЗТМ}_{i,j,k} = \text{HyperT}(F \cap B(\eta(\omega_i, \varphi_j), r_k))$ с вероятностью $p_i q_j s_k$ и $\text{ЗТМ}_{i,j,k} = 0$ с вероятностью $1 - p_i q_j s_k$.

После заполнения гипертрейс-матрицы последовательно обрабатываются ее глубинные, вертикальные и горизонтальные строки с помощью функционалов HyperP , $\text{Hyper}\Omega$ и $\text{Hyper}\Theta$ соответственно. Например,

$$\begin{aligned} \text{HyperP} &= \frac{\sum_r G(\omega, \varphi, r)}{\max_r G(\omega, \varphi, r)}; \\ \text{Hyper}\Omega &= \max_{\varphi} G(\omega, \varphi) - \min_{\varphi} G(\omega, \varphi); \\ \text{Hyper}\Theta &= \frac{\min_{\omega} G(\omega)}{\max_{\omega} G(\omega)}. \end{aligned}$$

В результате получается число $\text{Res}(F)$ — гипертриpletный признак 3D изображения F в виде композиции четырех функционалов [15]:

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

Для нахождения признака сечения используется трейс-преобразование [16]. Так, сканирование получаемых сечений F_{sect} осуществляется решеткой параллельных прямых $l(\theta, \rho)$ с расстоянием $\Delta\rho$ между линиями, где ρ и θ — полярные координаты прямой в плоскости сечения (рис. 3, а). Взаимное положение 2D изображения F_{sect} и каждой сканирующей линии $l(\theta, \rho)$ характеризуется числом $g = T(F_{\text{sect}} \cap l(\theta, \rho))$, описывающим любой заданный признак пересечений прямой с изображением.

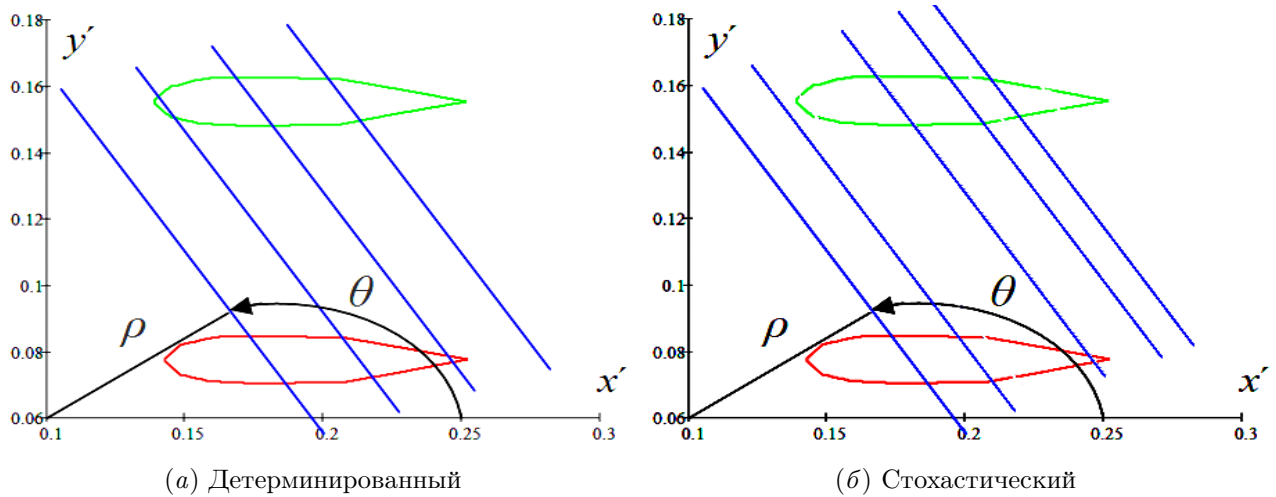


Рис. 3 Типы сканирования 2D изображения сечения

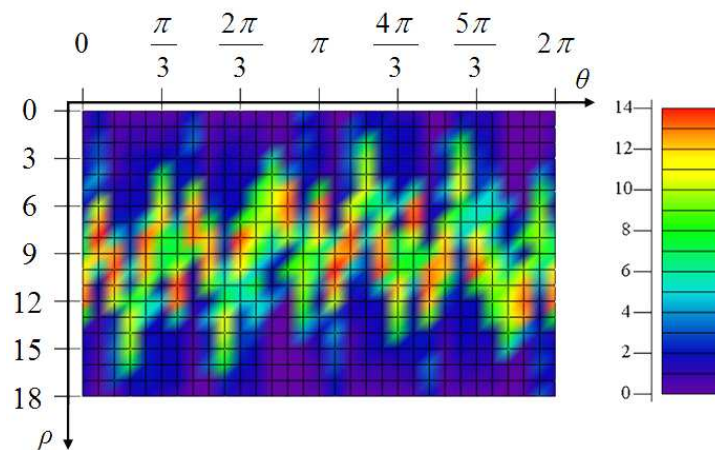


Рис. 4 Пример графического представления трейс-матрицы ТМ

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 2D трейс-матрицы ТМ, у которой ось 0θ направлена горизонтально, а ось 0ρ — вертикально (рис. 4). При стохастическом сканировании задаются вероятностные распределения (см. рис. 3, б):

$$N_1 = \{x_1(\theta), \dots, x_t(\theta)\};$$

$$N_2 = \{y_1(\rho), \dots, y_w(\rho)\},$$

где

$$\sum_i x_i(\theta) = 1; \quad \sum_j y_j(\rho) = 1.$$

Элемент матрицы показывает значение выбранного признака пересечения прямой изображения сечения F_{sect} . Другими словами, паре (θ_i, ρ_j) соответствует элемент матрицы ТМ с номером (i, j) и значением $T(F_{\text{sect}})$, которое характеризует признак пересечения

прямой и 2D изображения сечения: $T(F_{\text{sect}}) = F_{\text{sect}} \cap l(\theta_i, \rho_j)$. Выбирая параметры сканирования случайным образом на основе заданных распределений в виде множеств N_1 и N_2 , получаем матрицу случайных величин ТМ, где $\text{ТМ}_{i,j,k} = T(F_{\text{sect}} \cap l(\theta_i, \rho_j))$ с вероятностью $x_i y_j$ и $\text{ТМ}_{i,j,k} = 0$ с вероятностью $1 - x_i y_j$.

После заполнения матрицы ТМ последовательно обрабатываются ее вертикальные и горизонтальные строки с помощью функционалов P и Θ соответственно. Например, $P = \max_{\rho} g(\theta, \rho)$ и $\Theta = \min_{\theta} P(\theta)$. В результате получается число $\Pi(F_{\text{sect}})$ — триплетный признак 2D изображения сечения F_{sect} в виде композиции трех функционалов [17]:

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

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

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

Таким образом, благодаря структуре признаков в виде композиции функционалов, входящих в $\Pi(F_{\text{sect}})$ и $\text{Res}(F)$, возможно получение большого числа признаков, свойства которых можно регулировать [18].

Стандартная решающая процедура на основе триплетных признаков строится следующим образом [19].

Рассмотрим множество $M = \bigcup_{i=1}^m C_i$, состоящее из m подмножеств (классов) C_i , при этом в подмножестве C_i содержится h_i элементов (изображений). Выберем из данного множества M подмножество $M' \subset M$, мощность которого равна $\sum_{i=1}^m h_i/2$ для обучения системы, т.е. $M' = \bigcup_{i=1}^m A_i$. Оставшиеся подмножества B_i будут нужны для испытания обученной системы, а также контроля ее качества.

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 вида признака для всех изображений множества A_i :

$$\mu_k^{A_i} = \frac{2}{h_i} \sum_{s=1}^{h_i} \text{Res}_k^{A_i(s)}.$$

Среднеквадратическое отклонение k -го признака по множеству A_i :

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

Обозначим через t тестовое 3D изображение из какого-либо подмножества B_i . Тогда его k -й признак будет равен Res_k^t . Расстояние между тестовым 3D изображением и i -м классом A_i с учетом весовых значений определяется следующим образом:

$$d(t, A_i) = \sum_k \frac{|\text{Res}_k^t - \mu_k^{A_i}|}{\sigma_k^{A_i}}.$$

Распознающая система тестовое изображение t относит к классу A_j , если $d(t, A_j) = \min_i d(t, A_i)$.

4 Описание опорной сетки на сфере

Каждая сканирующая плоскость $B(\eta(\omega, \varphi), r)$ определяется парой углов ω и φ . Учитывая тот факт, что в сканирующей сетке все плоскости параллельны, пара углов ω и φ однозначно определяет сканирующую сетку параллельных плоскостей. Стандартный перебор всех пар углов $(\omega; \varphi)$ на единичной сфере образует множество точек — опорную сетку на сфере. Плотность точек этой сетки на сфере характеризует плотность всех сканирующих сеток плоскостей в пространстве. На рис. 5 представлены примеры стохастической и равномерной сетки на сфере.

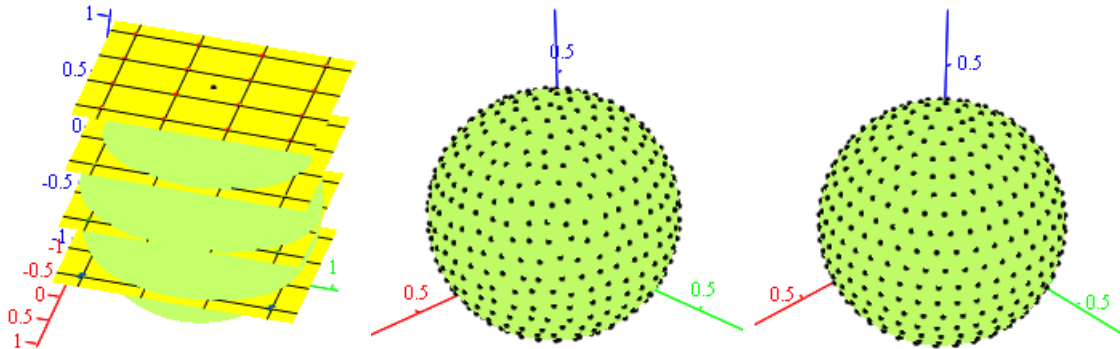


Рис. 5 Построение равномерной сетки на сфере детерминированным (**Saff_fix**) и стохастическим (**Saff_rnd**) способами для $N = 500$

Если при повороте сферы вокруг своего центра опорная сетка на ней перейдет сама в себя, то и все сканирующие плоскости под разными углами наклона перейдут сами в себя. Это означает, что получаемые сечения и вычисляемые по ним признаки не изменятся, что будет говорить об инвариантности распознавания к повороту 3D изображения в пространстве.

Таким образом, необходимо достичь равномерной плотности узлов сетки на сфере, чтобы ошибка наложения сетки при поворотах друг на друга была как можно меньше. Также и увеличение количества точек опорной сетки уменьшает колебание вычисляемого значения признака.

Опишем по одному способу построения равномерной опорной сетки детерминированного и стохастического типов.

При детерминированном способе генерирования равномерного распределения точек на сфере сетка будет иметь строго фиксированный характер, заданный правилом ее построения. Каждый способ имеет свои плюсы и минусы, но основным недостатком является возможность построения сетки не для любого количества узлов.

Равномерное распределение точек на сфере можно получить при помощи спирали на ее поверхности (алгоритм **Saff**) [20]. Спираль приближается последовательностью точек, расстояние между которыми приблизительно равно расстоянию между витками. Каждая точка на сфере $[\cos \omega \cdot \sin \varphi, \sin \omega \cdot \sin \varphi, \cos \varphi]$ будет задаваться при помощи сферических координат ω и φ , которые будут определяться следующим образом. Пусть для каждого k будут последовательно вычислены следующие выражения:

$$h_k = -1 + \frac{2(k-1)}{N-1}, \quad \omega_k = \arccos(h_k), \quad 1 \leq k \leq N;$$

$$\varphi_1 = \varphi_N = 0, \quad \varphi_k = \left(\varphi_{k-1} + \frac{3,6}{\sqrt{N}} \frac{1}{\sqrt{1-h_k^2}} \right) \equiv \text{mod } 2\pi, \quad 2 \leq k \leq N-1;$$

$$\text{Grid}_k = [\cos \omega_k \cdot \sin \varphi_k, \sin \omega_k \cdot \sin \varphi_k, \cos \varphi_k].$$

Тогда спираль может быть приближена любым количеством точек, расположенных по спирали. Расстояние между витками приблизительно равно расстоянию между узлами сетки. Построенная сетка будет обладать очень низкой потенциальной энергией, что говорит о ее равномерности. Преимуществом данного метода является возможность построения сетки для любого количества точек. Недостатком данного метода является построение точек в области полюса, где наблюдается отклонение от истинного направления описывающей спирали.

При стохастическом способе генерирования равномерного распределения точек на сфере используются генераторы псевдослучайных чисел, которые позволяют получать равномерные сетки на сфере для любого количества точек. Авторы решили модифицировать алгоритм **Saff** 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. Данную процедуру можно описать следующими формулами:

$$\text{Grid}'_k = [\cos \omega_k \cdot \sin \varphi_k, \sin \omega_k \cdot \sin \varphi_k, \cos \varphi_k], \quad 1 \leq k \leq N;$$

$$\text{Grid}''_k = \left[\text{Grid}'_{k,1} + \text{rnd} \left(\frac{1}{\sqrt{N}} \right); \text{Grid}'_{k,2} + \text{rnd} \left(\frac{1}{\sqrt{N}} \right); \text{Grid}'_{k,3} + \text{rnd} \left(\frac{1}{\sqrt{N}} \right) \right];$$

$$\text{Grid}_{k,i} = \frac{\text{Grid}''_{k,i}}{\sqrt{(\text{Grid}''_{k,1})^2 + (\text{Grid}''_{k,2})^2 + (\text{Grid}''_{k,3})^2}}, \quad i = 1, 2, 3,$$

где N — общее количество точек на сфере; $\text{rnd}(x)$ — функция, выдающая случайные равномерные числа в диапазоне от 0 до x .

5 Исследование быстродействия распознавания трехмерного изображения в зависимости от количества опорных точек на сфере

Учитывая тот факт, что сканирование сетками плоскостей под разными углами обзора исходного 3D изображения достаточно емко по времени, актуальным становится выявление путей ускорения работы распознающей системы. Так как количество сеток сканирующих плоскостей под разными углами обзора исходного пространственного объекта зависит от количества узлов сетки на сфере, целесообразной представляется задача анализа быстродействия распознавания 3D изображения в зависимости от количества опорных точек на сфере.

Для этого был осуществлен эксперимент, который состоял из двух частей. Первая часть ставила своей целью исследовать зависимость точности вычисления признака от количества узлов опорной сетки на сфере в виде фиксированной развертки с использованием детерминированного сканирования сеткой параллельных плоскостей и фиксированной опорной сетки на сфере. Вторая же часть была той же самой с той лишь разницей, что использовалась стохастическая сетка сканирования и стохастическая опорная сетка на сфере.

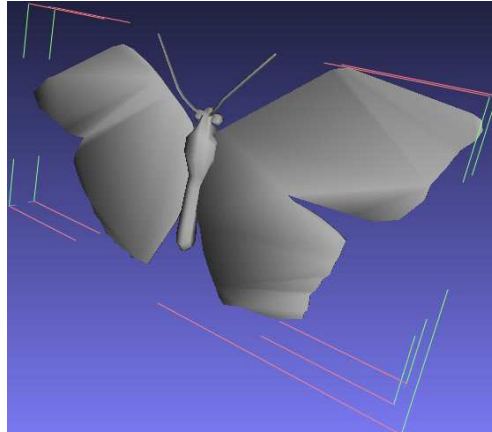


Рис. 6 Трехмерная модель бабочки m11 базы The Princeton Shape Benchmark

Достаточно сложно оценить точно произвольный признак аналитически или численно для произвольно взятого пространственного объекта. Поэтому в качестве признака была взята максимальная длина отрезка, который может быть помещен внутри пространственного объекта:

$$L(F) = \text{Hyper}\Theta \circ \text{Hyper}\Omega \circ \text{HyperP} \circ \text{HyperT}(\Theta \circ \text{P} \circ \text{T}).$$

Здесь $\text{T} = \max_t f(\theta, \rho, t)$; $\text{P} = \max_\rho \text{T}(\theta, \rho)$; $\Theta = \max_\theta \text{P}(\theta)$; $\text{HyperT} = G(\omega, \varphi, r)$; $\text{HyperP} = \text{Row3D} \cdot \Delta r$, где Row3D — количество ненулевых элементов в глубинных строках матрицы ЗТМ; $\text{Hyper}\Omega = \max_\varphi G(\omega, \varphi)$; $\text{Hyper}\Theta = \max_\omega G(\omega)$.

Две максимально удаленные друг от друга точки поверхности пространственного объекта можно рассчитать, перебрав попарно все вершины 3D фигуры изображения и вычислив получаемые расстояния. Расстояние с наибольшим значением и будет искомым отрезком.

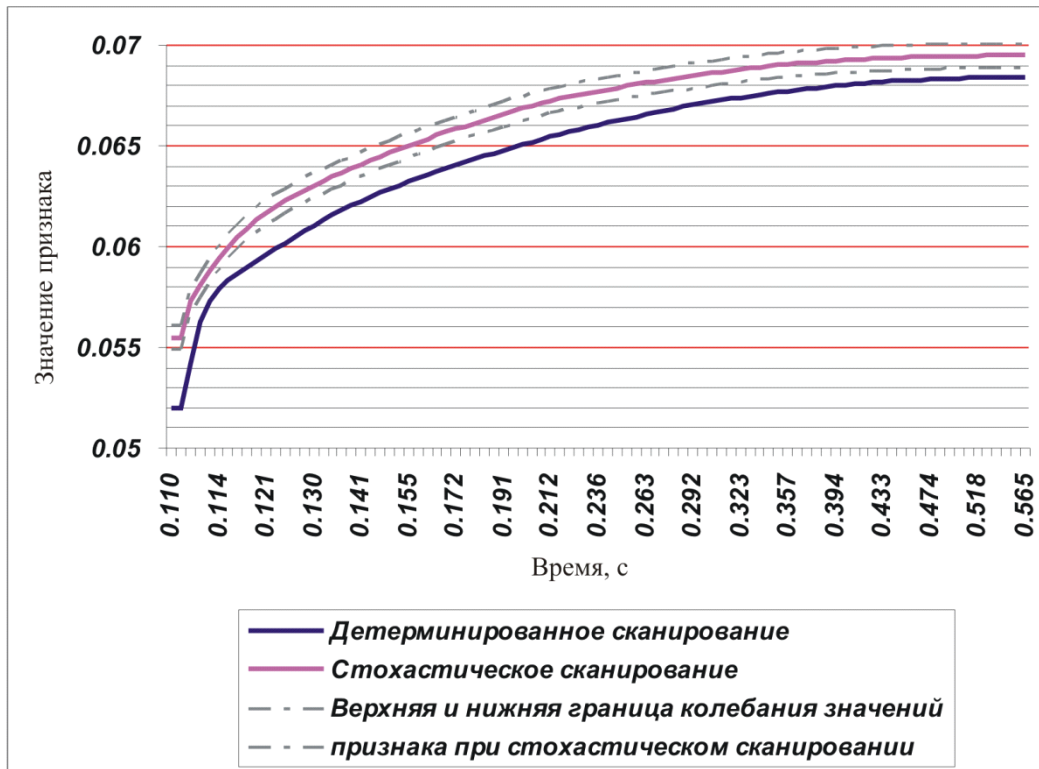
В качестве объекта была взята 3D модель бабочки с кодовым номером m11 из базы данных Prinston [21] (рис. 6). Модель имеет 9305 полигонов. Эксперимент проводился на компьютере со следующими характеристиками: ОС Windows 7 с архитектурой процессора Core i5-3570K CPU; тактовая частота 3,4 ГГц; оперативная память 8 ГБ.

В качестве алгоритмов построения равномерной сетки на сфере были взяты алгоритмы Saff_fix и Saff_rnd, описанные в настоящей статье выше. Для стохастического сканирования результаты эксперимента были усреднены для 100 испытаний.

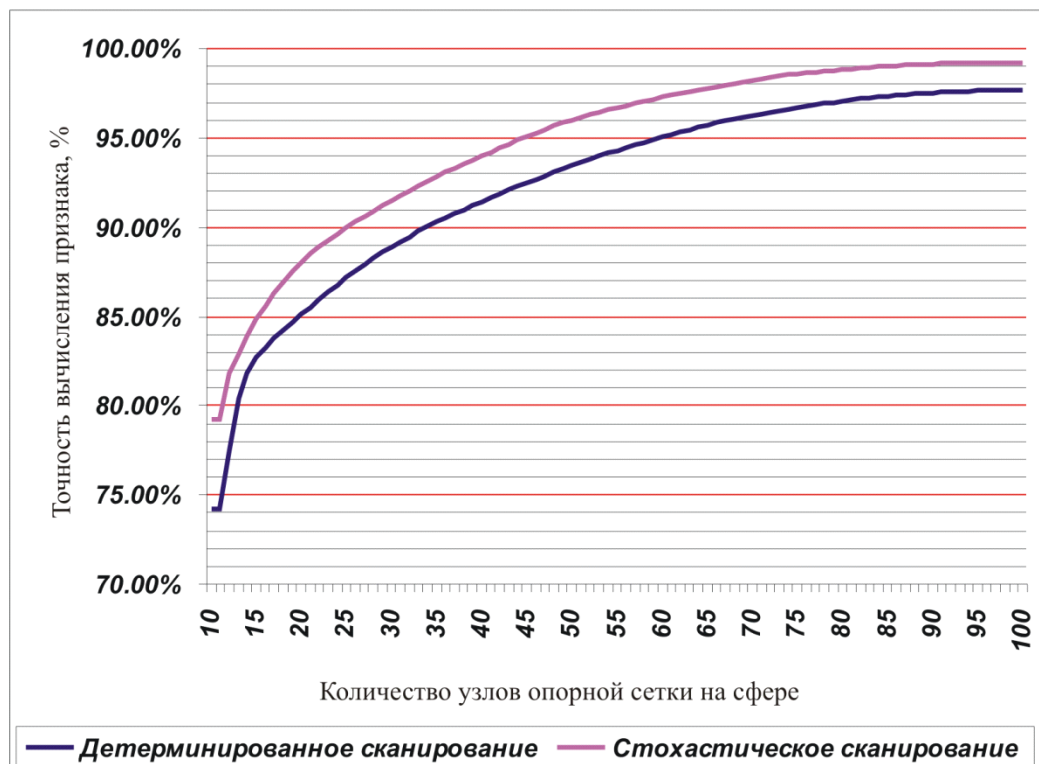
Сводные результаты эксперимента приведены на рис. 7.

На рис. 7, а показано быстродействие вычисления детерминированным и стохастическим способами сканирования признаков 3D изображения. Из графиков наглядно видно преимущество стохастического сканирования по сравнению с детерминированным. В частности, при одинаковой точности вычисления значения признака, равного 0,068 (истинное точное значение составило 0,07004), при стохастическом сканировании данное значение достигается за 0,252 с, тогда как при детерминированном сканировании — за 0,394 с. Таким образом, прирост быстродействия вычисления признака одной и той же точности вычисления 97,06% при стохастическом сканировании составляет: $(0,394 - 0,252)/0,394 = 36,06\%$.

На рис. 7, б показаны те же самые результаты эксперимента, но уже в категориях: точность вычисления признака от количества узлов опорной сетки на сфере.



(а)



(б)

Рис. 7 Быстродействие (а) и точность (б) вычисления признака 3D изображения двумя видами сканирования

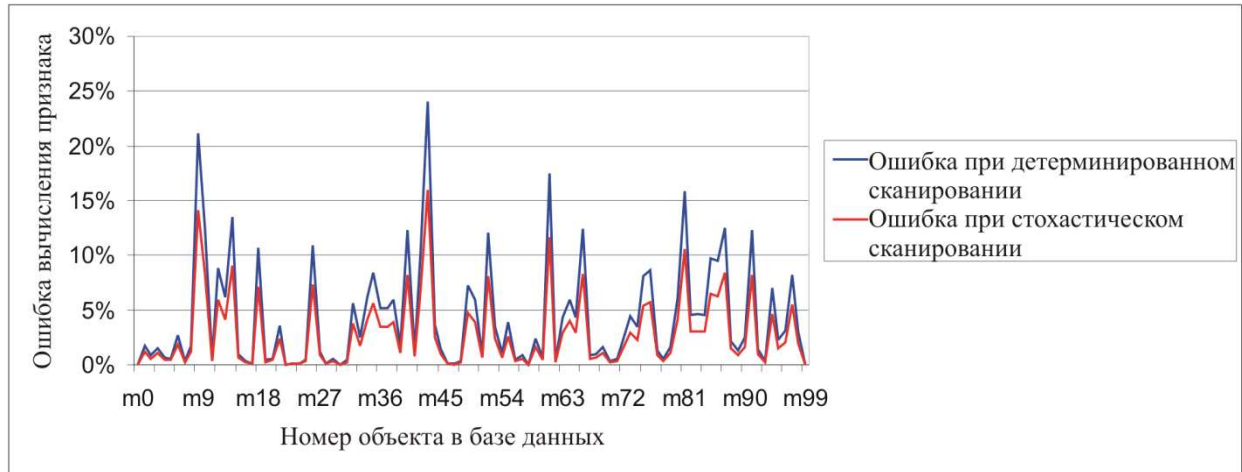


Рис. 8 Оценка ошибки вычисления признака для первых 100 моделей базы [21] в 20 узлах опорной сетки

Как видно из графиков, при добавлении одного узла сетки прирост времени вычисления признака растет с большей скоростью по сравнению с приростом точности вычисления признака (графики похожи по виду на логарифмическую функцию). В частности, видно, что при 20 узлах опорной сетки точность вычисления признака при детерминированном сканировании равна 85,10%, а при стохастическом способе — 88,03%, т.е. разница в точности вычисления признака почти 3%.

Ниже проведен эксперимент по оценке ошибки вычисления признака для первых ста 3D моделей пространственных объектов базы (рис. 8) [21]. Был взят другой признак, который несложно вычислить аналитически — максимально возможная площадь сечения объекта плоскостью:

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

Здесь $T(F_{\text{sect}} \cap l(\theta, \rho)) = \sum_t f(\theta, \rho, t)$; $P = \sum_{\rho} g(\theta, \rho) \cdot \Delta\rho$; $\Theta = \max_{\theta} g(\theta)$; $\text{HyperT}(F \cap B(\eta(\omega, \varphi), r)) = \Pi(F_{\text{sect}}) = G(\omega, \varphi, r)$; $\text{HyperP} = \max_r G(\omega, \varphi, r)$; $\text{Hyper}\Omega = \max_{\varphi} G(\omega, \varphi)$; $\text{Hyper}\Theta = \max_{\omega} G(\omega)$, где $\Delta\rho$ — расстояние между параллельными прямыми в плоскости сечения; $f(\theta, \rho, t)$ — длина t -го отрезка, высекаемого ρ -й прямой под θ -м углом наклона в плоскости сечения F_{sect} ; $\Pi(F_{\text{sect}}) = G(\omega, \varphi, r)$ — признак сечения (площадь), получаемого пересечением r -й плоскости $B(\eta(\omega, \varphi), r)$ под парой углов (ω, φ) обзора объекта.

Опорная сетка на сфере имела всего 20 узлов (для быстроты вычисления). Погрешность вычисления признака рассчитывается по формуле:

$$\varepsilon = \frac{|\text{Res}_i - \text{True}_i|}{\text{True}_i} \cdot 100\%,$$

где Res_i — значение признака для i -го изображения; True_i — истинное значение признака. Для стохастического эксперимента расчетное значение усреднялось для 100 испытаний.

Преимущество стохастического типа сканирования очевидно перед детерминированным при той же скорости вычисления в данном эксперименте. Поскольку далее будет действовать решающая процедура, которая по вектору признаков 3D изображения определит

класс изображения и которая одинакова для обоих типов распознавания 3D изображений, можно констатировать, что стохастическое распознавание пространственных объектов по соотношению «точность–быстродействие» превосходит детерминированное распознавание с фиксированного сеткой сканирования.

6 Заключение

Элементы теории стохастической геометрии, применявшиеся ранее только лишь для 2D изображений, впервые получили свое развитие на 3D изображения. Разрабатываемый метод сканирования 3D изображений плоскостями позволяет достичь полного инвариантного описания пространственных объектов и решить на этой основе проблему распознавания 3D изображений при их произвольной ориентации в пространстве и уровня масштабирования.

В настоящее время особенно актуальной становится задача ускорения вычислений сканирующей системы и достижения распознавания изображений в режиме реального времени. Для этого был проведен эксперимент, чтобы проанализировать особенности и результаты соотношения «точность–быстродействие» при применении гипертрейс-преобразования.

Учитывая данные проведенного эксперимента, а также результаты работ [13, 14], были выявлены следующие возможные пути улучшения быстродействия распознавания 3D изображения методами стохастической геометрии:

1. Разработать решающие процедуры на основе формирования триплетных признаков и дерева решений для повышения скорости распознавания. В частности, повышение скорости распознавания за счет построения естественной решающей процедуры, учитывающей особенности формирования триплетных признаков. Предполагается минимизировать количество дополнительных вычислительных операций.
2. Оценить временные затраты и сложность предлагаемых решающих процедур. В частности, планируется получить сравнительные оценки быстродействия распознавания 3D изображений, а также сделать анализ теоретической сложности разработанных решающих процедур.

Авторы планируют развить данный метод для анализа не только бинарных и монохромных 3D изображений [22], но также цветных и текстурных 3D изображений. Аналогичные результаты уже были получены при анализе цветных и текстурных 2D изображений в [23].

Литература

- [1] *Vasil'ev K. K., Dement'ev V. E., Andriyanov N. A.* Doubly stochastic models of images // *Pattern Recogn. Image Anal.*, 2015. Vol. 25. No. 1. P. 105–110. doi: 10.1134/S1054661815010204.
- [2] *Kiy K. I.* Segmentation and detection of contrast objects and their application in robot navigation // *Pattern Recogn. Image Anal.*, 2015. Vol. 25. No. 2. P. 338–346. doi: 10.1134/S1054661815020145.
- [3] *Myasnikov V. V.* Analysis of efficient linear local features of digital signals and images // *Pattern Recogn. Image Anal.*, 2016. Vol. 26. No. 1. P. 22–23. doi: 10.1134/S105466181601017X.
- [4] *Witten I. H., Frank E., Hall M. A.* Data mining: Practical machine learning tools and techniques. — The Morgan Kaufmann ser. in data management systems. — 3rd ed. — Burlington, MA, USA: Morgan Kaufmann, 2011. 664 p.
- [5] *Zhang N., Donahue J., Girshick R., Darrell T.* Part-based R-CNNs for fine-grained category detection // *Computer vision — ECCV 2014* / Eds. D. Fleet, T. Pajdla, B. Schiele,

- T. Tuytelaars. — Lecture notes in computer science ser. — Springer, 2014. Vol. 8689. P. 834–849. doi: 10.1007/978-3-319-10590-1_54.
- [6] *Song S., Xiao J.* Sliding shapes for 3D object detection in depth images // Computer vision — ECCV 2014 / Eds. D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars. — Lecture notes in computer science ser. — Springer, 2014. Vol. 8694. P. 634–651. doi: 10.1007/978-3-319-10599-4_41.
- [7] *Wang C., Huang K.-Q.* VFM: Visual feedback model for robust object recognition // J. Comput. Sci. Technol., 2015. Vol. 30. Iss. 2. P. 325–339. doi: 10.1007/s11390-015-1526-1.
- [8] *Andreux M., Rodolà E., Aubry M., Cremers D.* Anisotropic Laplace–Beltrami operators for shape analysis // Computer vision — ECCV 2014 / Eds. L. Agapito, M.M. Bronstein, C. Rother. — Image processing, computer vision, patterns recognition, and graphics ser. — Springer, 2015. Vol. 8928. P. 299–312. doi: 10.1007/978-3-319-16220-1_21.
- [9] *Lmaati E. A., Oirrak A. E., Kaddioui M. N., Ouahman A. A., Sadgal M.* 3D model retrieval based on 3D discrete cosine transform // Int. Arab J. Inf. Technol., 2010. Vol. 7. No. 3. P. 264–270.
- [10] *Boucher M., Evans A. C., Siddiqi K.* Anisotropic diffusion of tensor fields for fold shape analysis on surfaces // Information processing in medical imaging / Eds. G. Szekely, H. Hahn. — Image processing, computer vision, pattern recognition, and graphics ser. — Berlin – Heidelberg: Springer-Verlag, 2011. Vol. 6801. P. 271–282.
- [11] *Litman R., Bronstein A.* Learning spectral descriptors for deformable shape correspondence // IEEE T. Pattern Anal., 2014. Vol. 36. Iss. 1. P. 171–180.
- [12] *Elhachloufi M., Oirrak A. El., Aboutajdine D., Kaddioui M. N.* Affine invariant descriptors of 3D object using multiple regression model // Int. J. Comput. Sci. Inform. Technol., 2011. Vol. 3. No. 1. P. 1–10.
- [13] *Fedotov N. G., Moiseev A. V., Syetov A. A., Lizunkov V. G., Kindaev A. Y.* New methods of three-dimensional images recognition based on stochastic geometry and functional analysis // IOP Conf. Ser. Mat. Sci., 2017. Vol. 177. Iss. 1. P. 012047. doi: 10.1088/1757-899X/177/1/012047. <http://stacks.iop.org/1757-899X/177/i=1/a=012047>.
- [14] *Fedotov N. G., Ryndina S. V., Semov A. A.* Trace transform of three-dimensional objects: Recognition, analysis and database search // Pattern Recogn. Image Anal., 2014. Vol. 24. No. 4. P. 566–574. doi: 10.1134/S105466181404004X.
- [15] *Федотов Н. Г., Семов А. А., Моисеев А. В.* Анализ условий, влияющих на свойства конструируемых признаков 3D-изображений // Компьютерная оптика, 2016. Т. 40. № 6. С. 887–894. doi: 10.18287/2412-6179-2016-40-6-887-894.
- [16] *Fedotov N. G.* The theory of image-recognition features based on stochastic geometry // Pattern Recogn. Image Anal., 1998. Vol. 8. No. 2. P. 264–266.
- [17] *Fedotov N. G., Mokshanina D. A.* Recognition of images with complex half-tone texture // Measurement Techniques, 2011. Vol. 53. No. 11. P. 1226–1232. doi: 10.1007/s11018-011-9646-7.
- [18] *Федотов Н. Г., Семов А. А., Моисеев А. В.* Интеллектуальные возможности гипертрейс-преобразования: конструирование признаков с заданными свойствами // Машинное обучение и анализ данных, 2014. Т. 1. № 9. С. 1200–1214.
- [19] *Федотов Н. Г., Семов А. А., Моисеев А. В.* Минимизация признакового пространства распознавания 3D изображения на основе стохастической геометрии и функционального анализа // Машинное обучение и анализ данных, 2015. Т. 1. № 13. С. 1796–1814. doi: 10.21469/22233792.1.13.03.
- [20] *Rakhmanov E., Saff E. B., Zhou Y. M.* Minimal discrete energy on the sphere // Math. Res. Lett., 1994. Vol. 1. No. 1. P. 647–662.
- [21] Princeton Shape Benchmark of 3D models database. <http://shape.cs.princeton.edu/benchmark/>.

- [22] Федотов Н. Г., Семов А. А. Программный комплекс анализа и распознавания 3D изображений на основе пространственного трейс-преобразования со случайными параметрами сканирования. Свидетельство об официальной регистрации программ для ЭВМ № 2015612257 Роспатента от 16.02.15.
- [23] Fedotov N. G., Mokshanina D. A. Recognition of halftone textures from the standpoint of stochastic geometry and functional analysis // Pattern Recognition Image Anal., 2010. Vol. 20. No. 4. P. 551–556. doi: 10.1134/S1054661810040164.

Поступила в редакцию 03.09.2017

Performance investigation of three-dimensional image recognition by stochastic geometry methods independent on the number of reference points on the sphere*

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

fedotov@pnzgu.ru; matematik_aleksey@mail.ru; moigus@mail.ru

¹Penza State University, 40 Krasnaya Str., Penza, Russia

²Comearth, 11a Gagarina Str., Penza, Russia

Background: A new developed approach to the three-dimensional (3D) images recognition, giving the object invariant description for any its spatial orientations is proposed. This method has many advantages and 3D images data mining capabilities. In particular, in parallel with the spatial object recognition, it is possible to analyze the original image. Due to building a rigorous mathematical model, it is possible to design analytically features with predetermined properties.

Methods: The suggested approach is based on the modern methods of stochastic geometry and functional analysis. Hypertrace transform creates a 3D trace-image of the original spatial object due to scan of the parallel planes grid from different view angles. Created on this trace-image basis, hypertrace matrix is a convenient tool for analyzing 3D images in contrast to other mathematical methods.

Results: Stochastic scan with random parameters is more efficient than the determinate scan in terms of the 3D images recognition “reliability–performance” relation. The conducted experiments results are shown. These results demonstrate both theoretical and practical significance and effectiveness of the proposed method.

Concluding Remarks: The evaluation task of 3D image recognition performance independent on the number of reference points on the sphere with the use of various kinds of scanning are analyzed. Potential further ways to accelerate the recognition system are proposed.

Keywords: 3D image recognition; hypertrace-transform; performance of the recognition system; reference grid on the sphere; invariant description; stochastic scan

DOI: 10.21469/22233792.3.3.01

References

- [1] Vasil’ev, K. K., V. E. Dement’ev, and N. A. Andriyanov. 2015. Doubly stochastic models of images. *Pattern Recogn. Image Anal.* 25(1):105–110. doi: 10.1134/S1054661815010204.
- [2] Kiy, K. I. 2015. Segmentation and detection of contrast objects and their application in robot navigation. *Pattern Recogn. Image Anal.* 25(2):338–346. doi: 10.1134/S1054661815020145.

*The research was supported by the Russian Foundation for Basic Research (grant No. 15-07-04484).

- [3] Myasnikov, V. V. 2016. Analysis of efficient linear local features of digital signals and images. *Pattern Recogn. Image Anal.* 26(1):22–23. doi: 10.1134/S105466181601017X.
- [4] Witten, I. H., E. Frank, and M. A. Hall. 2011. *Data mining: Practical machine learning tools and techniques*. The Morgan Kaufmann ser. in data management systems. 3rd ed. Burlington, MA: Morgan Kaufmann. 664 p.
- [5] Zhang, N., J. Donahue, R. Girshick, and T. Darrell. 2014. Part-based R-CNNs for fine-grained category detection. *Computer vision — ECCV 2014*. Eds. D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars. Lecture notes in computer science ser. Springer. 8689:834–849. doi: 10.1007/978-3-319-10590-1_54.
- [6] Song, S., and J. Xiao. 2014. Sliding shapes for 3D object detection in depth images. *Computer vision — ECCV 2014*. Eds. D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars. Lecture notes in computer science ser. Springer. 8694:634–651. doi: 10.1007/978-3-319-10599-4_41.
- [7] Wang, C., and K.-Q. Huang. 2015. VFM: Visual feedback model for robust object recognition. *J. Comput. Sci. Technol.* 30(2):325–339. doi: 10.1007/s11390-015-1526-1.
- [8] Andreux, M., E. Rodolà, M. Aubry, and D. Cremers. 2014. Anisotropic Laplace–Beltrami operators for shape analysis. *Computer vision — ECCV 2014*. Eds. L. Agapito, M. M. Bronstein, and C. Rother. Image processing, computer vision, pattern recognition, and graphics ser. Springer. 8928:299–312. doi: 10.1007/978-3-319-16220-1_21.
- [9] Lmaati, E. A., A. E. Oirrak, M. N. Kaddioui, A. A. Ouahman, and M. Sadgal. 2010. 3D model retrieval based on 3D discrete cosine transform. *Int. Arab J. Inf. Technol.* 7(3):264–270.
- [10] Boucher, M., A. C. Evans, and K. Siddiqi. 2011. Anisotropic diffusion of tensor fields for fold shape analysis on surfaces. *Information processing in medical imaging*. Eds. G. Szekely and H. Hahn. Image processing, computer vision, pattern recognition, and graphics ser. Berlin–Heidelberg: Springer-Verlag. 6801:271–282.
- [11] Litman, R., and A. Bronstein. 2014. Learning spectral descriptors for deformable shape correspondence. *IEEE T. Pattern Anal.* 36(1):171–180.
- [12] Elhachloufi, M., A. El. Oirrak, D. Aboutajdine, and M. N. Kaddioui. 2011. Affine invariant descriptors of 3D object using multiple regression model. *Int. J. Comput. Sci. Inform. Technol.* 3(1):1–10.
- [13] Fedotov, N. G., A. V. Moiseev, A. A. Syemov, V. G. Lizunkov, and A. Y. Kindaev. 2017. New methods of three-dimensional images recognition based on stochastic geometry and functional analysis. *IOP Conf. Ser. Mat. Sci.* 177(1):012047. doi: 10.1088/1757-899X/177/1/012047. Available at: <http://stacks.iop.org/1757-899X/177/i=1/a=012047> (accessed December 29, 2017).
- [14] Fedotov, N. G., S. V. Ryndina, and A. A. Semov. 2014. Trace transform of three-dimensional objects: Recognition, analysis and database search. *Pattern Recogn. Image Anal.* 24(4):566–574. doi: 10.1134/S105466181404004X.
- [15] Fedotov, N. G., A. A. Syemov, and A. V. Moiseev. 2016. Analiz usloviy, vliyayushhikh na svoystva konstruiruemyykh priznakov 3D-izobrazheniy [Analysis of conditions that influence the properties of the constructed 3D’=image features]. *Komp’juternaya optika* [Computer optics] 40(6):887–894. doi: 10.18287/2412-6179-2016-40-6-887-894.
- [16] Fedotov, N. G. 1998. The theory of image-recognition features based on stochastic geometry. *Pattern Recogn. Image Anal.* 8(2):264–266.
- [17] Fedotov, N. G., and D. A. Mokshanina. 2011. Recognition of images with complex half-tone texture. *Measurement Techniques* 53(11):1226–1232.
- [18] Fedotov, N. G., A. A. Syemov, and A. V. Moiseev. 2014. Intellektual’nye vozmozhnosti i giperpreobrazovaniya: konstruirovaniye priznakov s zadannymi svoystvami [Intelligent capabilities hy-

- pertrace transform: Constructing features with predetermined properties]. *Machine Learning Data Anal.* 1(9):1200–1214.
- [19] Fedotov, N. G., A. A. Syemov, and A. V. Moiseev. 2015. Minimizatsiya priznakovogo prostranstva raspoznavaniya 3D izobrazheniya na osnove stokhasticheskoy geometrii i funktsional'nogo analiza [Feature space minimization of 3D image recognition based on stochastic geometry and functional analysis]. *Machine Learning Data Anal.* 1(13):1796–1814. doi: 10.21469/22233792.1.13.03.
- [20] Rakhmanov, E., E. B. Saff, and Y. M. Zhou. 1994. Minimal discrete energy on the sphere. *Math. Res. Lett.* 1(1):647–662.
- [21] Princeton Shape Benchmark of 3D models database. Available at: <http://shape.cs.princeton.edu/benchmark/> (accessed May 11, 2017).
- [22] Fedotov, N. G., and A. A. Syemov. February 16, 2015. Programmnyy kompleks analiza i raspoznavaniya 3D izobrazheniy na osnove prostransvennogo treys-preobrazovaniya so sluchaynymi parametrami skanirovaniya [Software for 3D images analysis and recognition based on the spatial trace transform with random scan parameters]. Official registration certificate for computer programs No. 2015612257 of Rospatent.
- [23] Fedotov, N. G., and D. A. Mokshanina. 2010. Recognition of halftone textures from the standpoint of stochastic geometry and functional analysis. *Pattern Recognition Image Anal.* 20(4):551–556. doi: 10.1134/S1054661810040164.

Received September 03, 2017