

Машинное обучение и анализ данных

Журнал «Машинное обучение и анализ данных» публикует новые теоретические и обзорные статьи с результатами научных исследований в области искусственного интеллекта, теоретических основ информатики и приложений. Цель журнала — развитие теории машинного обучения, интеллектуального анализа данных и методов проведения вычислительных экспериментов. Принимаются статьи на английском и русском языках.

Журнал включен в российский индекс научного цитирования РИНЦ. Информация о цитировании статей находится на сайте Российского индекса научного цитирования, ISSN 2223-3792, номер свидетельства о регистрации ЭЛ № ФС 77-55486. Журнал зарегистрирован в системе Crossref, doi <http://dx.doi.org/10.21469/22233792>.

- Новостной сайт <http://jmla.org/>
- Электронная система подачи статей <http://jmla.org/papers/>
- Правила подготовки статей <http://jmla.org/papers/doc/authors-guide.pdf>

Тематика журнала:

- классификация, кластеризация, регрессионный анализ;
- алгебраический подход к проблеме синтеза корректных алгоритмов;
- многомерный статистический анализ;
- выбор моделей и сложность;
- глубокое обучение;
- статистическая теория обучения;
- методы прогнозирования временных рядов;
- методы обработки и распознавания сигналов;
- методы оптимизации в задачах машинного обучения и анализа данных;
- методы анализа больших данных;
- методы визуализации данных;
- обработка и распознавание речи и изображений;
- анализ и понимание текста;
- информационный поиск;
- прикладные задачи анализа данных.

Редакционный совет

Ю. Г. Евтушенко, акад.
Ю. И. Журавлёв, акад.
Д. Н. Зорин, проф.
К. В. Рудаков, акад.

Редколлегия

К. В. Воронцов, д.ф.-м.н.
А. Г. Дьяконов, д.ф.-м.н.
И. А. Матвеев, д.т.н.
Л. М. Местецкий, д.т.н.
В. В. Моттль, д.т.н.
М. Ю. Хачай, д.ф.-м.н.

Координаторы

В. Р. Бунакова
Ш. Х. Ишкина
М. П. Кузнецов
А. П. Мотренко

Редактор: В. В. Стрижов, д.ф.-м.н. (strijov@ccas.ru)

Вычислительный центр им. А. А. Дородницына ФИЦ ИУ РАН
Московский физико-технический институт
Факультет управления и прикладной математики
Кафедра «Интеллектуальные системы»

Москва, 2017

Journal of Machine Learning and Data Analysis

The journal Machine Learning and Data Analysis publishes original research papers and reviews of the developments in the field of artificial intelligence, theoretical computer science and its applications. The journal aims to promote the theory of machine learning and data mining and methods of conducting computational experiments. Papers are accepted in English and Russian.

The journal is included in the Russian science citation index RSCI. Information about citation to articles can be found at the Russian science citation index website. ISSN 2223-3792. Mass media registration certificate ЭЛ № ФС 77-55486. The Crossref journal doi is <http://dx.doi.org/10.21469/22233792>.

- Journal news and archive <http://jmla.org/>
- Open journal system for papers submission <http://jmla.org/papers/>
- Style guide for authors <http://jmla.org/papers/doc/authors-guide.pdf>

The scope of the journal:

- classification, clustering, regression analysis;
- multidimensional statistical analysis;
- Bayesian methods for regression and classification;
- model selection and complexity;
- deep learning;
- Statistical Learning Theory;
- time series forecasting techniques;
- methods of signal processing and speech recognition;
- optimization methods for solving machine learning and data mining problems;
- methods of big data analysis;
- data visualization techniques;
- methods of image processing and recognition;
- text analysis, text mining and information retrieval;
- applied data analysis problems.

Editorial Council

Yu. G. Evtushenko, acad.
K. V. Rudakov, acad.
Yu. I. Zhuravlev, acad.
D. N. Zorin, prof.

Editorial Board

A. G. Dyakonov, D.Sc.
M. Yu. Khachay, D.Sc.
I. A. Matveev, D.Sc.
L. M. Mestetskiy, D.Sc.
V. V. Mottl, D.Sc.
K. V. Vorontsov, D.Sc.

Editorial Support

V. R. Bunakova
Sh. Kh. Ishkina
M. P. Kuznetsov
A. P. Motrenko

Editor-in-Chief: V. V. Strijov, D.Sc. (strijov@ccas.ru)

Dorodnicyn Computing Centre FRC CSC RAS
Moscow Institute of Physics and Technology
Department of Control and Applied Mathematics
Division “Intelligent Systems”

Moscow, 2017

Содержание

Н. Г. Федотов, А. А. Сёмов, А. В. Моисеев

Исследование быстродействия распознавания трехмерного изображения методами стохастической геометрии в зависимости от количества опорных точек на сфере 176

В. М. Старожилец, Ю. В. Чехович

Об идентификации статистической модели транспортных потоков с использованием групп автомобильно-транспортных средств 193

Н. А. Самсонов, А. Н. Гнеушев

Дескриптор в аккумуляторном пространстве Хафа градиентного поля изображения для детектирования пешеходов 203

Contents

<i>N. G. Fedotov, A. A. Syemov, and A. V. Moiseev</i>	
Performance investigation of three-dimensional image recognition by stochastic geometry methods independent on the number of reference points on the sphere	176
<i>V. M. Starozhilets and Yu. V. Chehovich</i>	
About identification of a statistical model of traffic flows using vehicle groups	193
<i>N. A. Samsonov and A. N. Gneushev</i>	
Textural descriptor in the Hough accumulator space of the gradient field for detecting pedestrians	203

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

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

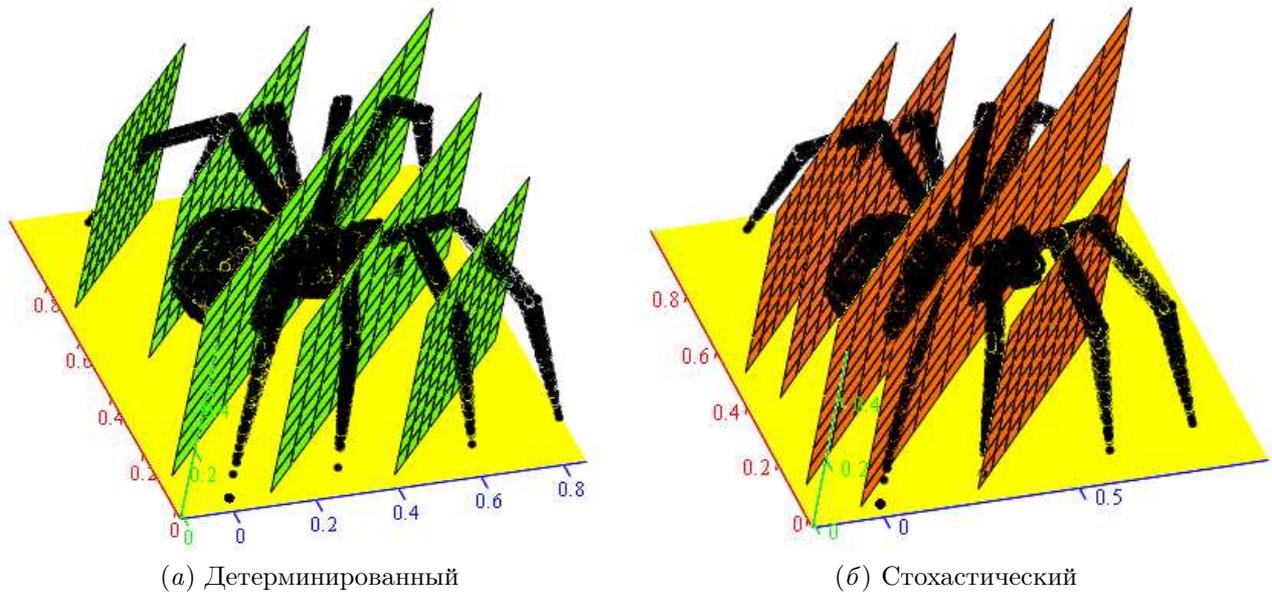


Рис. 1 Типы сканирования 3D изображения объекта

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)

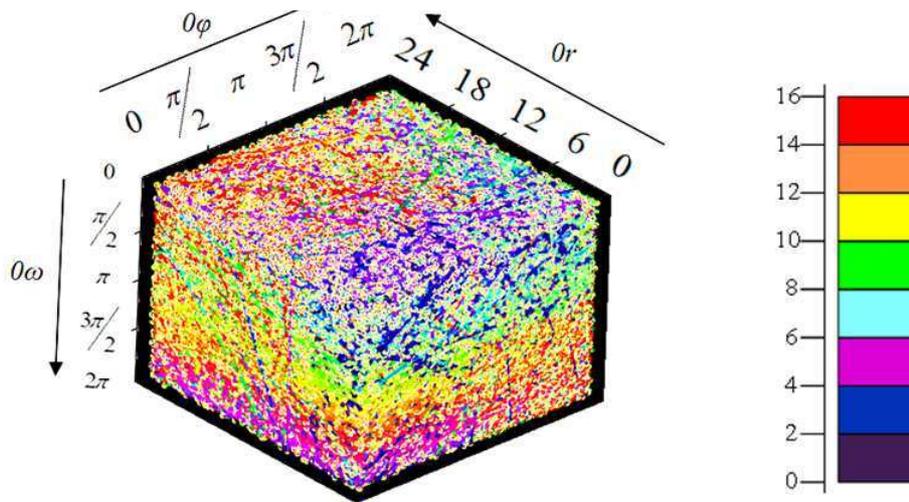


Рис. 2 Пример графического представления гипертрейс-матрицы ЗТМ

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 $\text{Res}(F)$ — гипертриплетный 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 $g = \text{T}(F_{\text{sect}} \cap l(\theta, \rho))$, описывающим любой заданный признак пересечений прямой с изображением.

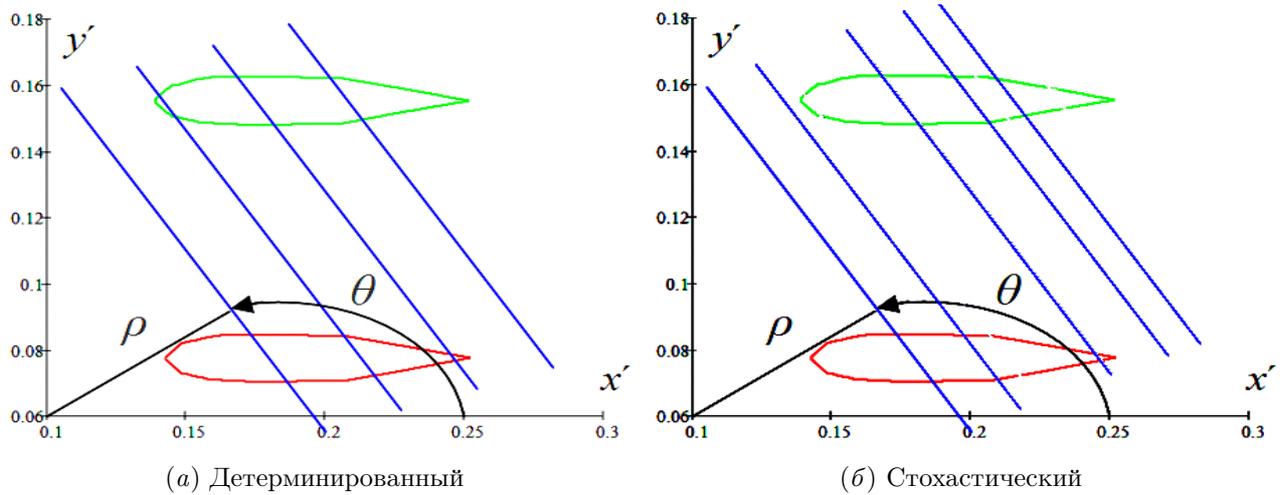


Рис. 3 Типы сканирования 2D изображения сечения

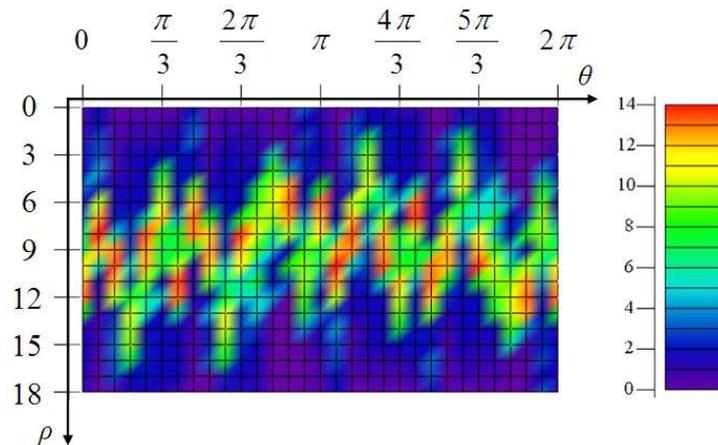


Рис. 4 Пример графического представления трейс-матрицы ТМ

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 :

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

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.

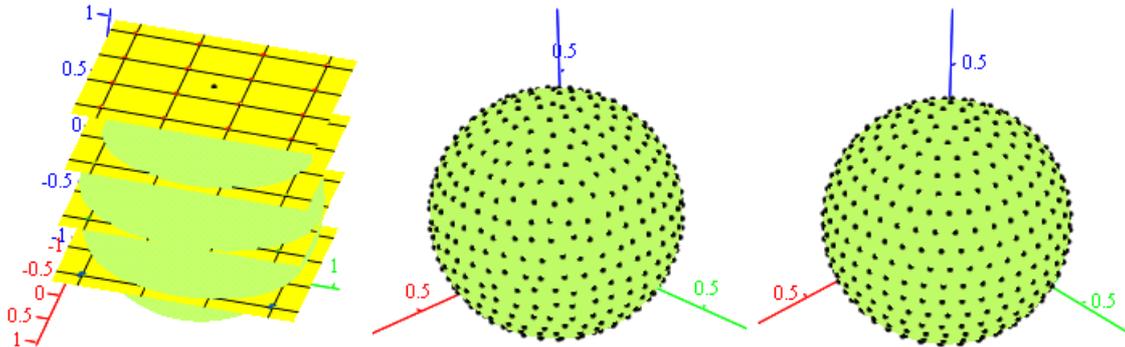


Рис. 5 Построение равномерной сетки на сфере детерминированным (**Saff_fix**) и стохастическим (**Saff_rnd**) способами для $N = 500$

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.

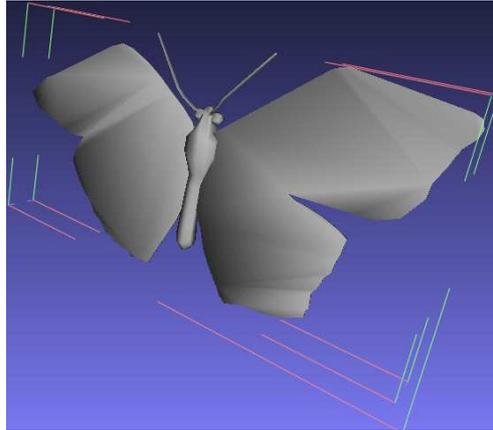


Рис. 6 Трехмерная модель бабочки m11 базы The Princeton Shape Benchmark

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.

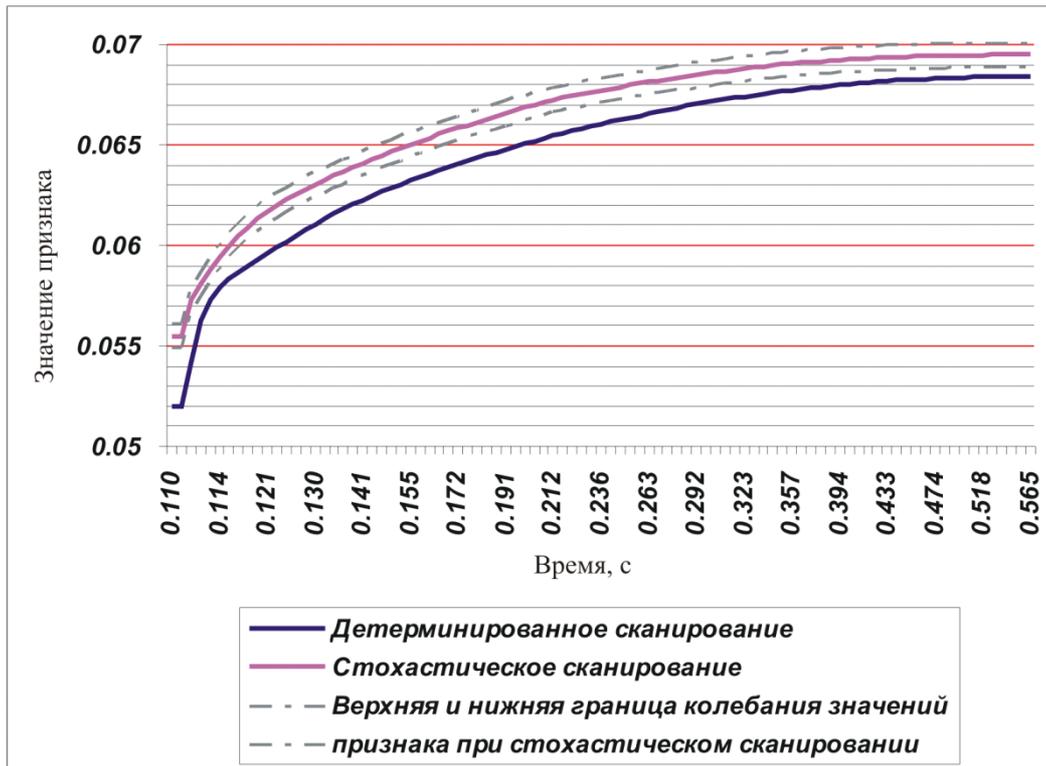
В качестве объекта была взята 3D модель бабочки с кодовым номером m11 из базы данных Prinston [21] (рис. 6). Модель имеет 9305 полигонов. Эксперимент проводился на компьютере со следующими характеристиками: ОС Windows 7 с архитектурой процессора Core i5-3570K CPU; тактовая частота 3,4 ГГц; оперативная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

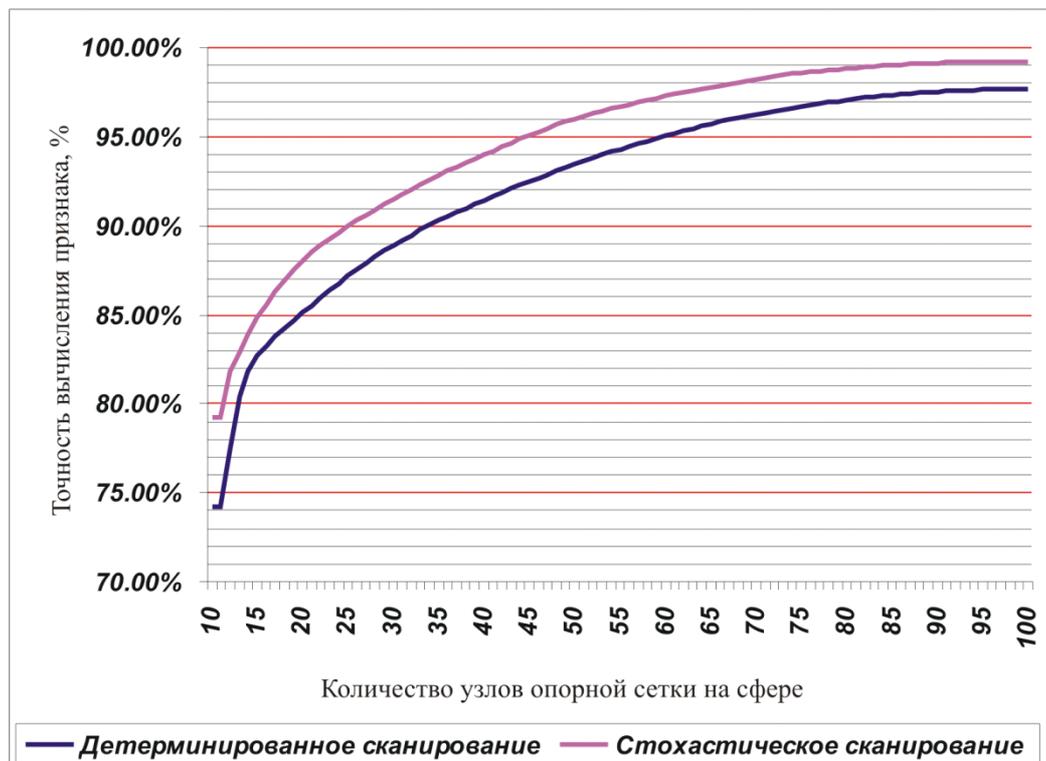
Сводные результаты эксперимента приведены на рис. 7.

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

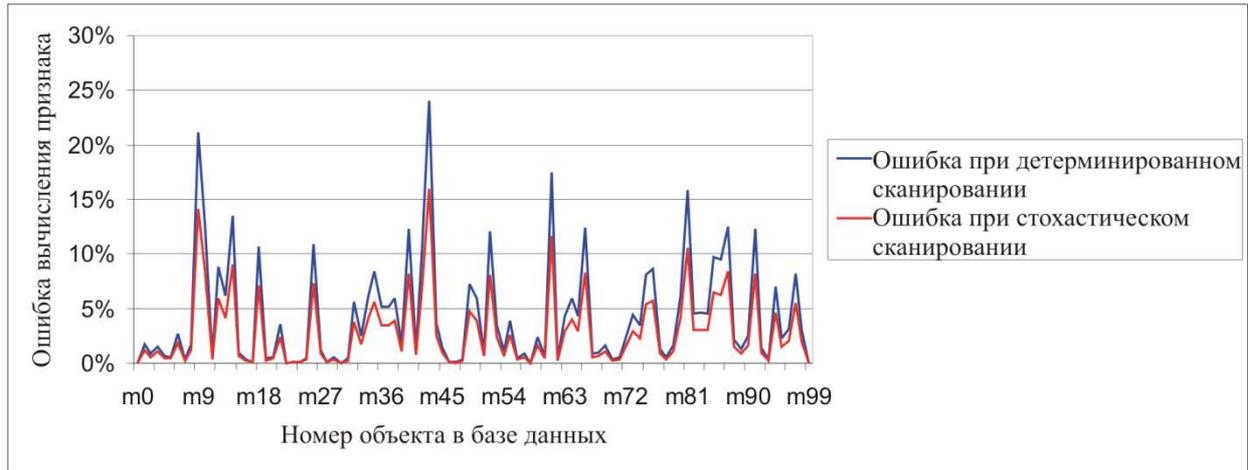


Рис. 8 Оценка ошибки вычисления признака для первых 100 моделей базы [21] в 20 узлах опорной сетки

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’juternaja 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. Intellectual’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

Об идентификации статистической модели транспортных потоков с использованием групп автомобильно-транспортных средств

В. М. Старожилец^{1,2}, *Ю. В. Чехович*^{1,2}
starvsevol@gmail.com; chehovich@forecsys.ru

¹ФИЦ «Информатика и управление» РАН, Россия, г. Москва, ул. Вавилова, 44/2

²Московский физико-технический институт, Россия, г. Долгопрудный, Институтский пер., 9

Предлагается статистическая модель транспортных потоков для моделирования движения транспортных средств на автомагистралях, идентифицируемая на данных из гетерогенных источников. Модель симулирует движение групп транспортных средств по магистрали с использованием фундаментальной диаграммы на выбранном участке автодороги для расчета скорости группы. Проводятся вычислительные эксперименты для подтверждения работоспособности модели, а также анализа ее поведения при моделировании ситуации перекрытия одной из полос автомагистрали. Критерием качества выбрана среднеквадратичная ошибка между предсказанным числом проехавших автотранспортных средств и их реальным числом. Используются данные дорожных датчиков Центра организации дорожного движения, а также данные, полученные с помощью видеосъемки.

Ключевые слова: *моделирование транспортных потоков; МКАД; дорожные датчики*

DOI: 10.21469/22233792.3.3.02

1 Введение

Данная работа посвящена описанию модели, предназначенной для моделирования транспортных потоков на автомагистрали (например, Московской кольцевой автомобильной дороге (МКАД)) с использованием анонимных данных с GPS-треков и дорожных датчиков. Процедура комплексирования данных из этих двух источников подробно рассмотрена в [1]. Также в данной работе проводятся эксперименты как на данных с дорожных датчиков, так и полученных видеосъемкой, показывающих работоспособность представленной модели.

Моделирование транспортных потоков основано на их сходстве с жидкой или газовой средой. В частности, базовая модель Лайтхилла–Уизема–Ричардса (Lighthill–Whitham–Richards, LWR) [2–4] основана на предположении о существовании взаимно-однозначной зависимости между скоростью и плотностью потока автомобильно-транспортных средств (АТС) и сохранении числа АТС в транспортной сети. В современном макроскопическом подходе транспортный поток описывается нелинейной системой гиперболических дифференциальных уравнений в частных производных второго порядка в различных постановках [5–12]. В данной работе будет, однако, рассмотрена модель, которая не пытается поставить в соответствие потоку АТС поток жидкости или газа, а использует общие знания о характеристиках участка автомагистрали (фундаментальную диаграмму [13] на ней), а также усредненные физические размеры автомобилей для моделирования состояния магистрали в каждый момент времени.

История развития задач моделирования транспортной сети, а также существующие модели широко рассмотрены в [14]. В [15] приводится общая схема моделирования, схожая с предлагаемой, и проводится моделирование большого участка автомагистрали, что

совпадает с нашей конечной целью. Основное отличие данной модели от уже представленных — это то, что рассматривается движение неразделимых групп автомобилей по магистрали (которые, однако, могут соединяться между собой) вместо движения самих автомобилей, считая скорость транспортных средств в группе одинаковой. Хотя данное приближение довольно грубое, будет показано, что его достаточно для получения результатов, хорошо совпадающих с реальными измерениями, при любых режимах автомагистрали [16], а также моделирования перекрытия полос на магистрали.

2 Описание модели

2.1 Структура модели

В данном разделе определим топологию, моделирующую структуру автомагистрали в предлагаемой модели. Будем считать, что только вершины графа являются стоками и источниками АТС и никаких дополнительных стоков и источников не бывает. Временной интервал одного шага симуляции в этой работе обозначен как τ .

Пусть $\mathbf{G} = \{\mathbf{V}, \mathbf{E}\}$ — связный ориентированный граф, представляющий собой модель автомагистрали, где \mathbf{V} — множество ребер; \mathbf{E} — множество упорядоченных пар вершин. Пусть также $\tilde{\mathbf{G}} = \{\mathbf{L}, \mathbf{N}, \mathbf{Ex}, \mathbf{In}, \mathbf{Tr}, \mathbf{M}, \mathbf{N}_{\text{exit}}\}$ — свойства участков автомагистрали, соответствующих данным ребрам и вершинам, где $L_{i,k} \in \mathbf{L}$ — длина участка автомагистрали, соответствующего ребру (i, k) в метрах; $n_{i,k} \in \mathbf{N}$ — число полос, по которым разрешено движение АТС в ребре (i, k) , $\text{Ex}_i \in \mathbf{Ex}$ — максимальное число АТС, которое может съехать с дороги при достижении данной вершины i за время τ ; $\text{In}_{i,k} \in \mathbf{In}_i \in \mathbf{In}$ — число АТС, которые должны въехать на автомагистраль с вершины i за время τ в момент времени k ; $\text{Tr}_i \in \mathbf{Tr}$ — максимальное число АТС, которые могут переехать с ребра (m, i) на ребро (i, l) ; $M_{i,k} \in \mathbf{M}_i \in \mathbf{M}$ — доля АТС, которая съезжает с автомагистрали при достижении вершины i в момент времени k ; $N_{\text{exit},i} \in \mathbf{N}_{\text{exit}}$ — количество АТС, которые стоят в очереди на съезд с автомагистрали в узле i . При этом $\delta(\mathbf{G}) = 1$ и $\Delta(\mathbf{G}) = 2$ — минимальная и максимальная степень вершины в графе соответственно. В данной работе не проводится моделирование всей автомагистрали, а только небольших ее участков, поэтому $|\mathbf{V}| = 2$.

Как уже упоминалось выше, в данной статье не рассматривается движение каждого отдельного автомобиля, а только движение групп автомобилей $\mathbf{A}_l \in \bar{\mathbf{A}}_{i,j}$, где $\bar{\mathbf{A}}_{i,j}$ — множество всех автомобильных групп в ребре (i, j) , упорядоченное по убыванию расстояния до начала ребра, $l \in \mathbb{N}_+$. Множество \mathbf{A}_l представимо в виде $\{\text{Pos}_l, V_l, N_l\}$, где $\text{Pos}_l \in [0, L_{i,j}]$ — положение группы АТС в ребре (i, j) , $V_l \in \mathbb{Q}_+$ — скорость группы АТС, $N_l \in \mathbb{Q}_+$ — число АТС в группе. Заметим, что на данный момент нигде не предполагается, что число АТС в группе съезжающих либо въезжающих АТС должно быть целым числом.

2.2 Процедура расчета

Теперь, когда структура модели определена, опишем процедуру расчета движения АТС в данной модели. Окончательная цель предлагаемого алгоритма — это по имеющимся данным о числе въехавших на автомагистраль АТС получить число проезжающих по любому участку магистрали АТС в любой момент времени. Начнем с методики расчета скорости движения группы АТС. С помощью фундаментальной диаграммы для каждого ребра (i, j) определяется функция $V = f_{i,j}(\rho)$, $f_{i,j} : \mathbb{Q}_+ \rightarrow \mathbb{Q}_+$, где $\rho \in \mathbb{R}_+$ — плотность АТС.

Процедура построения этой функции детально описана в работах [1, 13]. Также важным фактором является то, что скорость для конкретной группы АТС рассчитывается исходя из плотности АТС на участке автомагистрали перед ним либо же по плотности АТС на следующем участке автомагистрали, если группа достигла конца своего участка. Таким образом, для каждого $\mathbf{A}_l \in \bar{\mathbf{A}}_{i,j}$ скорость определяется следующим образом:

$$V_l = f_{i,j}(\rho_l), \quad (1)$$

где $\rho_l \in \mathbb{Q}_+$ — плотность АТС на участке автомагистрали перед группой АТС \mathbf{A}_l :

$$\rho_l = \begin{cases} \frac{N_{\text{exit},j} + \sum_{m=l+1}^{|\bar{\mathbf{A}}_{i,j}|} N_m}{(L_{i,j} - \text{Pos}_l) N_{i,j}}, N_m \in \mathbf{A}_m \in \bar{\mathbf{A}}_{i,j}, \text{Pos}_l < L_{i,j}; \\ \frac{N_{\text{exit},k} + \sum_{m=1}^{|\bar{\mathbf{A}}_{j,k}|} N_m}{L_{j,k} N_{j,k}}, N_m \in \mathbf{A}_m \in \bar{\mathbf{A}}_{j,k}, \text{Pos}_l = L_{i,j}. \end{cases}$$

Теперь опишем саму процедуру моделирования. Расчет начинается либо в вершине j : $\forall i \in \mathbf{V}! \exists (j, i)$, либо в любой вершине графа, идет в направлении против ориентации ребер. Положим, что $L_{\text{avg}} \in \mathbb{Q}_+$ — средняя длина автомобилей в метрах; $N_{\text{max},(j,m)} = L(j, m) n_{(j,m)} / L_{\text{avg}}$ — максимальное число АТС на ребре (j, m) ; $b_{l,i,j} \in \{0, 1\}$ — маркер того, съезжали ли из группы АТС l на ребре (i, j) автомобили с автомагистрали; $\widetilde{\text{Tr}}_{i,k}$ — число АТС, проезжающих с ребра m, i на ребро i, j за временной интервал $[\tau(k-1), \tau k]$. Вид алгоритма 1 представлен ниже.

3 Вычислительный эксперимент

Проводятся три эксперимента на данных о въехавших и съехавших с участка автомагистрали АТС, полученных ручной съемкой. Съемка велась на пересечении Варшавского шоссе с Новоданиловским и Нагорными проездами, длина участка автомагистрали — 1250 м, а также с помощью дорожных датчиков на МКАД между 99 и 101 км. Тут проводилось два эксперимента — моделирование числа съехавших АТС по числу въехавших и моделирование перекрытия одной полосы на съезде с участка автомагистрали.

Данные с дорожных датчиков и полученные ручной съемкой представляют собой последовательность проехавших по участку автотрассы АТС за каждый фиксированный интервал времени. Более детально данные с дорожных датчиков и их недостатки описаны в работе [1].

Все графики в данном разделе имеют вид $n(k)$, где n — число проехавших автомобилей за временной интервал τ с порядковым номером k , т. е. время от начала моделирования имеет вид τk . Как критерий качества выбрана величина среднеквадратичной ошибки предсказанной величины n :

$$S = \sqrt{\frac{1}{K} \sum_{k=1}^K (n(k) - \bar{n}(k))^2},$$

где K — число временных интервалов; \bar{n} — зафиксированное ручной съемкой либо дорожными датчиками число проехавших по участку магистрали АТС. Временной интервал τ в данной работе равен 2 мин.

Алгоритм 1 Алгоритм моделирования автомагистрали

Вход: $\mathbf{In} \in \tilde{\mathbf{G}}$; T — время конца моделирования; j' — вершина начала моделирования;

$\mathbf{V}' = \mathbf{V}$ — множество непросмотренных вершин; $k = -1$.

Выход: $\forall k \tau \forall (i, j) \in \mathbf{E}$ множества $\bar{\mathbf{A}}_{i,j}$

пока $\tau k < T$

$\mathbf{V}' = \mathbf{V}$; $j = j'$; $k = k + 1$

пока $|\mathbf{V}'| \neq 0$

Создаем новую группу АТС $\mathbf{A}_l \in \bar{\mathbf{A}}_{i,j}$, $l = |\bar{\mathbf{A}}_{i,j}| + 1$, $V_l = 0$, $\text{Pos}_l = 0$

если $\text{Tr}'_{i,k} + \text{In}_{i,k} + \sum_{s=1}^{|\bar{\mathbf{A}}_{i,j}|} N_s < N_{\max,(i,j)}$, **то**

$N_l = \text{Tr}'_{i,k} + \text{In}_{i,k}$

иначе

$N_l = N_{\max,(i,j)} - \sum_{s=1}^{|\bar{\mathbf{A}}_{i,j}|} N_s$;

$\text{In}_{i,k+1} = \text{In}_{i,k+1} + (\text{Tr}'_{i,k} + \text{In}_{i,k} + \sum_{s=1}^{|\bar{\mathbf{A}}_{i,j}|} N_s - N_{\max,(i,j)})$

для всех $\mathbf{A}_l \in \bar{\mathbf{A}}_{i,j}$, $l \in [1, |\bar{\mathbf{A}}_{i,j}|]$

$N_{\text{tmp}} = 0$; $V_l = f_{i,j}(\rho_l)$ (1).

Рассчитываем новое положение группы АТС как

если $\exists \text{Pos}_{l+1}$, **то**

$$\text{Pos}_l = \begin{cases} \text{Pos}_l + V_l \tau, & \text{Pos}_l + V_l \tau < \text{Pos}_{l+1} - ((N_l + N_{l+1})/2)L_{\text{avg}}; \\ \text{Pos}_{l+1} - ((N_l + N_{l+1})/2)L_{\text{avg}}, & \text{Pos}_l + V_l \tau \geq \text{Pos}_{l+1} - ((N_l + N_{l+1})/2)L_{\text{avg}}, \end{cases}$$

иначе

$$\text{Pos}_l = \begin{cases} \text{Pos}_l + V_l \tau, & \text{Pos}_l + V_l \tau < L_{i,j}; \\ L_{i,j}, & \text{Pos}_l + V_l \tau \geq L_{i,j} \end{cases}$$

если $\text{Pos}_l = L_{i,j}$ и $l = |\bar{\mathbf{A}}_{i,j}|$, **то**

если $b_{l,i,j} = 0$, **то**

$N_{\text{exit},j} = N_{\text{exit},j} + M_{j,k} N_l$; $N_l = (1 - M_{j,k}) N_l$; $b_{l,i,j} = 1$

если $N_{\text{exit},m} + \sum_{s=1}^{|\bar{\mathbf{A}}_{j,m}|} N_s + N_{\text{tmp}} < N_{\max,(j,m)}$, **то**

$N_{\text{proc}} = N_l$

иначе

$N_{\text{proc}} = N_{\max,(j,m)} - (N_{\text{exit},m} + \sum_{s=1}^{|\bar{\mathbf{A}}_{j,m}|} N_s + N_{\text{tmp}})$

если $N_{\text{proc}} > \text{Tr}_j$, **то**

$N_{\text{proc}} = \text{Tr}_j$

$N_{\text{tmp}} = N_{\text{tmp}} + N_{\text{proc}}$; $\text{Tr}'_{j,k+1} = \text{Tr}'_{j,k+1} + N_{\text{proc}}$; $N_l = N_l - N_{\text{proc}}$

если $N_l = 0$, **то**

Удаляем группу АТС \mathbf{A}_l

если $\text{Pos}_{l+1} - \text{Pos}_l \leq ((N_l + N_{l+1})/2)L_{\text{avg}}$ и $N_l + N_{l+1} \leq \varepsilon_n$, **то**

если $\text{Pos}_{l+1} = L_{i,j}$ и $l + 1 = |\bar{\mathbf{A}}_{i,j}|$, **то**

$N_{\text{exit},j} = N_{\text{exit},j} + M_{j,k} N_l$; $N_l = (1 - M_{j,k}) N_l$; $b_{l,i,j} = 1$

$N_l = N_l + N_{l+1}$; $V_l = V_{l+1}$; $\text{Pos}_l = \text{Pos}_{l+1}$

Удаляем группу АТС \mathbf{A}_{l+1}

$\mathbf{V}' = \mathbf{V}' \setminus j$, $i = j$

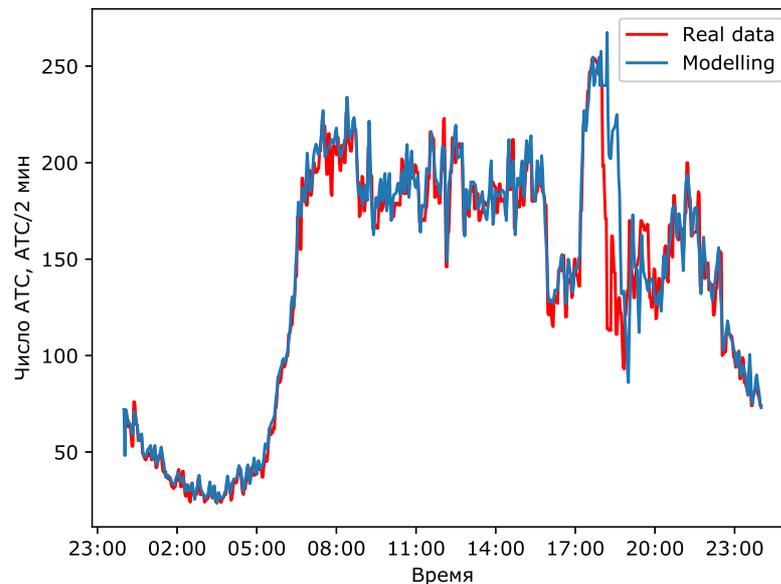


Рис. 1 График полученного с помощью модели числа съехавших АТС (красная линия) в сравнении с числом съехавших АТС, зафиксированных дорожным датчиком (синяя линия), за один день. Среднеквадратичная ошибка $S = 18,4$

3.1 Эксперимент с данными дорожных датчиков

В эксперименте использовали построенную для данного участка автомагистрали фундаментальную диаграмму [13]. В приведенной статье диаграмма строилась на тех же данных, которыми авторы оперируют в этой. Полный временной интервал эксперимента — одна неделя, графики приведены за один день. В первом эксперименте проводится проверка результатов модели в простейшем случае моделирования числа съехавших АТС по числу въехавших на участке автомагистрали без въездов и съездов. Результаты показаны на рис. 1. Среднеквадратичная ошибка $S = 18,4$.

3.2 Эксперимент с перекрытием полосы

Во втором эксперименте проводится моделирование ситуации, когда одна из полос на автомагистрали перекрывается. Сам эксперимент проводится на том же участке автомагистрали и за тот же промежуток времени, что и первый в подразд. 3.1. Данных для расчета среднеквадратичной ошибки у авторов нет, и этот эксперимент был поставлен, чтобы проанализировать поведение модели в такой ситуации. Результаты за тот же день, что и в первом эксперименте, можно увидеть на рис. 2.

Видно, что АТС не могут съехать из-за ограничения на максимальное число проезжающих с одного ребра на другое автомобилей, что приводит к появлению горизонтальной линии на графике. Однако через некоторое время после того, как поток должен был спастись, что видно на графике реальных данных за тот же временной промежуток, дорога освобождается и результат моделирования приходит в соответствие с реальными данными.

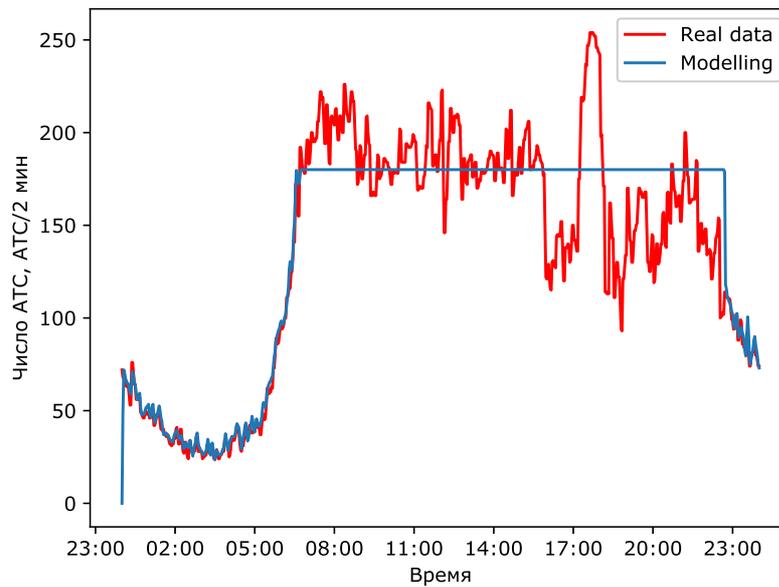


Рис. 2 График полученного с помощью модели числа съехавших АТС (красная линия) в сравнении с числом съехавших АТС зафиксированных дорожным датчиком (синяя линия), за один день. Одна полоса перекрыта

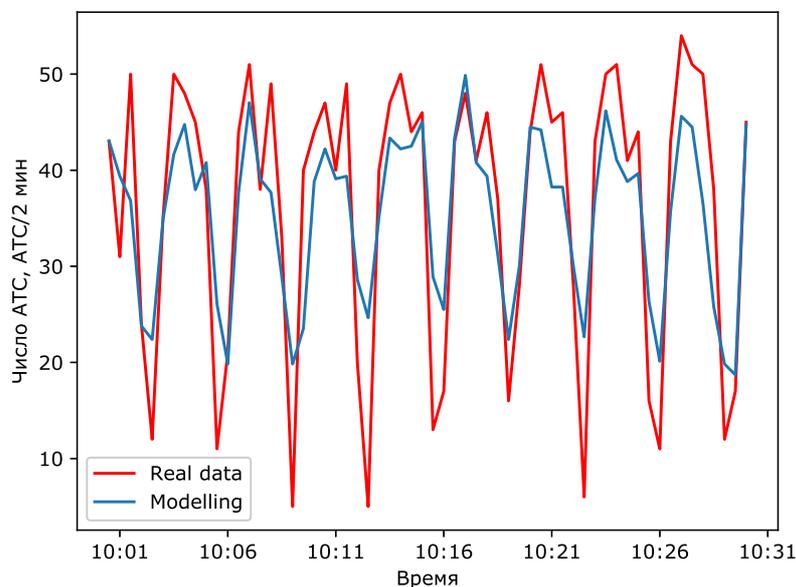


Рис. 3 График полученного с помощью модели числа съехавших АТС (красная линия) в сравнении с числом съехавших АТС зафиксированных видеосъемкой (синяя линия) за 30 мин. Среднеквадратичная ошибка $S = 8,08$

3.3 Эксперимент с данными видеосъемки

В данном эксперименте на въезде и съезде с выбранного участка автомагистрали стоят светофоры, что учитывается при анализе результатов эксперимента. Временной интервал

эксперимента — 30 мин. Из-за небольших объемов имеющихся данных авторы не имели возможности построить фундаментальную диаграмму для данного участка автомагистрали и воспользовались диаграммой для МКАД из первого эксперимента. Полученный результат приведен на рис. 3. В данном эксперименте из-за светофора на съезде число съехавших АТС за некоторые временные интервалы практически равен нулю. В то же время из-за светофора на въезде модель показывает, что даже не будь светофора, на въезде число автомобилей, подъезжающее к съезду в данные моменты времени, также невелико. Оказывается, что светофоры синхронизированы. Среднеквадратичная ошибка $S = 8,08$.

4 Обсуждение результатов

В работе изложен новый алгоритм моделирования числа проехавших АТС для задачи моделирования транспортных потоков с использованием данных из разнородных источников [1] и проведены три эксперимента, показывающих его состоятельность.

Эксперимент из подразд. 3.1 показывает работоспособность модели для моделирования прямого участка автомагистрали. На рис. 1 видно, что модель адекватно симулирует дорожную ситуацию при любом числе проезжающих АТС.

В эксперименте в подразд. 3.2 исследуется реакция модели на перекрытие одной из полос автомагистрали. То, что после спада потока АТС прогноз модели снова стал совпадать с реальными данными, показывает ее состоятельность в данной ситуации.

Эксперимент на данных видеосъемки из подразд. 3.3 не такой наглядный, как из подразд. 3.1, из-за светофора на съезде, наличие которого нельзя учесть в предлагаемой модели. Однако даже на нем модель ведет себя адекватно, во многом из-за того что светофоры на въезде понижают поток АТС так, что во время перекрытия съезда поток подъезжающих к нему АТС минимален.

5 Благодарности

Авторы благодарят ООО «Яндекс» за предоставленные анонимные трековые данные, Центр организации дорожного движения за данные датчиков движения. Авторы также чрезвычайно признательны Я. А. Холодову и А. Е. Алексеенко за консультации и обсуждение результатов.

Литература

- [1] *Старожилец В. М., Чехович Ю. В.* Комплексирование данных из разнородных источников в задачах моделирования транспортных потоков // *Машинное обучение и анализ данных*, 2016. Т. 2. № 3. С. 260–276.
- [2] *Lighthill M.J., Whitham G.B.* On kinematic waves. II. A theory of traffic flow on long crowded roads // *P. Roy. Soc. Lond. A Mat.*, 1955. Vol. 229. P. 317–345. doi: 10.1098/rspa.1955.0089.
- [3] *Richards P.I.*, Shock waves on the highway // *Oper. Res.*, 1956. Vol. 4. No. 1. P. 42–51. doi: 10.1287/opre.4.1.42.
- [4] *Whitham J.B.* Linear and nonlinear waves. — Wiley, 1974. 656 p.

- [5] *Daganzo C. F.* Requiem for second-order fluid approximations of traffic flow // *Transport. Res. B Meth.*, 1995. Vol. 29. No. 4. P. 277–286. doi: 10.1016/0191-2615(95)00007-Z.
- [6] *Payne H. J.* Models of freeway traffic and control // *Math. Models Public Syst.*, 1998. No. 4. P. 51–61.
- [7] *Papageorgiou M.* Some remarks on macroscopic traffic flow modelling // *Transport. Res. A Pol.*, 1998. Vol. 32. No. 5. P. 323–329. doi: 10.1016/S0965-8564(97)00048-7.
- [8] *Aw A., Michel Rasche M.* Resurrection of “second order” models of traffic flow // *SIAM J. Appl. Math.*, 2000. Vol. 60. No. 3. P. 916–938. doi: 10.1137/S0036139997332099.
- [9] *Zhang M.* A non-equilibrium traffic model devoid of gas-like behavior // *Transport. Res. B Meth.*, 2002. Vol. 36. No. 3. P. 275–290. doi: 10.1016/S0191-2615(00)00050-3.
- [10] *Zhang M.* Anisotropic property revisited — does it hold in multi-lane traffic? // *Transport. Res. B Meth.*, 2003. Vol. 37. No. 6. P. 561–577. doi: 10.1016/S0191-2615(02)00030-9.
- [11] *Siebel F., Mauser W.* On the fundamental diagram of traffic flow // *SIAM J. Appl. Math.*, 2006. Vol. 66. No. 4. P. 1150–1162. doi: 10.1137/050627113.
- [12] *Siebel F., Mauser W.* Synchronized flow and wide moving jams from balanced vehicular traffic // *Phys. Rev. E*, 2006. Vol. 73. No. 6. P. 066108. doi: 10.1103/PhysRevE.73.066108.
- [13] *Алексеевко А.Е., Холодов Я.А., Холодов А.С., Горева А.И., Васильев М.О., Чехович Ю.В., Мишин В.Д., Старожилец В.М.* Разработка, калибровка и верификация модели движения трафика в городских условиях. Ч. I // *Компьютерные исследования и моделирование*, 2015. Т. 7. №6. С. 1185–1203.
- [14] *Wageningen-Kessels V., Hoogendoorn S. P., Vuk C., et al.* Traffic flow modeling: A genealogy. — TRB, 2014. 156 p.
- [15] *Kotsialos A., Papageorgiou M., Diakaki C., Pavlis Y., Middelham F.* Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET // *IEEE T. Intell. Transp.*, 2002. Vol. 3. No. 4. P. 282–292. doi: 10.1109/TITS.2002.806804.
- [16] *Kerner B.* The physics of traffic. — Berlin: Springer, 2004. 681 p. doi: 10.1007/978-3-540-40986-1.

Поступила в редакцию 17.09.2017

About identification of a statistical model of traffic flows using vehicle groups

V. M. Starozhilets^{1,2} and Yu. V. Chehovich^{1,2}

starvsevol@gmail.com; chehovich@forecsys.ru

¹Federal Research Center “Computer Science and Control” of RAS, 44/2 Vavilova Str., Moscow, Russia

²Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow, Russia

A statistical model of traffic flows for modeling speed and number of cars on highways identified on data from heterogeneous sources is proposed. The model simulates movement of car groups along the highway using corresponding to the selected road segment fundamental diagram to calculate the car groups speed. Computational experiments are provided to confirm the

adequateness of the model. Also, its behavior in situation of blocking one of the lanes of the highway is analyzed. The criterion of quality is the root-mean-square error between the predicted number of passed vehicles and the actual number of vehicles. Data from traffic detectors, provided by Traffic Management Center, and data obtained by video recording are used in this study.

Keywords: *transport flow modelling; Moscow Ring Road; traffic detectors*

DOI: 10.21469/22233792.3.3.02

References

- [1] Starozhilets, V. M., and Yu. V. Chekhovich. 2016. Kompleksirovanie dannykh iz raznorodnykh istochnikov v zadachakh modelirovaniya transportnykh potokov [Aggregation of data from different sources in traffic flow tasks]. *Machine Learning Data Anal.* 2(3):260–276.
- [2] Lighthill, M. J., and G. B. Whitham. 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads // *P. Roy. Soc. Lond. A Mat.* 229:281–345. doi: 10.1098/rspa.1955.0089.
- [3] Richards, P. I. 1956. Shock waves on the highway. *Oper. Res.* 4(1):42–51. doi: 10.1287/opre.4.1.42.
- [4] Whitham, J. B. 1974. *Linear and nonlinear waves*. Wiley. 656 p.
- [5] Daganzo, C. F. 1955. Requiem for second-order fluid approximations of traffic flow. *Transport. Res. B Meth.* 29(4):277–286. doi: 10.1016/0191-2615(95)00007-Z.
- [6] Payne, H. J. 1998. Models of freeway traffic and control. *Math. Models Public Syst.* 4:51–61.
- [7] Papageorgiou, M. 1988. Some remarks on macroscopic traffic flow modelling. *Transport. Res. A Pol.* 32(5):323–329. doi: 10.1016/S0965-8564(97)00048-7.
- [8] Aw, A., and M. Michel Rasclé. 2000. Resurrection of “second order” models of traffic flow. *SIAM J. Appl. Math.* 60(3):916–938. doi: 10.1137/S0036139997332099.
- [9] Zhang, M. 2002. A non-equilibrium traffic model devoid of gas-like behavior. *Transport. Res. B Meth.* 36(3):275–290. doi: 10.1016/S0191-2615(00)00050-3.
- [10] Zhang, M. 2003. Anisotropic property revisited — does it hold in multi-lane traffic? *Transport. Res. B Meth.* 37(6):561–577. doi: 10.1016/S0191-2615(02)00030-9.
- [11] Siebel, F., and W. Mauser. 2006. On the fundamental diagram of traffic flow. *SIAM J. Appl. Math.* 66(4):1150–1162. doi: 10.1137/050627113.
- [12] Siebel, F., and W. Mauser. 2006. Synchronized flow and wide moving jams from balanced vehicular traffic. *Phys. Rev. E* 73(6):066108. doi: 10.1103/PhysRevE.73.066108.
- [13] Alekseenko, A. E., Yu. A. Kholodov, A. S. Kholodov, A. I. Goreva, M. O. Vasil’ev, Yu. V. Chekhovich, V. D. Mishin, and V. M. Starozhilets. 2015. Razrabotka, kalibrovka i verifikatsiya modeli dvizheniya trafika v gorodskikh usloviyakh. Ch. I [Development, calibration, and verification of mathematical model for multilane urban road traffic flow. Part I]. *Komp’yuternye issledovaniya i modelirovanie* [Computer Research and Modeling] 7(6):1185–1203.
- [14] Wageningen-Kessels, V., S. P. Hoogendoorn, C. Vuik, *et al.* *Traffic flow modeling: A genealogy*. TRB, 2014. 156 p.

- [15] Kotsialos, A., M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham. 2002. Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET. *IEEE T. Intell. Transp.* 3(4):282–292. doi: 10.1109/TITS.2002.806804.
- [16] Kerner, B. 2004. *The physics of traffic*. Berlin: Springer. 681 p. doi: 10.1007/978-3-540-40986-1.

Received September 17, 2017

Дескриптор в аккумуляторном пространстве Хафа градиентного поля изображения для детектирования пешеходов*

Н. А. Самсонов¹, А. Н. Гнеушев²

nikita.samsonov@phystech.edu; gneushev@ccas.ru;

¹Московский физико-технический институт, Россия, г. Долгопрудный, Институтский пер., 9

²ФИЦ «Информатика и управление» РАН, Россия, г. Москва, ул. Вавилова, 44/2

Рассматривается задача выделения признаков для распознавания пешеходов на изображении. Наиболее популярный и эффективный подход к выбору признаков для распознавания объекта на изображении — использование дескриптора на основе гистограмм ориентаций градиентов (Histograms of Oriented Gradients, HOG). В работе предлагается использовать аккумуляторное пространство Хафа для обобщения HOG дескриптора путем получения проекций не только ориентаций, но и положений границ на локальном участке изображения — гистограмм аккумуляторного пространства Хафа (Hough Accumulator Histograms, HAH). Аккумуляторное пространство Хафа строится на основе лучевого преобразования Радона градиентного поля изображения. Предлагаемые методы были протестированы совместно с линейным классификатором метода опорных векторов (Support Vector Machine, SVM) на базе изображений пешеходов INRIA. Результаты экспериментов показали лучшую разделяющую способность новых дескрипторов, уменьшение ложных срабатываний детекторов в сравнении с HOG.

Ключевые слова: *histograms of oriented gradients, HOG; Hough accumulator histograms, HAH; дескриптор; пешеходы; преобразование Радона; преобразование Хафа*

DOI: 10.21469/22233792.3.3.03

1 Введение

Автоматическое детектирование объектов на изображениях и видео является одной из основных задач компьютерного зрения. Цель детектирования — определить наличие и расположение искомого объекта на видеоизображении для последующего анализа. Такие объекты на изображении, как люди, являются неотъемлемой частью дорожной сцены, играют важную роль в развитии дорожной ситуации. Задача локализации человека на видеоизображениях широко востребована в таких областях, как мониторинг и анализ дорожных ситуаций, обнаружение дорожно-транспортных происшествий, контроль за соблюдением правил дорожного движения, системы безопасности и следящие системы, беспилотные автомобили, робототехника, системы помощи водителю.

Распознавание человека на изображении — сложная задача. Ее сложность обусловлена несколькими причинами: изменение структуры изображения вследствие движения объекта, неравномерная освещенность изображения, большая вариабельность изображений человека из-за разных ракурсов (поз, размеров, углов поворота), частичные перекрытия фигуры человека другими объектами.

Как правило, множество признаков, которые используются для распознавания объекта, определяются его характерной структурой на изображении. Изображение человека

*Работа выполнена при финансовой поддержке РФФИ, грант №16-07-01171.

может быть представлено совокупностью контуров, силуэтов частей тела, которые являются контурными признаками и представляются на изображении как максимальные перепады значений яркости [1]. Однако контуры на изображении человека могут иметь достаточно сложную и разнообразную структуру, и для описания этой структуры необходимо учитывать совокупность контурных признаков — локальных дескрипторов для описывания изображения объекта в целом [2, 3]. На сегодняшний день наиболее популярны подходы, основанные на градиентах яркости в локальных областях изображения и учитывающие совокупность контурных признаков — локальных дескрипторов для описания изображения объекта в целом. Так, в подходе, описанном в [4, 5], строится каскад бинарных классификаторов, где признаки, напоминающие базисные вейвлеты Хаара, вычисляются по локальным прямоугольным областям изображения. Однако детекторы, основанные на данном методе, являются неустойчивыми по отношению к таким факторам, как неравномерное изменение освещения и вариабельность изображения объекта. Эффективная схема построения интегрального вектора признаков для задачи детектирования человека на основе локальных контурных дескрипторов — гистограмм ориентаций градиентов в локальных областях изображения — используется в методе HOG [6] и подходах, которые его развивают [7, 8]. HOG-дескриптор обладает рядом преимуществ: он устойчив к неравномерному освещению, вычислительно эффективен, показывает лучшие результаты для задачи детектирования пешеходов, однако он недостаточно полно описывает локальные контурные особенности объекта, используя лишь информацию об ориентации, вследствие чего детектор на его основе имеет неудовлетворительную для многих реальных приложений долю ложных срабатываний при приемлемом уровне пропуска. Гистограммы ориентаций градиентов можно рассматривать как аппроксимацию проекций модуля градиента на основные направления. Так, в работах [9, 10] предлагаются методы на основе преобразования Радона для детектирования сложных фигур на изображении. Описанные в работах дескрипторы строятся как проекции в пространстве преобразования Радона бинарных изображений, являются инвариантными относительно поворота, переноса и масштабирования, показывая при этом хорошие результаты, однако для полутоновых изображений результатов получено не было.

В данной работе предлагается обобщить HOG-дескриптор с помощью определения нового дескриптора — НАН — для детектирования пешеходов на изображениях. Для построения аккумуляторного пространства Хафа используется подход на основе лучевого поперечного преобразования Радона градиентного поля изображения. Использование проекций в аккумуляторном пространстве Хафа с целью описания изображения позволяет учитывать в НАН-дескрипторе не только локальное распределение ориентаций контурных признаков, но также и их положение. Таким образом, предложенный метод преодолевает недостатки [6], показывая, что использование дополнительной информации о положении локальных контурных особенностей объекта позволяет превзойти HOG по качеству детектирования и уменьшить уровень пропуска при приемлемом числе ложных срабатываний.

2 Постановка задачи

Определим задачу построения признакового пространства G как получение интегрального вектора признаков $\mathbf{g}(f) = (g_1, \dots, g_n)^T$ для входного изображения f , где n — размерность пространства признаков. Вектор признаков $\mathbf{g}(f)$ будем строить на основе общей схемы метода HOG [6], используя классификацию методом SVM при построении эффективного детектора пешехода. Критерием эффективности детектора будем считать уровень ошибок 1-го и 2-го рода на тестовой выборке пешеходов INRIA.

В соответствии со схемой НОГ [6], изображение $f(x, y)$, где x, y — координаты точки изображения размером $X \times Y$, разбивается на смежные области — квадратные ячейки размером $K \times K$ пикселей, $I = X/K$ ячеек по горизонтали и $J = Y/K$ ячеек по вертикали. В каждой ячейке (i, j) , где i, j — индексы ячейки по вертикали и горизонтали соответственно, выделяются локальные признаки, которые определяются дескриптором ячейки — вектором $\mathbf{u}_{i,j} = (u_1, \dots, u_l)^T$, где l — размерность вектора. Четыре смежные ячейки объединяются в пересекающиеся блоки. Таким образом, каждый блок содержит $4K \times 4K$ пикселей и имеет общие ячейки с соседними блоками. Дескриптор блока определяется объединением дескрипторов собственных ячеек, вектором $\mathbf{v}_{i,j} = \mathbf{u}_{i,j} \cup \mathbf{u}_{i+1,j} \cup \mathbf{u}_{i,j+1} \cup \mathbf{u}_{i+1,j+1}$, где i, j — индексы блока. Общее количество блоков — $(I - 1)(J - 1)$. Под операцией объединения \cup двух векторов \mathbf{u}_1 и \mathbf{u}_2 с размерностью l_1 и l_2 соответственно будем понимать результирующий вектор из пространства размерности $l_1 + l_2$, первые l_1 компонент которого являются компонентами вектора первого аргумента, последние l_2 — компонентами вектора второго аргумента.

Дескриптор блока нормируется с помощью одной из двух норм: L_2 нормы

$$\tilde{\mathbf{v}} = N_{L_2}(\mathbf{v}) = \frac{\mathbf{v}}{\sqrt{\|\mathbf{v}\|_2^2 + \varepsilon^2}} \tag{1}$$

либо L_2 -hys нормы [6]

$$\tilde{\mathbf{v}} = N_{L_2\text{-hys}}(\mathbf{v}) = N_{L_2}(\min(N_{L_2}(\mathbf{v}), h)), \tag{2}$$

где операция \min применяется покомпонентно к вектору первого аргумента; h — пороговое значение, которое используется для ограничения значений компонент вектора в операции \min . Данное пороговое ограничение определено с целью фильтрации шумовых всплесков, которые могут появляться в компонентах дескриптора. В данной работе, как и в работе [6], в качестве порогового значения используется $h = 0,2$.

Интегральный дескриптор определяется объединением всех блочных дескрипторов, т. е. вектором

$$\mathbf{g} = \bigcup_j^{J-1} \bigcup_i^{I-1} \tilde{\mathbf{v}}_{i,j}.$$

Интегральные дескрипторы \mathbf{g} множества изображений формируют признаковое пространство G . Для решения задачи детектирования пешеходов необходимо определить разбиение этого пространства на два непересекающихся класса: первый характеризует пешеходов; второй — фон, не содержащий пешехода. Ограничимся использованием линейного классификатора, который задается выражением:

$$s(\mathbf{g}) = \text{sign} \left(\sum_{j=1}^n w_j g_j - w_0 \right) = \text{sign} (\langle \mathbf{w}, \mathbf{g} \rangle - w_0),$$

где вектор $\mathbf{w} = (w_1, \dots, w_n)^T \in R^n$ и скалярный порог $w_0 \in R$ определяют разделяющую гиперплоскость $\langle \mathbf{w}, \mathbf{g} \rangle = w_0$ в пространстве G .

Вектор весов \mathbf{w} классификатора и скалярный порог w_0 находятся из процедуры обучения на основе метода SVM [6] по специально подготовленной обучающей выборке из базы изображений INRIA, содержащий два подмножества изображений: с положительными примерами, содержащими пешеходов, и отрицательными примерами, содержащими фон.

Критерием качества классификации $s(\mathbf{g})$ на специально подготовленной тестовой выборке из базы изображений INRIA будем считать отношение $MR = FN / (TP + FN)$ — доля неверно отвергнутых классификатором изображений (Miss Rate), к $FPPW = FP / (TN + FP)$ — доля неверно принятых изображений (False Positive Per Window), где FN — количество неверно отвергнутых классификатором положительных примеров; TP — количество верно классифицированных положительных примеров; TN — количество верно классифицированных отрицательных примеров; FP — количество неверно классифицированных отрицательных примеров.

3 Построение текстурных признаков для детектирования пешеходов в аккумуляторном пространстве Хафа

3.1 Лучевое поперечное преобразование Радона градиентного поля изображения

Рассмотрим преобразование Радона [11] функции $f \in L_2(\mathbb{R}^2)$ вдоль прямой $L : \mathbf{x} = \rho \mathbf{n} + t \mathbf{a}$, которое определяется формулой:

$$\mathcal{R}_{\alpha, \rho}(f) = \int_L f(\rho \mathbf{n} + t \mathbf{a}) dt, \quad (3)$$

где t — параметр прямой L ; $\mathbf{n} = (\cos \alpha, \sin \alpha)^T$ и $\mathbf{a} = (-\sin \alpha, \cos \alpha)^T$ — нормаль и направляющий вектор прямой; α задает ориентацию вектора нормали \mathbf{n} ; $\rho = x \cos \alpha + y \sin \alpha$ — модуль радиуса-вектора ближайшей к началу координат точки на прямой. Таким образом, α и ρ — параметры прямой L , задающие пространство Радона \mathcal{R} .

Преобразование Радона обладает свойством обратимости [11]. Таким образом, при переходе в пространство Радона информация о функции f не теряется.

На основе преобразования Радона (3) определим лучевое поперечное преобразование [11] градиента функции $\nabla f \in L_2(\mathbb{R}^2)$ (рис. 1):

$$\mathcal{R}_{\alpha, \rho}^{\perp}(\nabla f) = \int_L \langle \mathbf{n}, \nabla f(\rho \mathbf{n} + t \mathbf{a}) \rangle dt. \quad (4)$$

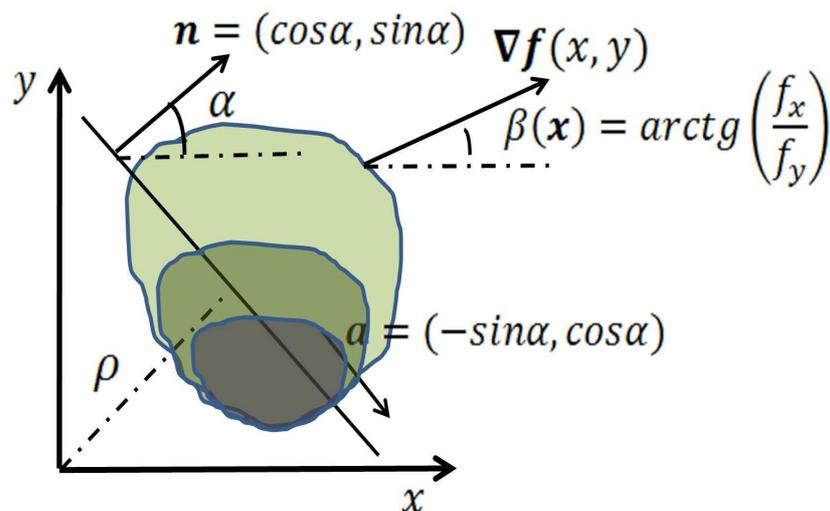


Рис. 1 Преобразование Радона градиентного поля

Выражение (4) можно переписать в виде:

$$\begin{aligned} \mathcal{R}_{\alpha,\rho}^{\perp}(\nabla f) &= \int_L \langle \mathbf{n}, \nabla f(\rho\mathbf{n} + t\mathbf{a}) \rangle dt = \int_L |\nabla f(\rho\mathbf{n} + t\mathbf{a})| \cos(\alpha - \beta(\mathbf{x})) dt = \\ &= \int_L |\nabla f(\rho\mathbf{n} + t\mathbf{a})| q(\alpha, \beta(\mathbf{x})) dt, \end{aligned} \quad (5)$$

где $\nabla f(\mathbf{x})$ — вектор градиента в точке $\mathbf{x} = (x, y)^T$; $\beta(\mathbf{x})$ — угол, задающий ориентацию градиента в точке \mathbf{x} ; $\mathbf{n} = (\cos \alpha, \sin \alpha)^T$; $\mathbf{a} = (-\sin \alpha, \cos \alpha)^T$ — нормаль и направляющий вектор прямой L соответственно; α задает ориентацию вектора нормали \mathbf{n} ; $q(\alpha, \beta(\mathbf{x})) = \cos(\alpha - \beta(\mathbf{x}))$ — функция, задающая вклад модуля градиента в точке \mathbf{x} в направлении α .

3.2 Гистограммные признаки в аккумуляторном пространстве Хафа

Рассмотрим дискретный случай лучевого поперечного преобразования Радона (5). Разобьем пространство параметров α, ρ на ячейки, для чего зададим разбиение осей на бины: $k = 1, \dots, N$ — индексы разбиения по оси α ; $l = 1, \dots, M$ — индексы разбиения по оси ρ . Получим множество ячеек $\{(\alpha_k, \rho_l)\}$, задающее аккумуляторное пространство Хафа (рис. 2), где (α_k, ρ_l) — центр ячейки с соответствующими индексами. Аккумуляторное пространство Хафа для поля градиентов ∇f строится путем сбора двумерной гистограммы. Из формулы преобразования Радона (5) для значения ячейки (α_k, ρ_l) аккумуляторного пространства получаем выражение:

$$\mathcal{H}_{\alpha_k, \rho_l}^{\perp}(\nabla f) = \sum_{x,y} |\nabla f(x, y)| q(\alpha_k, \beta(\mathbf{x})) [\rho_l = x \cos \alpha_k + y \sin \alpha_k], \quad (6)$$

где операция $[\cdot]$ возвращает индикаторную величину, равную 0, если выражение в скобках не выполняется, и 1 — в противном случае.

Аккумуляторное пространство Хафа, описываемое выражением (6), можно рассматривать в качестве пространства, характеризующего контурные признаки изображения, и строить на его основе дескрипторы. Определим локальные дескрипторы ячеек изображения в аккумуляторном пространстве Хафа. Из формулы (6) следует, что для локальной

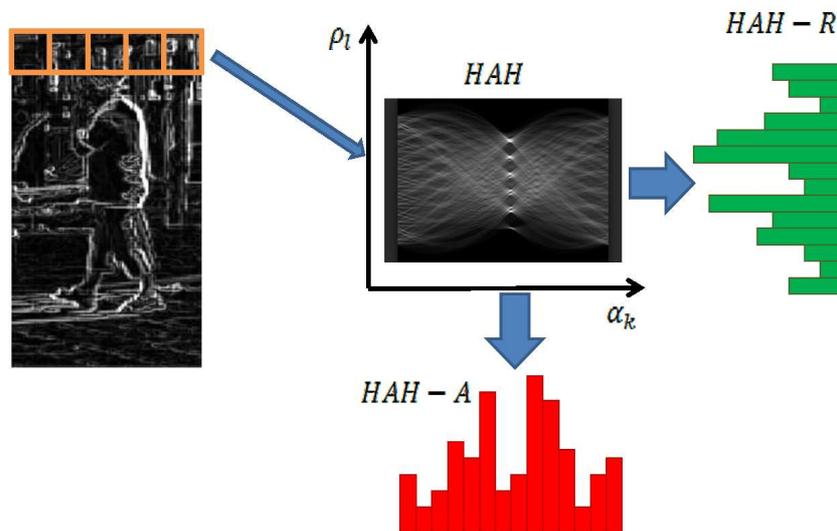


Рис. 2 Преобразование Хафа градиентного поля

ячейки (i, j) изображения значение в точке (α_k, ρ_l) признакового пространства описывается формулой:

$$\mathcal{H}_{\alpha_k, \rho_l}(i, j, \nabla f) = \sum_{(x, y) \in \text{Cell}_{i, j}} |\nabla f(x, y)| q_{\mathcal{H}^\perp}(\alpha_m, \beta(\mathbf{x})) [\rho_l = x \cos \alpha_k + y \sin \alpha_k]. \quad (7)$$

В качестве функции вклада $q_{\mathcal{H}^\perp}(\alpha_m, \beta(\mathbf{x}))$ будем использовать функцию:

$$q_{\mathcal{H}^\perp}(\alpha_m, \beta(\mathbf{x})) = |\cos(\alpha_k - \beta(\mathbf{x}))|, \quad (8)$$

где модуль берется для того, чтобы вклады противоположных по направлению градиентов яркости на изображении учитывались одинаково.

Для уменьшения размерности пространства признаков (7) построим его проекции в пространство одного из параметров.

Определение 1. НАН-А дескриптор является проекцией пространства (7) на пространство параметра α и определяется выражением:

$$\mathbf{u}_{i, j}^{\text{НАН-А}} = \{u_{i, j}^k\}_{k=1}^N, \quad u_{i, j}^k = \sum_l \mathcal{H}_{\alpha_k, \rho_l}(i, j). \quad (9)$$

Полученный дескриптор совпадает по размерности с НОГ [6], но отличается от него видом интерполяционной функции.

Утверждение 1. НАН-А дескриптор $\mathbf{u}_{i, j}^{\text{НАН-А}}$, определяемый выражением (9) с функцией вклада

$$q(\alpha_k, \beta) = \left(1 - \frac{|\alpha_k - \beta(x, y)|}{\alpha_{k+1} - \alpha_k}\right) [|\alpha_k - \beta(x, y)| < (\alpha_{k+1} - \alpha_k)], \quad (10)$$

эквивалентен локальному дескриптору ячейки НОГ [6].

Для доказательства утверждения достаточно заметить, что функция вида (10) для q в сумме (9) является весовой функцией, линейно интерполирующей модуль градиента между соседними бинами параметра α . Именно данная интерполяция используется в дескрипторе НОГ.

Определение 2. НАН-R дескриптор является проекцией пространства (7) на пространство параметра ρ и определяется выражением:

$$\mathbf{u}_{i, j}^{\text{НАН-R}} = \{u_{i, j}^l\}_{l=1}^M, \quad u_{i, j}^l = \sum_k \mathcal{H}_{\alpha_k, \rho_l}(i, j). \quad (11)$$

Данный дескриптор, в отличие от НОГ [6], учитывает только положение границ объектов внутри ячеек.

На основе дескрипторов, задаваемых выражениями (9) и (11) можно построить дескриптор, учитывающий как ориентацию, так и положение границ контуров внутри ячеек.

Определение 3. НАН-AR дескриптор является объединением обеих проекций пространства (7) и определяется выражением:

$$\mathbf{u}_{i, j}^{\text{НАН-AR}} = \mathbf{u}_{i, j}^{\text{НАН-А}} \cup \mathbf{u}_{i, j}^{\text{НАН-R}}. \quad (12)$$

3.3 Построение детектора пешеходов на основе гистограммных признаков в аккумуляторном пространстве Хафа

Построение детектора пешеходов включает две подзадачи: выделение гистограммных признаков в аккумуляторном пространстве Хафа (6) для построения дескриптора и его использование для бинарной классификации (рис. 3). Для выделения признаков на изображении будем пользоваться схемой HOG [6]. В качестве локальных дескрипторов ячеек будем использовать дескрипторы (9), (11) и (12). При использовании в качестве локального дескриптора ячейки HAH-AR на этапе нормализации гистограмм блоков, где блочный дескриптор обозначается как $v_{i,j}^{HAH-AR}$, будем нормировать следующим образом: сначала отдельно нормируются (1) и (2) гистограммы:

$$\begin{aligned} v_{i,j}^{HAH-A} &= u_{i,j}^{HAH-A} \cup u_{i+1,j}^{HAH-A} \cup u_{i,j+1}^{HAH-A} \cup u_{i+1,j+1}^{HAH-A}, \\ v_{i,j}^{HAH-R} &= u_{i,j}^{HAH-R} \cup u_{i+1,j}^{HAH-R} \cup u_{i,j+1}^{HAH-R} \cup u_{i+1,j+1}^{HAH-R}. \end{aligned}$$

Затем строится нормализованная гистограмма блока:

$$v_{i,j}^{HAH-AR} = u_{i,j}^{HAH-AR} \cup u_{i+1,j}^{HAH-AR} \cup u_{i,j+1}^{HAH-AR} \cup u_{i+1,j+1}^{HAH-AR}.$$

При использовании в качестве локальных дескрипторов ячейки HAH-A или HAH-R (9) или (11) нормализация гистограммы блока осуществляется по формулам (1) либо (2).

Для задачи классификации пешеходов на изображении будем применять линейный метод SVM, как и в [6]. Будем обучать алгоритм SVM в два этапа. На первом этапе обучение производится на выборке, составленной из базы размеченных изображений пешеходов (положительные примеры) и случайной выборки областей из изображений фона (отрицательные примеры). После первого этапа полученный классификатор используется для нахождения областей на изображениях фона, на которых детектор ошибается и ложно срабатывает. Эти области — «сложные примеры» — добавляются к обучающей выборке отрицательных примеров первого этапа. На втором этапе алгоритм SVM обучается второй раз на тех же положительных примерах и расширенной выборке отрицательных примеров для получения финального вектора весов линейного классификатора.

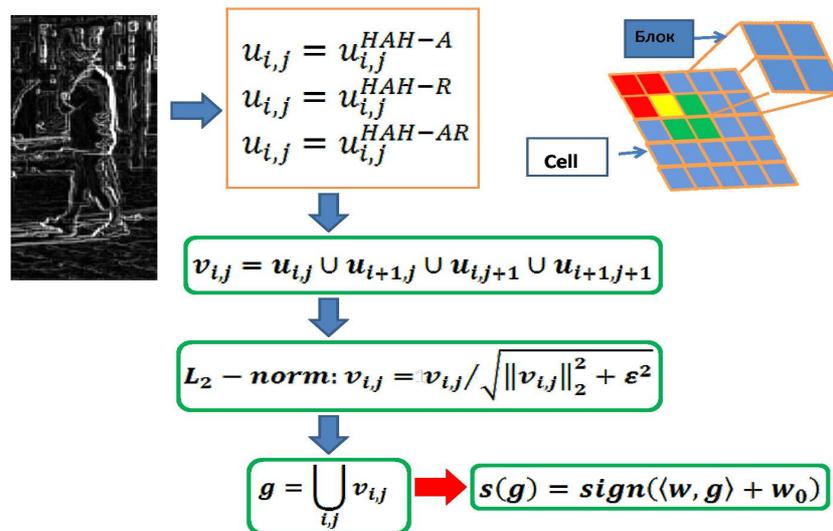


Рис. 3 Общая схема HAH-детектора

4 Результаты вычислительных экспериментов

Целью вычислительных экспериментов является проверка работы предлагаемых в статье дескрипторов на реальных данных, а также сравнение их качества детектирования и времени работы с аналогичными показателями HOG дескриптора [6].

Вычисления производились с помощью персонального компьютера с процессором Intel(R) Core(TM) i7 CPU с частотой 3,07 ГГц, 9 ГБ оперативной памяти, в системе ©Microsoft-Windows 7.

В качестве обучающей выборки была использована база изображений пешеходов INRIA [12]. База содержит 2478 изображений пешеходов размером 96×160 пикселей (рис. 4). Из каждого изображения выбирается центрированное окно размером 64 пикселей. Таким образом, 2478 окон составляют положительную часть обучающей выборки. В качестве изображений фона база данных содержит 1218 изображений (рис. 5) из которых в отрицательную часть обучающей выборки случайным образом выбирается 12 180 окон размером 64 пикселей. Число «сложных примеров» при использовании различных вариантов дескрипторов варьируется от 30 000 до 45 000.

В качестве локальных дескрипторов ячеек были использованы дескрипторы НАН-А, НАН-R и НАН-AR (9), (11) и (12), где множество параметра α аккумуляторного пространства Хафа разбивается на 9 бинов, а ρ — на 16 бинов. В зависимости от числа бинов интегральный дескриптор имеет следующую размерность: НАН-А — 3780, НАН-R — 6720 и НАН-AR — 10 500.

Для обучения классификатора была использована библиотека OpenCV [13]. Классификатор обучается с помощью линейного метода SVM [14]. В данных экспериментах в каче-



Рис. 4 Примеры позитивных изображений из обучающей выборки



Рис. 5 Примеры изображений фона из обучающей выборки, не содержащих пешеходов

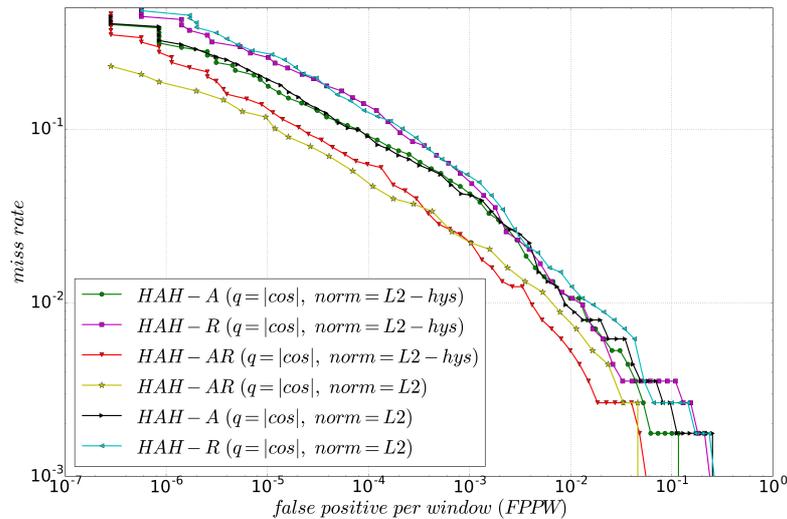


Рис. 6 Сравнение качества предлагаемых дескрипторов при различных вариантах нормализации гистограмм блоков

стве значений параметра метода C , который управляет балансом между максимизацией разделяющей полосы и суммарной ошибкой на объектах обучающей выборки, использовались значения из множества $\{0, 1, 10\}$. Выбиралось то значение параметра C , при котором достигался наилучший результат на тестовой выборке.

Полученные классификаторы были проверены на тестовой выборке изображений INRIA [12]. Тестовая выборка содержит 1132 изображения пешехода и 3 526 548 изображений фона размером 64×128 пикселей. Тестовые изображения фона были сформированы путем отбора областей из всевозможных окон с размерами 64×128 пикселей на 453 изображениях отрицательных примеров.

Для определения качества классификации были построены DET (detection error trade-off) кривые, где по оси абсцисс отложена доля неверно отвергнутых классификатором изображений (miss rate), а по оси ординат — доля неверно принятых решений (FPPW).

Для обнаружения влияния разных способов нормировки дескриптора при их обучении использовались два вида нормировки: L_2 (1) и L_2-hys (2), а в качестве функции вклада — функция (8).

На рис. 6 приведены DET кривые для различных вариантов нормализации дескрипторов блоков. Из него видно, что качество детектирования при использовании L_2 -нормы повышается на 32% при $10^{-4}FPPW$ по сравнению с L_2-hys нормализацией. Это связано с тем, что функция вклада (8) сглаживает вклад модуля градиента яркости по всем направлениям в каждой точке аккумуляторного пространства Хафа, что уменьшает вероятность всплесков. Таким образом, использование L_2 нормализации совместно с функцией вклада (8) имеет преимущества перед L_2-hys за счет того, что не теряется информативность признаков из-за порогового ограничения, определенного в функции (2) с целью фильтрации шумовых всплесков, которые могут появляться при использовании функции вклада (10).

На рис. 7 и в таблице отражены результаты для лучших вариантов предложенных в работе дескрипторов и НОГ-дескриптора [6].

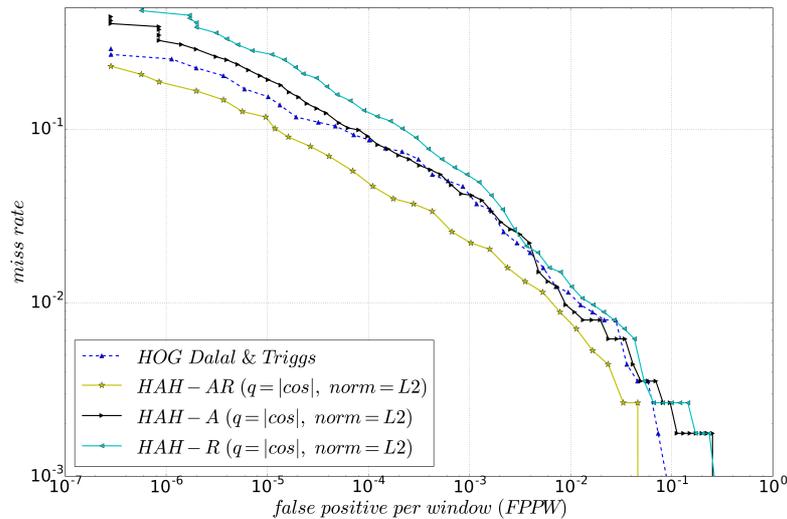


Рис. 7 Сравнение качества предлагаемых дескрипторов с HOG-дескриптором

Сравнение miss rate (MR) на уровне 10^{-4} FPPW
и скорости работы

Дескриптор	MR	t , мс
НОГ [6]	0,08	0,6
НАН-А($q = \cos $, $\text{norm} = L_2$)	0,08	1,4
НАН-Р($q = \cos $, $\text{norm} = L_2$)	0,13	1,5
НАН-АР($q = \cos $, $\text{norm} = L_2$)	0,04	2,2

На рис. 8 представлены результаты работы НАН-АР детектора на тестовых изображениях большого масштаба.

5 Заключение

В работе представлено описание метода построения текстурных признаков для обнаружения пешеходов на изображении. Описан подход построения нового дескриптора в аккумуляторном пространстве Хафа с использованием лучевого поперечного преобразования Радона градиентного поля изображения. Описаны несколько вариантов дескрипторов, построенных в рамках данного подхода. Представлены результаты вычислительного эксперимента с использованием предложенных дескрипторов и их сравнение с НОГ-дескриптором по производительности. Предложенный метод существенно уменьшает ошибки детектирования пешеходов по сравнению с НОГ-дескриптором за счет использования информации о положении контуров объекта. Основным недостатком предлагаемого подхода является довольно низкая производительность, вызванная использованием функции вклада, требующей интерполяции модуля градиента в каждой точке изображения во все бины по углам аккумуляторного пространства Хафа. Частично подобную проблему можно решить использованием других функций вклада, которые позволят уменьшить количество вычислительных операций.

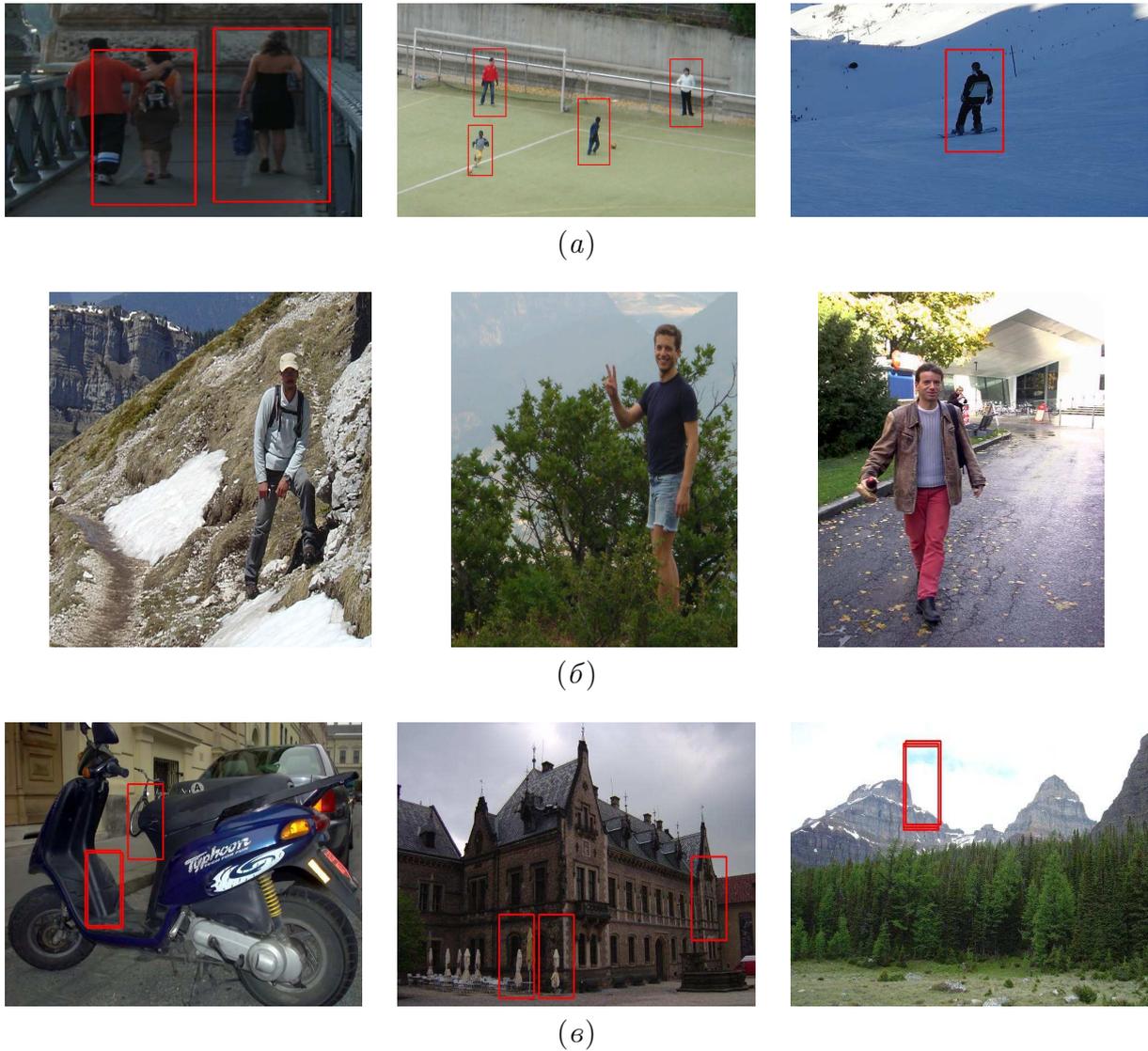


Рис. 8 Результаты положительной работы (а), пропущенные изображения (б) и ложные срабатывания (в) детектора NAN-AR при пороговом значении, соответствующем 10^{-4} FPPW

Литература

- [1] Гнеушев А. Н., Мурьнин А. Б. Адаптивный градиентный метод выделения контурных признаков объектов на изображениях реальных сцен // Изв. РАН. Теория и системы управления, 2003. № 6. С. 128–135.
- [2] Гнеушев А. Н. Построение и оптимизация текстурно-геометрической модели изображения лица в пространстве базисных функций Габора // Изв. РАН. Теория и системы управления, 2007. № 3. С. 85–96.
- [3] Гнеушев А. Н. Оптимизация текстурно-геометрической модели изображения для оценивания параметров лица // Автоматика и телемеханика, 2012. № 1. С. 159–168.
- [4] Viola P., Jones M. J. Rapid object detection using a boosted cascade of simple features // CVPR, 2001. Vol. 1. P. 511–518.
- [5] Viola P., Jones M. J., Snow D. Detecting pedestrians using patterns of motion and appearance // 9th Conference (International) on Computer Vision, 2003. Vol. 1. P. 734–741.

- [6] *Dalal N., Triggs B.* Histograms of oriented gradients for human detection // IEEE CVPR. — San Diego, CA, USA, 2005.
- [7] *Felzenszwalb P. F., Girshick B. R., McAllester D., Ramanan D.* Object detection with discriminatively trained part based models // IEEE T. Pattern Anal., 2010. Vol. 32. No. 9. P. 1627–1645.
- [8] *Sun D., Watanada J.* Detecting pedestrians and vehicles in traffic scene based on boosted HOG features and SVM // IEEE 9th Symposium (International) on Intelligent Signal Processing, 2015.
- [9] *Tabbone S., Wendling L., Salmon J. P.* A new shape descriptor defined on the Radon transform // Comput. Vis. Image Und., 2006. Vol. 102. No. 1. P. 42–51.
- [10] *Nacereddine, N., Tabbone S., Ziou D., Hamami L.* Shape-based image retrieval using a new descriptor based on the Radon and wavelet transforms // 20th Conference (International) on Pattern Recognition. — Istanbul, Turkey, 2010.
- [11] *Светов И. Е.* Формулы обращения для восстановления двумерных гармонических векторных полей по известным лучевым преобразованиям // Сиб. электрон. матем. изв., 2015. № 12. С. 436–446.
- [12] INRIA Person Dataset. <http://pascal.inrialpes.fr/data/human/>.
- [13] Open Source Computer Vision Library. <http://opencv.org/releases.html>.
- [14] *Воронцов К. В.* Лекции по методу опорных векторов. — 2007. <http://www.ccas.ru/voron/download/SVM.pdf>.

Поступила в редакцию 19.09.2017

Textural descriptor in the Hough accumulator space of the gradient field for detecting pedestrians*

N. A. Samsonov¹ and A. N. Gneushev²

nikita.samsonov@phystech.edu; gneushev@ccas.ru;

¹Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow, Russia

²Federal Research Center “Computer Science and Control” of RAS, 44/2 Vavilova Str., Moscow, Russia

The problem of selecting features for recognizing pedestrians on an image is considered. The most popular and effective approach to selecting features is using descriptors based on Histograms of Oriented Gradients (HOG). In this paper, it is proposed to use the Hough accumulator space to generalize the HOG descriptor by obtaining projection not only of orientations, but also of the positions of the boundaries in the local area of the image Hough Accumulator Histograms (HAH). The Hough accumulator space is built on the basis of the beam Radon transform of the gradient field of the image. The proposed methods were tested together with linear support vector machine (SVM) classifier on the INRIA pedestrian database. The results of the experiment have shown the best separating ability of new descriptors and reduction of false detections in comparison with HOG.

Keywords: *HOG, histograms of oriented gradients; HAH, Hough accumulator histograms; descriptor; pedestrians; Radon field transform; Hough transform*

DOI: 10.21469/22233792.3.3.03

*The work was supported by the Russian Foundation for Basic Research, grant No. 16-07-01171.

References

- [1] Gneushev, A. N., and A. B. Murynin. 2003. Adaptive gradient method for extracting contour features of objects in images of real-world scenes. *J. Comput. Sys. Sci. Int.* 42(6):973–980.
- [2] Gneushev, A. N. 2007. Construction and optimization of a texture–geometric model of a face image in the space of basic Gabor functions. *J. Comput. Sys. Sci. Int.* 46(3):418–428.
- [3] Gneushev, A. N. 2012. Optimization of the texture–geometric image model for estimation of the face parameters. *Autom. Rem. Contr.* 73(1):144–152.
- [4] Viola, P., and M. J. Jones. 2001. Rapid object detection using a boosted cascade of simple features. *CVPR* 1:511–518.
- [5] Viola, P., M. J. Jones, and D. Snow. 2003. Detecting pedestrians using patterns of motion and appearance. *9th Conference (International) on Computer Vision*. 1:734–741.
- [6] Dalal, N., and B. Triggs. 2005. Histograms of oriented gradients for human detection. *IEEE CVPR*. San Diego, CA.
- [7] Felzenszwalb, P. F., B. R. Girshick, D. McAllester, and D. Ramanan. 2010. Object detection with discriminatively trained part based models. *IEEE T. Patt. Anal.* 32(9):1627–1645.
- [8] Sun, D., and J. Watanada, 2015. Detecting pedestrians and vehicles in traffic scene based on boosted HOG features and SVM. *IEEE 9th Symposium (International) on Intelligent Signal Processing*.
- [9] Tabbone, S., L. Wendling, and J. P. Salmon. 2006. A new shape descriptor defined on the Radon transform. *Comput. Vis. Image Und.* 102(1):42–51.
- [10] Nacereddine, N., S. Tabbone, D. Ziou, and L. Hamami. 2010. Shape-based image retrieval using a new descriptor based on the Radon and wavelet transforms. *20th Conference (International) on Pattern Recognition*. Istanbul, Turkey.
- [11] Svetov, I. 2015. Formuly obrashcheniya dlya vosstanovleniya dvumernykh garmonicheskikh vektornukh poley po izvestnym luevym preobrazovaniyam [Inversion formulas for recovering the harmonic 2D-vector fields by known ray transform]. *Sib. Electron. Math. Izv.* 12:436–446.
- [12] INRIA Person Dataset. Available at: <http://pascal.inrialpes.fr/data/human/> (accessed June 4, 2017).
- [13] Open Source Computer Vision Library. Available at: <http://opencv.org/releases.html> (accessed May 16, 2017).
- [14] Vorontsov, K. V. 2007. Lectures on support vector machine. <http://www.ccas.ru/voron/download/SVM.pdf> (accessed June 25, 2017).

Received September 19, 2017