

Федотов Н.Г., Сёмов А.А., Курносков А.А. Проблемы распознавания 3D изображений у машин и людей: сравнительная характеристика. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей XIV Междунар. научно-техн. конф. – Пенза: ПДЗ, 2014. – С. 185-193.

УДК 004.93: 519.7: 130.2

ББК 32.81

## ПРОБЛЕМЫ РАСПОЗНАВАНИЯ 3D ИЗОБРАЖЕНИЙ У МАШИН И ЛЮДЕЙ: СРАВНИТЕЛЬНАЯ ХАРАКТЕРИСТИКА

Н.Г. Федотов, А.А. Сёмов, А.А. Курносков

### 3D IMAGE RECOGNITION PROBLEMS IN MACHINES AND HUMANS: COMPARATIVE CHARACTERISTICS

N.G. Fedotov, A.A. Syemov, A.A. Kurnosov

**Аннотация.** Настоящая статья посвящена теме распознавания 3D изображений. Во введении обосновывается актуальность данной темы. Далее описываются основные проблемы распознавания 3D изображений: от построения до анализа и восприятия. Анализ проблем машинного распознавания 3D образов дается в контексте сравнения аналогичных способностей у людей, результаты сводятся в таблицу. На основе излагаемых идей вводится новый метод – геометрическое пространственное трейс-преобразование. Данный подход основан на методах стохастической геометрии и функционального анализа. Кратко дается его описание, обосновывается универсальность.

**Ключевые слова:** распознавание 3D образов, человеческий и машинный анализ данных изображения, пространственное трейс-преобразование, стохастическая геометрия, сканирование сеткой параллельных плоскостей.

**Abstract.** Present article is dedicated by the subject of 3D images' recognition. In the introduction actuality of given theme are substantiated. The following the basic problems of recognizing 3D images are described: from a construction to an analysis and a perception. Problems' analysis of 3D images' machine recognition is given in the context of comparing similar capabilities in humans, results tabulated. Based on the expounded ideas a new method – geometry spatial trace transform – are introduced. This approach is based on the methods of stochastic geometry and functional analysis. Briefly its describe are given, its universality is substantiated.

**Keywords:** 3D image recognition, the human and machine analysis of the image data, the spatial trace transform, stochastic geometry, scanning a grid of parallel planes.

### Введение

Для современного этапа развития теории распознавания образов актуально расширение круга рассматриваемых задач распознавания на пространственные изображения, в то время как ранее внимание исследователей было сосредоточено на решении задач анализа и распознавания 2D изображений.

Зрительные анализаторы человека способны воспринимать больше информации в секунду, нежели слуховые. Поэтому увеличивается количество зрительной информации и средств подачи зрительных данных в различных областях человеческой деятельности.

В современных условиях интенсивного развития компьютерных технологий возрастает роль методов обработки информации, представленной в виде изображений. Данная тенденция находит свое воплощение в таких областях, как аэрокосмические исследования, зрение роботов, трехмерная компьютерная то-

мография, исследования Земли из космоса, автоматизация эксперимента [1]. Одной из главных задач развития машинного зрения является получение более качественных методов обработки информации, получаемой в виде изображения, на основании которых даже в условиях шума и искажений можно получать в системах класса «искусственный интеллект» верные решения.

В настоящее время преимущества анализа и распознавания 3D объектов у человека перед любой машинной системой очевидно. Зрительное восприятие человека начинается не с выделения каких-либо деталей, как у машины, а наоборот, его исходной точкой является целостность образа (гельштат психология). В связи с этим актуальной становится задача сравнительного анализа возможностей зрительных систем человека и машины, чтобы возможности человека можно было реализовать в машинном зрении и искусственном интеллекте.

### **Проблема распознавания изображений у машин и людей**

Исходными данными для зрительного восприятия у людей являются изображения, проецируемые оптической системой глаза на глазное дно, а у машин – изображения, проецируемые оптической системой сканера на входную матрицу (сетчатку) системы.

Процессы восприятия образа, как у людей, так и у машин, протекают в двух тесно связанных фазах:

фаза анализа, выявляющая по определенным критериям характерные признаки: движение, формы, пространственное положение, цвета;

синтезирующая фаза (воображение), создающая на основе выявленных признаков зрительный образ.

Учитывая вышесказанное, можно выделить три основные причины ошибок распознавания образов у людей и машин: неправильное построение изображения, его неправильная обработка и/или интерпретация (таблица). Рассмотрим подробно каждую из них.

Прежде всего проблема адекватного анализа и восприятия образа может быть связана с самой простой и очевидной причиной – в зрительную систему поступают неверные данные из-за неправильно построенного изображения. Это может быть связано с нарушениями в формировании и развитии органов (у людей) или неисправностями в оборудовании распознающей машины.

Для человека данная проблема имеет важное и первостепенное значение, т.к. влияет не только на возможность распознавания образов, но и на само развитие человека, его внутренних и социальных качеств [2]. Так, при нарушении зрения происходит снижение зрительного восприятия у слабовидящих и слепых детей, что проявляется в трудности наблюдения и анализа окружающих предметов. Многие признаки предметов визуального характера (величина, цвет, форма) не воспринимаются такими детьми, что резко обедняет их чувственный опыт, затрудняет ориентировку в пространстве и, как следствие, тормозит развитие наглядно-образного мышления.

*Сравнительная характеристика трудностей решения  
проблемы распознавания образов у человека и машины*

Категория Причина	Человек	Машина
1. Неправильное построение изображения	Вызвано нарушениями работы органов, связанных со зрительным восприятием	Вызвано неправильно настроенным программным и/или аппаратным обеспечением, или неточным оборудованием.
2. Неправильная обработка изображения	Связано с особенностями психики человека, его мотивами и опытом, эмоциональным состоянием	Обусловлено несовершенством алгоритмов обработки изображений, а также с низкими разрешающими способностями и возможностями оборудования (помехи, низкое качество изображения, малое разрешение)
3. Неправильная интерпретация изображения	Связано с феноменом зрительных иллюзий	Связано с несовершенством теоретических методов распознавания образов (особенно в области интеллектуального анализа), а также феноменом зрительных иллюзий

Применительно к машинному зрению данная проблема решается правильной настройкой программного обеспечения, синхронизацией поступающих данных от аппаратных устройств, а также устранением причины поломки того или иного оборудования.

Помимо неправильного построения видимого изображения, трудности в решении проблемы распознавания образов могут быть вызваны неправильной обработкой поступающего изображения, его коррекции.

Особенности обработки изображения у человека зависят от факторов, отражающих психосоматическое состояние человека.

В машинном зрении неточная обработка изображений обусловлена несовершенством используемых алгоритмов. Эффективность машинного зрения при распознавании 3D изображения значительно уступает возможностям человека, что объясняется не слабой мощностью вычислительных средств, а недостаточно развитыми теоретическими методами. Машина неспособна мыслить «свободно» и отойти от заложенного алгоритма.

Кроме того, на неправильную обработку изображений даже при наличии мощного искусственного интеллекта может оказать низкая разрешительная способность сканирующих лазеров и видеокамер роботов. Изображения плохого качества имеют сильные помехи, не прорисовывают тонкие контуры линий, мелкие предметы воспринимаются как окружающий фон.

Третья причина трудностей при распознавании образов – неправильная интерпретация изображения, которая в основном возникает из-за феномена зрительных иллюзий. Зрительные иллюзии, приводящие к ошибке в зрительном восприятии, связаны с неточностью или неадекватностью процессов коррекции

зрительного образа или намеренными нарушениями законов геометрии и оптики в видимом 2D или 3D образе.

Так, например, на рис. 1 изображена комната Эймса – помещение, сконструированное определенным образом, для того чтобы вызывать оптическую иллюзию [3]. Комната воспринимается как обычная прямоугольная, чему способствует даже «шахматный» рисунок пола. Однако на самом деле комната имеет вид прямоугольной трапеции. Дальняя стена расположена под острым углом к одной стене (левый край изображения).



Рис. 1. Иллюзии с нарушением законов перспективы: комната Эймса

Аналогичные иллюзии также присущи машинному зрению и искусственному интеллекту. Так, в последнее время многие компьютерные модели пытались имитировать процесс обработки мозгом визуальной информации. Ученые обнаружили, что искусственные имитаторы мозга могли быть обмануты иллюзией. «Если мы будем строить роботы с искусственным интеллектом, используя модели нашего интеллекта, то это означает, что эти роботы могут также видеть иллюзии, подобно человеку. Имитируя удивительную точность, гибкость и надежность в распознавании объектов, мы также копируем возможные ошибки в расчетах, которые проявляются в зрительных иллюзиях», – писал издательству «Live Science» Астрид Земан, нейробиолог-исследователь из австралийского университета Macquarie University [4].

### Теория пространственного трейс-преобразования

На основе вышеприведенных рассуждений, а также других исследований [5–7] авторами был разработан новый метод – геометрическое пространственное трейс-преобразование. Его суть заключается в следующем.

Пусть  $B(h,r)$  – плоскость с нормалью  $h = [\cos j \psi \sin \omega, \sin j \psi \sin \omega, \cos \omega]$ ,  $r$ ,  $\omega$  и  $\psi$  – сферические координаты плоскости.

Исходный объект  $F$  сканируется сеткой параллельных плоскостей под различными углами  $\omega$  и  $\psi$  с шагом  $\Delta\omega$  и  $\Delta\psi$  до завершения оборота обхода вокруг него в  $2\pi$  радиан по каждому углу с дискретным шагом  $\Delta$  между сканирующими плоскостями (рис. 2). Результат  $G$  пересечения плоскости  $B(h,r)$  и объекта  $F$  вычисляется при помощи гиперфункционала *HyperT*. Множество чисел  $G$  формируется гипертрейс матрица ЗТМ, у которой ось  $0\omega$  направлена горизонтально, ось  $0\psi$  – вертикально, ось  $0r$  – вглубь.

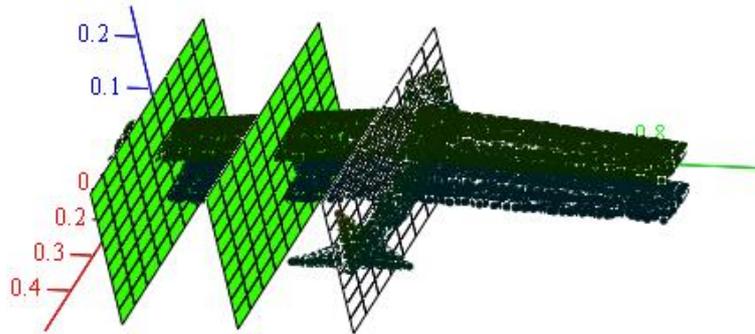


Рис. 2. Процесс сканирования сеткой параллельных плоскостей

Далее при помощи гиперфункционалов  $HyperP$ ,  $Hyper\Omega$  и  $Hyper\Theta$  вычисляется признак  $Res(F)$  исходного 3D объекта [8]:

$$Res(F) = HyperQ \circ HyperW \circ HyperR \circ HyperT(F_{sect}).$$

Признаки получившихся в сечении фигур  $F_{sect}$  извлекаются при помощи 2D трейс-преобразования, введенного в [9]. Признак 2D изображения имеет структуру в виде композиции трех функционалов:

$$P(F_{sect}) = HyperT(F_{sect}) = Q \circ R \circ T(F_{sect} \ I \ l(q, r)),$$

где  $\rho$  и  $\theta$  – полярные координаты прямых в плоскости сечения  $F_{sect}$ .

Данный метод обладает определенной универсальностью, так как схема сканирования не привязана к геометрическим особенностям исходной модели, а благодаря большому числу используемых видов функционалов и их композиционной структуре можно подбирать и конструировать различные признаки [8], которые будут наиболее эффективны при распознавании заданной базы объектов. Предлагаемая методика ориентирована на объекты любой сложности и конфигурации.

Более подробно про новое геометрическое пространственное трейс-преобразование можно найти в работах [10–12]. Преимущество нового метода можно найти в работах [13, 14], а апробацию результатов на 2D и 3D изображениях можно найти в [15–17].

## Заключение

Проблема развития зрительных способностей имеет большое значение для человека, т.к. определяет индивидуальные особенности его развития, играет важную роль в организации его познавательных процессов.

Данная проблема также играет важную роль в такой области, как зрение роботов. Зрительная память ориентирует робота в пространстве, помогает взаимодействию робота с окружающим миром, позволяет идентифицировать объекты. На основе параллелей между характеристиками памяти человека и робота могут быть построены новые идеи коммуникации роботов [18].

Разработанное новое геометрическое преобразование обладает определенной универсальностью, т.к. схема сканирования не привязана к особенностям объекта. Главный инструмент чтения и понимания 3D изображений – пространственное трейс-преобразование, которое осуществляет сканирование исходного пространственного объекта, осуществляется сеткой параллельных плоскостей с разных углов обзора.

Данное пространственное трейс-преобразование осуществляет обработку и анализ 3D изображения в одной технике сканирования благодаря специфичной структуре формируемых признаков [8], что позволяет ускорить время работы распознающей системы. Это свойство присутствует и в способностях и возможностях человека.

*Работа выполнена при поддержке гранта РФФИ, проект № 12-07-00501.*

#### Библиографический список

1. Алпатов А.В., Калинин Р.Е. Способ количественной визуализации формы правого желудочка сердца человека в целях эхокардиографических исследований // Актуальные вопросы клинической морфологии : сб. науч. трудов. – Рязань, 2000. – С. 80-84.
2. Степанова А.Е. Работа с зерном как средство сенсомоторного развития детей с нарушением зрения // Педагогика: традиции и инновации : материалы III Междунар. науч. конф. – Челябинск, 2013. – С. 135-137.
3. Комната Эймса – Интернет-ресурс. – URL: <http://fotoxp.ru/komnata-ejmsa-o-fenomene-zritel'nogo-vospriyatiya/>
4. Интервью нейробиолога Астрида Земана – Интернет-ресурс. – URL: <http://wordscience.org/prostaya-vizualnaya-illyuziya-obmanyvaet-kompyuter.html>.
5. Семов А.А. Об одном подходе к распознаванию 3D-изображений // Надежность и качество : труды Междунар. симпозиума. – Пенза: Изд-во ПГУ. – 2013. – Т. 1. – С. 350-351.
6. Федотов Н.Г. Триплетные признаки распознавания изображений. Применение для анализа трехмерных поверхностей // «XXI век: итоги прошлого и проблемы настоящего – плюс» : научно-методический журнал. Серия: технические науки. Информационные технологии. – Пенза: Изд-во Пенз. гос. технол. акад., 2013. – Вып. 10(14). – С. 61-68.
7. Fedotov N.G., Ryndina S.V., Syemov A.A. Trace transform of spatial images / 11th International conference on Pattern Recognition and Image Analysis: New Information technologies (PRIA-11-2013). Conference Proceedings (V. I-II). – Samara: IPSI RAS, 2013. V. 1. – P. 186-189.
8. Федотов Н.Г., Семов А.А. 3d трейс-преобразование и его свойства // XXI век: итоги прошлого и проблемы настоящего – плюс : научно-методический журнал. Серия: технические науки. Информационные технологии. – Пенза: Изд-во Пенз. гос. технол. акад., 2013. – Вып. 10(14). – С. 68-74.
9. Fedotov N.G. The Theory of Image-Recognition Features Based on Stochastic Geometry // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications. – 1998. – V. 8. – № 2. – P. 264-266.
10. Fedotov N.G., Mokshanina D.A. Recognition of halftone textures from the standpoint of stochastic geometry and functional analysis // Pattern Recognition and Image Analysis, Advances in Mathematical Theory and Applications. – 2010. – Vol. 20. – № 4. – P. 551-556.
11. Fedotov N.G., Mokshanina D.A. Recognition of images with complex halftone texture // Measurement Techniques. – 2011. – Vol. 53. – № 11. – P. 1226-1232.
12. Fedotov N., Romanov S., Goldueva D. Application of triple features theory to the analysis of half-tone images and colored textures. Feature construction along sto-

chastic geometry and functional analysis. Computer and Information Science // Canadian Center of Science and Education Canada. – 2013. – Vol. 6. – № 4. – P. 17–24.

13. Сёмов А.А. Повышение надежности распознавания 3D-объектов на основе методов стохастической геометрии // Надежность и качество : труды Междунар. симпозиума / под ред. Н.К. Юркова. – Пенза: Изд-во ПГУ, 2014. – Т. 1. – С. 393-396.

14. Федотов Н.Г., Семов А.А. Основные преимущества и дополнительные возможности 3D трейс-преобразования // XXI век: итоги прошлого и проблемы настоящего – плюс : научно-методический журнал. Серия: технические науки. Информационные технологии. – Пенза: Изд-во Пенз. гос. технол. ун-та, 2014. – №03(19). – С. 77-83.

15. Федотов Н.Г., Семов А.А. Применение трейс-преобразования для распознавания 3D изображений // Современные методы и средства обработки пространственно-временных сигналов : сборник статей XI Всерос. научно-техн. конф. – Пенза: Приволжский Дом знаний, 2013. – С. 73-75.

16. Сёмов А.А., Федотов Н.Г. Новый метод распознавания и поиска 3D-объектов по базам данных // Открытые инновации – вклад молодежи в развитие региона : сб. материалов регионального молодежного форума: в 2-х т. – Пенза: Изд-во ПГУ, 2013. – Т. 1. – С. 192-194.

17. Семов А.А. Экспериментальная проверка свойств 3D трейс-преобразования // XXI век: итоги прошлого и проблемы настоящего – плюс : научно-методический журнал. Серия: технические науки. Информационные технологии. – Пенза: Изд-во Пенз. гос. техн. ун-та, 2014. – №03(19). – С. 83-89.

18. Шарапов Ю.А., Левченко Е.В., Яковлев В.И. Алгоритм и программная реализация вычисления характеристик кратковременной памяти человека в аспекте моделирования роботов // Современные проблемы науки и образования. – 2012. – № 5. – URL: [www.science-education.ru/105-7006](http://www.science-education.ru/105-7006).

**Федотов Николай Гаврилович**  
Пензенский государственный  
университет, г. Пенза, Россия  
E-mail: [fedotov@pnzgu.ru](mailto:fedotov@pnzgu.ru)

**Fedotov Nikolay Gavrillovich**  
Penza State University, Penza,  
Russia

**Сёмов Алексей Александрович**  
Пензенский государственный  
университет, г. Пенза, Россия  
E-mail: [mathematik\\_aleksey@mail.ru](mailto:mathematik_aleksey@mail.ru)

**Syomov Alexey Aleksandrovich**  
Penza State University, Penza,  
Russia

**Курнос Александр Алексеевич**  
Пензенский государственный  
университет, г. Пенза, Россия  
E-mail: [freedevmas@gmail.com](mailto:freedevmas@gmail.com)

**Kurnosov Alexander Alekseevich**  
Penza State University, Penza,  
Russia