

Московский государственный университет имени М. В. Ломоносова  
Факультет вычислительной математики и кибернетики

*На правах рукописи*

Баринова Ольга Вячеславовна

**Методы машинного обучения для построения трехмерных моделей  
антропогенных сцен**

Специальность 01.01.09 — дискретная математика и математическая кибернетика

АВТОРЕФЕРАТ  
диссертации на соискание учёной степени  
кандидата физико-математических наук

Москва – 2010

Работа выполнена на кафедре математических методов прогнозирования факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова.

Научный руководитель: Доктор физико-математических наук, профессор,  
член-корреспондент РАН  
Рудаков К. В.

Официальные оппоненты: Доктор физико-математических наук, профессор  
кафедры компьютерных методов физики  
физического факультета  
МГУ имени М. В. Ломоносова  
Чуличков А. И.;

Кандидат технических наук, доцент кафедры  
автоматики и телемеханики  
Тульского государственного университета  
Копылов А. В.

Ведущая организация: Институт прикладной математики РАН  
имени М. В. Келдыша.

Защита состоится «17 декабря» 2010 г в 11 часов на заседании диссертационного совета Д 501.001.44 в Московском государственном университете по адресу: 119991, ГСП-1, Москва, Ленинские горы, МГУ, 2-й учебный корпус, факультет ВМК, аудитория 685.

С диссертацией можно ознакомиться в библиотеке факультета ВМК МГУ.  
С текстом автореферата можно ознакомиться на официальном сайте ВМК МГУ имени М. В. Ломоносова <http://www.cmc.msu.ru> в разделе «Наука» - «Работа диссертационных советов» - «Д 501.001.44».

Автореферат разослан 17 ноября 2010 г.

Учёный секретарь  
диссертационного совета  
профессор

Трифонов Н. П.

# Общая характеристика работы

## Объект исследования, актуальность и структура работы

Одной из задач компьютерного зрения является задача извлечения информации о трехмерной структуре сцены из двумерных изображений. С развитием компьютерного зрения появились методы, позволяющие анализировать геометрию трехмерных сцен всего лишь по одному двумерному изображению. Класс изображений антропогенных сцен (городские сцены, снимки, сделанные в помещениях) представляет особый интерес, поскольку фотографии антропогенных сцен составляют большой процент любительских фотографий, и, кроме того, они часто содержат большое число линий, которые являются проекциями параллельных в трехмерной сцене прямых, лежащих на одной плоскости (например, линии границ окон зданий, границы стен зданий, дорожная разметка). Анализ геометрии антропогенных сцен используется в системах поиска объектов и распознавания для повышения точности и надежности их работы<sup>1</sup>. Также анализ геометрии антропогенных сцен по одному двумерному изображению используется для ориентации в пространстве роботов с одной камерой<sup>2</sup>, построении трехмерных компьютерных моделей зданий для создания трехмерных карт<sup>3</sup>.

Данная работа посвящена разработке новых методов анализа геометрии антропогенных сцен и построения их трехмерных компьютерных моделей, превосходящих по точности и надежности существующие методы. Построение трехмерных моделей антропогенных сцен в решающей степени основано на распознавании семейств параллельных прямых. Попытка применения методов машинного обучения для данной задачи привела к необходимости получения новых оценок для ошибок классификации, чему посвящена первая глава работы.

За основу метода машинного обучения, предложенного в первой главе работы, был взят широко используемый алгоритм бустинга<sup>4</sup>, который строит линейные комбинации простых классификаторов из некоторого заданного семейства. Известно, что при наличии шума в данных бустинг склонен к переобучению. В задачах компьютерного зрения процент шумовых объектов как правило бывает достаточно высоким, что негативно сказывается на работе бустинга. В данной работе рассматриваются причины этого явления, и выводятся новые оценки обобщающей способности для линейных комбинаций классификаторов, более точные, чем

---

<sup>1</sup> Hoiem, D., Efros, A.A., Hebert, M.: Putting objects in perspective. // International Journal of Computer Vision, 80, (2008), pp. 2137 – 2144.

<sup>2</sup> Saxena, A., Sun, M., Ng, A.: Learning 3-D Scene Structure from a Single Still Image. // In Proc. of ICCV workshop on 3D representation for Recognition (2007), pp. 1 – 8.

<sup>3</sup> Koutsourakis, P. Simon, L. Teboul, O. Tziritas, G. Paragios, N. Single view reconstruction using shape grammars for urban environments. // In. Proc. IEEE 12th International Conference on Computer Vision (2009), pp. 1795 - 1802

<sup>4</sup> Freund Y., Schapire R., A decision-theoretic generalization of on-line learning and an application to boosting. // Journal of Computer and System Sciences, 904 (1995), pp. 23-37.

существующая оценка<sup>5</sup>. Предлагается новый метод построения линейных комбинаций классификаторов, и доказывается, что предложенный метод минимизирует полученные оценки.

Как было сказано выше, линии, параллельные в реальной сцене, предоставляют ценную информацию для анализа геометрии антропогенных сцен. Однако существующие методы поиска прямых линий, в частности широко используемый метод Хафа<sup>6</sup>, и группировки их в семейства параллельных прямых<sup>7 8</sup>, не обладают достаточной точностью и надежностью. Во второй главе работы предлагается новый метод поиска прямых линий на карте краев изображения, а также метод их группировки в семейства параллельных линий, оба метода основаны на использовании аппарата графических моделей. В отличие от метода Хафа<sup>6</sup>, в предложенном методе поиска прямых линий не требуется решать проблему нахождения локальных максимумов, поэтому метод не требует использования таких эвристик, как подавление не-максимумов.

Согласно правилам перспективной проекции, проекции копланарных параллельных прямых пересекаются в одной точке на плоскости изображения, которая называется точкой схода<sup>9</sup>. Точка схода задает направление прямых линий и соответствующих плоскостей в трехмерном пространстве. Если на изображении присутствуют несколько семейств параллельных линий, лежащих на разных плоскостях в трехмерном пространстве, соответствующие им точки схода лежат на одной и той же прямой. Изображения антропогенных сцен часто содержат несколько семейств горизонтальных линий, лежащих на разных вертикальных плоскостях (например, линии окон лежат на стенах домов). В таком случае прямая, которая содержит соответствующие точки схода, называется горизонтом. Проекция параллельных вертикальных линий пересекаются в точке, которая называется зенитом. Во второй главе работы предлагается новый метод поиска геометрических примитивов различных уровней (прямые линии, точки схода, зенит и горизонт) в рамках одной вероятностной модели. Поиск оценки максимума апостериорной вероятности для предложенной вероятностной модели осуществляется методами дискретной оптимизации. В отличие от существующих методов<sup>7 8</sup>, где геометрические примитивы различных уровней обнаруживаются последовательно, в предлагаемом методе они обнаруживаются одновременно,

---

<sup>5</sup> Schapire, R., Freund, Y., Bartlett, P., and Wee Sun Lee.: Boosting the margin: A new explanation for the effectiveness of voting methods. // *Ann. Statist.* Volume 26, Number 5 (1998), 1651-1686.

<sup>6</sup> Illingworth J., Kittler J. A survey of the hough transform // *Computer Vision, Graphics, and Image Processing* Volume 44, Issue 1, October 1988, pp. 87-116.

<sup>7</sup> Kosecka, J., Zhang, W.: Video Compass. // *In Proc. of European Conference on Computer Vision 7* (2002) pp. 476-491

<sup>8</sup> Tardif, J.P.: Non-iterative approach for fast and accurate vanishing point detection. // *In: International Conference on Computer Vision.* (2009) pp.1250 – 1257.

<sup>9</sup> Criminisi A., Reid I., Zisserman A.: Single view metrology. // *International Journal on Computer Vision*, Volume 40, Number 2, (2000) pp. 123-148.

что позволяет избежать распространения ошибок и добиться более высокой точности и надежности метода.

Как правило, городскую сцену можно представить упрощенной трехмерной моделью, которая состоит из нескольких плоскостей, соответствующих земле и стенам зданий<sup>10</sup>. Существующие методы построения таких трехмерных моделей<sup>10 11</sup> работают недостаточно надежно и не позволяют получать модели приемлемого визуального качества. В третьей главе работы предлагается алгоритм автоматического построения трехмерных моделей по одной фотографии антропогенной сцены. Предложенный метод основан на распознавании образов и использует алгоритм машинного обучения, предложенный в первой главе работы. Поиск оптимальных параметров трехмерной модели сводится к оценке максимума апостериорной вероятности в графической модели условного случайного поля<sup>12</sup> и осуществляется методами дискретной оптимизации. Ориентация вертикальных плоскостей модели определяется на основе анализа прямых линий и точек схода. Предложенный метод позволяет получать трехмерные модели антропогенных сцен более высоко визуального качества по сравнению с существующими методами<sup>10 11</sup>.

## **Цель диссертационной работы**

Целью работы является разработка алгоритмов машинного обучения, обладающих высокой способностью к обобщению; разработка алгоритмов геометрического анализа антропогенных сцен по одному изображению; а также разработка алгоритма автоматического построения трехмерных моделей антропогенных сцен по одному изображению.

Основные задачи работы:

- Разработка алгоритма машинного обучения (machine learning), более устойчивого к шуму в данных, чем существующие методы.
- Разработка методов обнаружения геометрических примитивов на изображениях антропогенных сцен, как то прямые линии, точки схода, зенит и линия горизонта, более точного и надежного, чем существующие методы.
- Разработка алгоритма построения трехмерных моделей антропогенных сцен по одной фотографии, позволяющего получать модели более высокого визуального качества, чем существующие методы.

---

<sup>10</sup> Hoiem, D., Efros, A.A., Hebert, M.: Automatic photo pop-up. // ACM Trans. Graph. 24 (2005) pp. 577–584

<sup>11</sup> Saxena, A., Chung, S., Ng, A.: Depth Reconstruction from a Single Still Image. // International Journal on Computer Vision (2007), Volume 76, Number 1, 53-69

<sup>12</sup> Laferty, J., McCallum, A., Pereira, F.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. // In: Proc. 18th International Conf. on Machine Learning, (2001) pp.282–289

## **Научная новизна работы**

В диссертационной работе предложен новый алгоритм бустинга, основанный на использовании оценок условных вероятностей принадлежности классам для объектов из обучающей выборки. Разработанный алгоритм имеет такую же вычислительную сложность, как и стандартный алгоритм бустинга. В работе получены новые оценки способности к обобщению для линейных комбинаций классификаторов, не зависящие от числа простых классификаторов в линейной комбинации. Доказано, что предложенный метод минимизирует эти оценки.

Также в работе представлена новая новая вероятностная формулировка для поиска нескольких объектов на изображениях, в отличие от преобразования Хафа, не требующая поиска локальных максимумов и применения постобработки. Также предложен метод для решения задачи геометрического парсинга изображения, который одновременно осуществляет поиск геометрических примитивов разных уровней. В отличие от существующих методов поиска геометрических примитивов, работающих по принципу «снизу вверх», предложенный метод осуществляет поиск примитивов разного уровня в рамках одной вероятностной модели, что позволяет избежать распространения ошибок.

Также в работе предложен новый метод построения трехмерных моделей городских сцен. В отличие от существующих методов, в предложенном методе задача построения трехмерной модели формулируется как задача подгонки параметров трехмерной модели. В работе предложена пошаговая схема поиска параметров методом максимума апостериорной вероятности, основанная на методах дискретной оптимизации, что позволяет добиться более высокой скорости работы системы по сравнению с существующими методами.

## **Практическая значимость**

Результаты экспериментов на реальных задачах распознавания образов показали, что предложенные алгоритмы бустинга с вероятностными входами, позволяют повысить точность классификации, по сравнению с существующими методами бустинга.

Предложенные методы поиска прямых линий и геометрического парсинга были протестированы на стандартном наборе изображений, а также базе изображений, собранных автором работы. Экспериментальное сравнение методов, предложенных во второй главе, с существующими аналогами показало преимущество предложенных методов.

На основе разработанных методов бустинга создана система автоматической трехмерной реконструкции городских сцен по одной фотографии. Данная система была разработана по заказу корпорации Samsung в ходе совместного проекта лаборатории Компьютерной Графики и Мультимедиа МГУ и Samsung Advanced Institute of Technology. Метод трехмерной

реконструкции, лежащий в основе системы, запатентован в России [14, 15], США [16] и Южной Корее [17].

### **Личный вклад автора**

Автором разработаны метод обучения по прецедентам с вероятностными входами на основе бустинга, а также получено его теоретическое обоснование. Автором были разработаны и реализованы методы поиска прямых линий на изображениях и геометрического парсинга изображений. Под руководством автора и при непосредственном участии автора разработана система трехмерной реконструкции городских сцен по одной фотографии, в которой применены разработанные методы бустинга с вероятностными входами.

### **Результаты и положения, выносимые на защиту**

На защиту выносятся следующие основные результаты и положения:

1. Новый алгоритм машинного обучения (machine learning). Новые уточненные оценки обобщающей способности для линейных комбинаций классификаторов и доказано, что предложенный алгоритм минимизирует эти оценки.

2. Новый метод геометрического парсинга изображений, позволяющий находить геометрические примитивы разных уровней (пиксели краев, прямые линии, точки схода, зенит и горизонт) в рамках единой вероятностной модели, вывод в которой осуществляется методами дискретной оптимизации.

3. Новый метод автоматического построения трехмерных моделей антропогенных сцен по одной фотографии на основе машинного обучения и дискретной оптимизации.

### **Апробация работы**

Результаты работы докладывались и обсуждались на:

- научном семинаре 6th Annual Watson Workshop “Emerging Leaders in Multimedia and Signal Processing”, в исследовательском центре IBM, США, Хоторн;
- научном семинаре VASC - The Vision and Autonomous Systems Center в университете Карнеги-Меллон, США, Питтсбург, 2010;
- научном семинаре Joint Institutes Workshop INRIA - Microsoft Joint Center, Франция, Орсэ;
- научном семинаре по компьютерной графике и мультимедиа под руководством к. ф. м. н., доц. Ю. М. Баяковского (ф-т ВМК МГУ), Россия, Москва, 2009;

- научном семинаре отдела Интеллектуальных систем ВЦ РАН под руководством чл. корр. РАН, д. ф.-м. н., проф. К. В. Рудакова, Россия, Москва, 2009;
- научном семинаре кафедры АСВК ВМК МГУ под руководством Л. Н. Королева, Россия, Москва, 2009;
- международной конференции «European Conference on Computer Vision», Ираклион, Греция, 2010;
- международной конференции «Computer Vision And Pattern Recognition», Сан Франциско, США, 2010;
- 14-ой Всероссийской Конференции «Математические Методы Распознавания Образов», Суздаль, 2009;
- международной конференции «Machine Learning and Data Mining in Pattern Recognition», Лейпциг, Германия, 2009;
- международной конференции «European Conference on Computer Vision», Марсель, Франция, 2008;
- 18-ой международной конференции по компьютерной графике и машинному зрению «Graphicon-2008», Москва, Россия, 2008;
- 7-ой международной конференции «Интеллектуализация обработки информации», Алушта, Украина, 2008;
- юбилейной 50-ой научной конференции МФТИ, Долгопрудный, Россия, 2007;
- 13-ой всероссийской конференции «Математические Методы Распознавания Образов», Зеленогорский район, Пансионат Гелиос, 2007;
- 17-ой международной конференции по компьютерной графике и машинному зрению «Graphicon-2007», Москва, Россия, 2007;
- международной конференции «European Conference on Machine Learning», Варшава, Польша, 2007;

## **Публикации**

Всего автором диссертации опубликовано 25 научных работ, из них 12 по результатам диссертации, включая 2 статьи в рецензируемых научных журналах из списка ВАК [1,2], 6 статей в сборниках международных научных конференций [3, 3, 6, 7, 9, 13], 1 статью в сборнике всероссийской научной конференции [5] и 3 тезисные публикации в сборниках трудов международных и всероссийских конференций [8, 10, 11]. Автор диссертационной работы является соавтором 4 патентов в России, США и Южной Корее.

## Объем диссертации

Диссертация состоит из введения, трех глав, заключения, списка литературы и приложения. Содержание работы изложено на 155 страницах. Список литературы включает 122 наименования. В работе содержится 33 рисунка и 6 таблиц.

## Содержание работы

**Во введении** формулируются цели и задачи диссертации, показывается актуальность, научная новизна и практическая значимость работы. Описывается структура диссертации.

**В первой главе** рассматриваются задача классификации. Проводится анализ современных точек зрения на проблему переобучения при построении классификатора по обучающей выборке. В частности, рассматриваются существующие теории, объясняющие хорошую способность к обобщению бустинга. Предлагается новый алгоритм бустинга с вероятностными входами. Выводятся теоретические верхние оценки для линейных комбинаций простых классификаторов и доказывается, что предложенный алгоритм минимизирует полученные оценки. Приводится описание и результаты экспериментов, в которых предложенный алгоритм сравнивается со стандартным методом бустинга.

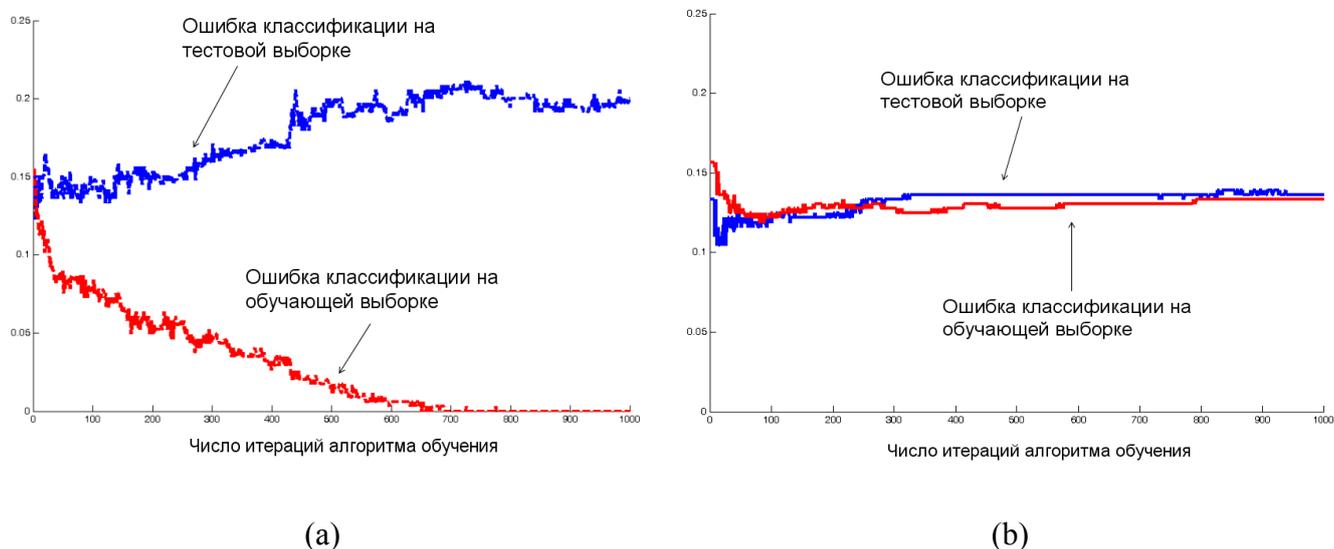
**В первом разделе** главы рассматривается задача классификации и проблема переобучения.

Большую популярность приобрёл подход к обучению по прецедентам, основанный на построении модели восстанавливаемой зависимости в виде параметрического семейства алгоритмов. В рамках данного подхода с помощью численной оптимизации подбираются такие значения параметров модели, при которых алгоритм допускает наименьшее число ошибок на заданной обучающей выборке прецедентов. Другими словами, осуществляется подгонка модели под выборку. Этот метод получил название минимизации эмпирического риска.

При использовании такого подхода на практике исследователи столкнулись с проблемой переобучения. Чем больше у алгоритма свободных параметров, тем меньшего числа ошибок на обучении можно добиться путём оптимизации. Однако по мере нарастания сложности модели «оптимальные» алгоритмы начинают слишком хорошо подстраиваться под конкретные данные, улавливая не только черты восстанавливаемой зависимости, но и ошибки измерения обучающей выборки, и погрешность самой модели. В результате ухудшается качество работы алгоритма вне обучающей выборки, или, иными словами, его способность к обобщению.

Для контроля над переобучением, как правило, используются такие методы как тестовая выборка или скользящий контроль. На Рис. 1(а) по горизонтальной оси отложено число слабых классификаторов в композиции, построенной стандартным методом бустинга. По вертикальной

оси отложена доля ошибочно классифицированных объектов. Красным показана доля ошибочно классифицированных объектов в обучающей выборке, синим показана доля ошибочно классифицированных объектов в тестовой выборке. Видно, что при использовании стандартного бустинга, в то время как ошибка обучения снижается, тестовая ошибка возрастает.



**Рисунок 1 (а)** – зависимость ошибки на тестовой и обучающей выборках в зависимости от числа простых классификаторов в композиции для стандартного метода AdaBoost; **(а)** – зависимость ошибки на тестовой и обучающей выборках в зависимости от числа простых классификаторов в композиции для предложенного метода бустинга с вероятностными входами.

Метод бустинга, предложенный Фрейдом и Шапиром является разновидностью взвешенного голосования классификаторов. Основная идея метода состоит в том, что базовые алгоритмы строятся последовательно, и процесс увеличения различий между ними управляется детерминированным образом. А именно, для каждого базового алгоритма, начиная со второго, веса обучающих объектов пересчитываются так, чтобы он точнее настраивался на тех объектах, на которых чаще всего ошибались все предыдущие простые классификаторы. Веса простых классификаторов в композиции вычисляются исходя из числа допущенных ими ошибок.

На данный момент наилучшее объяснение хорошей обобщающей способности бустинга дает теория отступов (margin theory)<sup>13</sup>, которая предоставляет верхнюю оценку ошибки классификации, не зависящую от числа слабых классификаторов.

**Определение 1.** *Отступом* для объекта  $(x,y)$  относительно классификатора  $f(x)$  называется величина  $yf(x)$ .

Основная теорема из теории отступов гласит, что ошибка классификации любой композиции простых гипотез, представляющей собой линейную комбинацию простых классификаторов из  $H$ , где все коэффициенты положительны и их сумма равна 1, ограничена

<sup>13</sup> Schapire, R., Freund, Y., Bartlett, P., and Wee Sun Lee.: Boosting the margin: A new explanation for the effectiveness of voting methods. In Machine Learning: Proceedings of the Fourteenth International Conference (1997)

сверху суммой значения эмпирической функции распределения отступов на обучающей выборке и дополнительного члена, который зависит от числа объектов в обучающей выборке, числа простых гипотез в семействе, и двух заданных констант.

**Теорема (Шапир и др.).** Пусть  $P$  - распределение над  $X \times Y$ . Пусть  $S$  - выборка из  $m$  объектов, выбранных независимо из распределения  $P$ . Предположим, что задан конечный набор слабых классификаторов  $H$ , принимающих значения из интервала  $[-1, +1]$ . Пусть также задана константа  $\delta > 0$ . Тогда с вероятностью не менее  $1 - \delta$  для выбранной случайным образом выборки  $S$ , все функции  $f \in C$  удовлетворяют следующему неравенству для всех  $\theta > 0$ :

$$P[yf(x) \leq 0] \leq P_S[yf(x) \leq \theta] + \varepsilon(m, |H|, \theta, \delta), \text{ причем}$$

$$\varepsilon(m, |H|, \theta, \delta) = O\left(\frac{1}{\sqrt{m}} \left(\frac{\log m \log |H|}{\theta^2} + \log(1/\delta)\right)^{1/2}\right)$$

Стоит отметить, что дополнительный член в этой оценке не зависит от числа простых гипотез в линейной комбинации и их коэффициентов. В случае фиксированной обучающей выборки данный член представляет собой константу. Таким образом, за счет выбора простых гипотез в композиции и подбора их коэффициентов можно минимизировать первое слагаемое. Следующая теорема показывает, что алгоритм AdaBoost<sup>14</sup> минимизирует его при увеличении числа итераций.

**Теорема (Шапир и др.)** Пусть метод обучения AdaBoost сгенерировал простые гипотезы с ошибками на обучающей выборке  $\varepsilon_1, \dots, \varepsilon_T$ . Тогда для любого  $\theta$ , имеем

$$P_S[yf(x) \leq \theta] \leq 2^T \prod_{t=1}^T \sqrt{\varepsilon_t^{1-\theta} (1-\varepsilon_t)^{1+\theta}}$$

Таким образом, теория отступов дает обоснование хорошей способности к обобщению для бустинга, несмотря на высокую размерность Вапника-Червоненкиса композиции классификаторов, которую строит бустинг. Однако данная теория не дает объяснения переобучению бустинга, которое происходит на практике.

Долгое время бустинг считался устойчивым к переобучению. Однако позднее многими исследователями было продемонстрировано, что методы бустинга крайне чувствительны к шуму в обучающей выборке. Эффект переобучения проявляется в росте тестовой ошибки с увеличением числа простых классификаторов в композиции, построенной бустингом. Данный эффект наблюдается даже при использовании простых классификаторов константной размерности Вапника-Червоненкиса, например пороговых правил. На данный момент не существует четкого объяснения причин такого поведения бустинга.

<sup>14</sup> Freund Y., Schapire R., A decision-theoretic generalization of on-line learning and an application to boosting Journal of Computer and System Sciences, no. 55. 1997

Стоит отметить, что склонность бустинга к переобучения не только ухудшает качество работы систем распознавания, но и затрудняет процесс разработки программного продукта, поскольку для того, чтобы избежать переобучения, требуется проводить дополнительные эксперименты.

В то время как теория отступов дает объяснение хорошей способности к обобщению бустинга, она не раскрывает причин переобучения бустинга, которое происходит на практике. Существует ряд методов, направленных на устранение эффекта переобучения бустинга, основанных на идеях, заимствованных из статистических методов устойчивого оценивания, а также наблюдениях за специфическим поведением алгоритма бустинга при наличии шумовых объектов в обучающей выборке. В ходе проведенного обзора было выявлено, что ни один из существующих на данный момент методов полностью не решает проблемы переобучения для бустинга.

**Во втором разделе** главы выводятся новые оценки обобщающей способности для линейных комбинаций классификаторов, которые применимы в случае, если известны некоторые оценки вероятности принадлежности каждому из классов для всех объектов обучающей выборки. Дополнительный член в предложенных оценках меньше, чем дополнительный член в оценке Шапира.

**Определение 1.** Будем называть задачу бинарной классификации в статистической постановке задачей с перекрывающимися классами, если существует множество ненулевой меры, для которого обе метки классов имеют положительную вероятность:  $P_x[p(+1|x) \times p(-1|x) > 0] > 0$ . Если множество, на котором обе апостериорные вероятности положительны, имеет меру нуль  $P_x[p(+1|x) \times p(-1|x) > 0] = 0$ , будем говорить, что классы не перекрываются.

Для задач с перекрывающимися классами введем обозначения:

$$p_{\max}(x) = \max(P(+1|x), P(-1|x))$$

$$p_{\min}(x) = \min(P(+1|x), P(-1|x))$$

В первых работах, посвященных бустингу, рассматривалась постановка задачи классификации в рамках концептуального обучения. Авторы алгоритма AdaBoost отмечали, что для случая перекрывающихся классов “AdaBoost не является оптимальным методом обучения”. Следует отметить, что для прикладных задач распознавания образов случай неперекрывающихся классов встречается крайне редко, поскольку представление реальных объектов в виде векторов признаков как правило несовершенно. Основная идея данной работы заключается в том, чтобы свести задачу с перекрывающимися классами к задаче концептуального обучения (или, другими словами, задаче с неперекрывающимися классами).

**Определение 2.** Будем называть две задачи бинарной классификации эквивалентными, если множества их решений совпадают.

В работе доказано утверждение о том, что для любой задачи в статистической постановке (в том числе, для задач с перекрывающимися классами) существует эквивалентная ей задача концептуального обучения.

**Теорема 1.** Для любой бинарной задачи классификации с перекрывающимися классами, заданной распределением  $P(X \times Y)$ , существует эквивалентная ей задача классификации с неперекрывающимися классами, заданная распределением  $\tilde{P}(X)$  и целевой функцией  $\tilde{c}(x): X \rightarrow Y$ , где

$$\tilde{P}(x) = (p_{\max}(x) - p_{\min}(x))P(x)/(1-2B),$$

$$\tilde{c}(x) = \arg \max_y (P(y|x)).$$

Причем для любого произвольного классификатора  $f$  верно следующее соотношение  $P[yf(x) \leq 0] \leq B + (1-2B) \times \tilde{P}[\tilde{c}(x)f(x) \leq 0]$ , где  $B$  – ошибка Байесовского классификатора на распределении  $P(X \times Y)$ .

Данная теорема дает способ свести задачу с перекрывающимися классами к задаче с неперекрывающимися классами. Данный результат использован в работе для вывода новых верхних оценок ошибки классификации для композиций простых гипотез.

Предположим, что задан конечный набор слабых классификаторов  $H$ . За множество  $C$  обозначим множество всех линейных комбинаций простых гипотез из  $H$ , где коэффициенты линейной комбинации положительны и их сумма равна 1.

Предположим, что  $S$  – выборка из  $m$  объектов, выбранных независимо из некоторого распределения над  $X \times Y$ . Обозначим за  $P_S[\chi(x, y)] = \frac{1}{m} \sum_{i=1}^m I_{[\chi(x, y)]}$  и  $E_S(\chi(x, y)) = \frac{1}{m} \sum_{i=1}^m \chi(x, y)$ . Символом  $E_y(\chi(x, y))$  будем обозначать условное матожидание величины  $\chi(x, y)$  по  $y$  при фиксированном  $x$ , то есть  $E_y(\chi(x, y)) = p(+1|x) \times \chi(x, +1) + p(-1|x) \times \chi(x, -1)$ .

Основными теоретическими результатами работы являются следующие теоремы.

**Теорема 2.** Пусть  $P$  – распределение над  $X \times Y$ . Предположим, что задан конечный набор слабых классификаторов  $H$ , принимающих значения из интервала  $[-1, +1]$ . Пусть также задана константа  $\delta > 0$ . Пусть  $S$  – выборка из  $m$  объектов, выбранных независимо из распределения  $P(X \times Y)$ , и для каждого объекта известны апостериорные вероятности принадлежности классам. Тогда с вероятностью не менее  $1-\delta$  для выбранной случайным образом выборки  $S$ , все функции  $f \in C$  удовлетворяют следующему неравенству для всех  $\theta > 0$ :

$$P[yf(x) \leq 0] \leq B + (1 - 2\hat{B}) \times E_S \left( \frac{p_{\max}(x) - p_{\min}(x)}{1 - 2\hat{B}} I_{[c(x)f(x) \leq \theta]} \right) + \tilde{\varepsilon}(m, |H|, \theta, \delta),$$

где  $B$  - ошибка Байесовского классификатора на распределении  $P(X \times Y)$ ;

$$\hat{B} = \frac{1}{m} \sum_{i=1}^m p_{\min}(x_i); \quad \tilde{\varepsilon}(m, |H|, \theta, \delta) = \varepsilon(m, |H|, \theta, \delta) - 2B \sqrt{\frac{\ln |H|}{m}} \text{ из Теоремы Шапиро.}$$

**Теорема 3.** Пусть  $P$  - распределение над  $X \times Y$ . Предположим, что задан конечный набор слабых классификаторов  $H$ , принимающих значения из интервала  $[-1, +1]$ . Пусть также задана константа  $\delta > 0$ ,  $\theta > 0$ . Пусть известны оценки апостериорных вероятностей для объектов из обучающей выборки, удовлетворяющие условию:  $|\hat{p}_{\max}(x_i) - p_{\max}(x_i)| < \alpha, i = 1, \dots, m$ . Тогда верна следующая верхняя оценка ошибки классификации по обучающей выборке:

$$P[yf(x) < 0] \leq B + (1 - 2\tilde{B}) \times E_S \left( \frac{\hat{P}_{\max} - \hat{P}_{\min}}{1 - 2\tilde{B}} I_{[\hat{E}_y, yf(x) < \theta + \beta(\alpha)]} \right) + \frac{2\alpha}{1 - 2\tilde{B}} + \tilde{\varepsilon}(m, |H|, \theta - \beta(\alpha), \delta),$$

$$\text{где } \tilde{\varepsilon}(m, |H|, \theta, \delta) = \varepsilon(m, |H|, \theta, \delta) - 2B \sqrt{\frac{\ln |H|}{m}} \text{ берется из Теоремы 6, } \tilde{B} = \frac{1}{m} \sum_{i=1}^m \hat{p}_{\min}(x_i),$$

$\alpha \beta(\alpha)$  непрерывно зависит от  $\alpha$ .

В третьем разделе предлагается новый метод построения линейных композиций классификаторов на основе бустинга, который получает на вход оценки вероятностей принадлежности классам для всех объектов обучающей выборки. Доказывается, что предложенный метод минимизирует оценки обобщающей способности, выведенные во втором разделе главы. Затем приводятся результаты экспериментов на нескольких открытых наборах данных, в которых предложенный метод сравнивается со стандартным методом бустинга.

В работе предложен итеративный алгоритм, на вход которому подается обучающая выборка и значения оценок условных вероятностей принадлежности классам для объектов из обучения. Следуя логике теории отступов, предложенный алгоритм минимизирует верхнюю оценку ошибки классификации из Теоремы 3 (в случае, когда известны точные значения апостериорных вероятностей) или Теоремы 4 (если используются оценки для апостериорных вероятностей).

**Теорема 4** Пусть предложенный алгоритм построил простые классификаторы с взвешенными ожидаемыми ошибками на тренировочной выборке  $\varepsilon_1, \dots, \varepsilon_T$ :  $\varepsilon_i = \sum_{j=1}^m D_j(i) p(y \neq h_i(x_j) | x_j)$ . Тогда для любого  $\theta$  выполняется неравенство:

$$E_S \left( \frac{(\hat{p}_{\max}(x) - \hat{p}_{\min}(x))}{1 - 2\tilde{B}} I_{[E_y, yf(x) \leq \theta]} \right) \leq 2^T \prod_{i=1}^T \sqrt{\varepsilon_i^{1-\theta} (1 - \varepsilon_i^{1+\theta})}$$

В работе используется алгоритм оценивания условных вероятностей принадлежности каждому из классов для объектов из обучающей выборки, на основе многократного применения метода Фридмана<sup>15</sup>. Для этого обучающая выборка несколько раз случайным образом делится на две равные части. На первой части стандартным методом бустинга строится композиция классификаторов, на второй части вычисляются выходы построенной композиции. К выходам композиции применяется преобразование Фридмана. В результате для каждого объекта получается несколько значений вероятностных выходов бустинга. Их среднее арифметическое принимается за оценку условной вероятности принадлежности первому классу. Полученные оценки условной вероятности подаются на вход предложенному методу бустинга с вероятностными входами.

**Таблица 1** Результаты экспериментов на нескольких наборах данных из репозитория UCI

Problem	AdaBoost, 1 итер.	ODDboost, 1 итер.	AdaBoost, 100 итер.	ODDboost, 100 итер.	AdaBoost, 500 итер.	ODDboost, 500 итер.
austra	13.2±0.5	13.2±0.5	14.2±0.2	12.1±0.1	16.0±0.1	13.1±0.1
australian	15.0±0.5	15.0±0.5	14.5±0.2	13.0±0.1	17.4±0.1	13.0±0.1
breast-cancer	8.5±0.3	7.5±0.3	4.5±0.1	4.2±0.0	5.4±0.0	4.4±0.1
checker9	47.0±0.4	46.0±0.4	28.8±0.4	27.5±0.2	29.5±0.4	27.0±0.2
german	39.0±0.3	38.0±0.3	23.7±0.2	26.6±0.2	25.7±0.3	25.0±0.1
madelon	37.5±0.3	37.6±0.3	40.1±0.5	38.7±0.2	41.5±0.3	38.0±0.1
magic04	26.4±0.2	26.4±0.2	17.0±0.5	17.6±0.2	17.9±0.1	16.4±0.1
page_blocks	6.3±0.1	6.4±0.1	3.4±0.2	3.6±0.1	3.3±0.1	2.9±0.1
pima-indians- diabetes	26.3±0.3	25.2±0.3	25.5±0.5	24.2±0.2	26.2±0.1	24.7±0.2
T	31.9±0.3	31.8±0.3	15.0±0.2	14.1±0.1	16.2±0.2	14.6±0.0
vote	4.4±0.2	4.4±0.2	3.7±0.2	4.3±0.2	4.6±0.2	4.3±0.2
XOR	54.0±0.3	57.0±0.3	46.2±0.4	47.4±0.3	46.8±0.4	49.0±0.3
Total :	2	1	6	12	1	14

В таблице 1 приведены результаты экспериментов на данных из репозитория UCI, в которых предложенный метод бустинга с вероятностными входами (ODDboost) сравнивается со стандартным методом AdaBoost. В последней строке поведена статистика, за каждый наилучший результат метод получает 2 балла, за каждое второе место метод получает 1 балл. Видно, что в большинстве случаев использование метода бустинга с вероятностными входами позволяет добиться более высокой точности классификации, чем при использовании стандартного метода AdaBoost.

**Во второй главе** предлагается новый метод поиска прямых, использующий вероятностную модель, схожую с той, что используется в методе Хафа. Также во второй главе предлагается новый метод поиска точек схода, линии горизонта и зенита на изображениях.

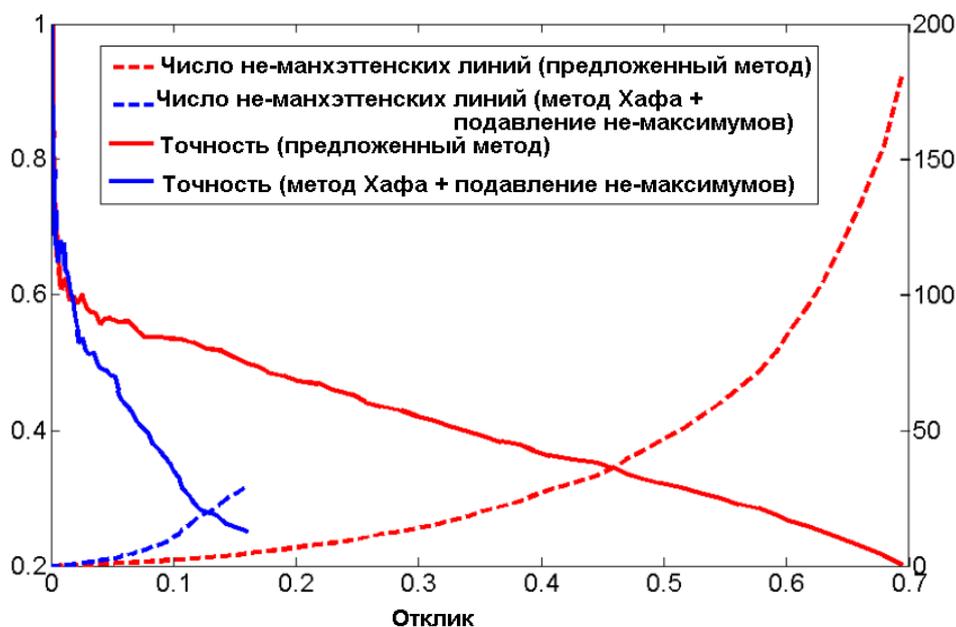
<sup>15</sup> Friedman, J., Hastie, T., & Tibshirani, R.: Additive Logistic Regression: a Statistical View of Boosting. The Annals of Statistics, 28, 2, (2000) 337-407

**В первом разделе** описывается новый метод для поиска прямых на изображениях, который имеет вероятностное обоснование и похож на преобразование Хафа. Однако, в отличие от преобразования Хафа предлагаемый метод осуществляет поиск нескольких объектов посредством вывода в вероятностной модели (минимизации энергии). Вероятностная природа предложенного метода означает, что предложенная модель может быть легко интегрирована с другими вероятностными моделями, например с глобальными геометрическими ограничениями.

Преобразование Хафа основано на том факте, что плотности вероятностей (голоса) можно легко определить или настроить при помощи машинного обучения на основании известного набора данных. Однако, вместо того, чтобы объединить голоса корректным с точки зрения вероятностной модели способом, преобразование Хафа просто суммирует голоса. В предлагаемой схеме голоса объединяются в вероятностной модели, при этом решением задачи является назначение значений переменных по принципу максимума апостериорной вероятности.

Максимизация апостериорной вероятности в данной задаче эквивалентна решению задачи о размещении предприятий. Известно, что эта задача является в общем случае NP-трудной, однако известно, что в терминах фактора аппроксимации оптимальным алгоритмом ее решения является простой жадный алгоритм. Жадный алгоритм начинает свою работу с конфигурации, в которой все гипотезы считаются неактивными, а все голосующие элементы принадлежат фону. Затем, на каждом шаге алгоритм включает одну из гипотез, одновременно переключая часть голосующих элементов на включенную гипотезу (голосующий элемент переключается на гипотезу только в том случае, если это переключение увеличивает значение апостериорной вероятности). На каждом шаге алгоритма гипотеза для включения выбирается таким образом, чтобы обеспечить максимальный прирост апостериорной вероятности.

Эксперименты на стандартном наборе данных, описанные в работе, показывают более высокую точность предлагаемого метода по сравнению с базовым методом преобразования Хафа с последующим поиском локальных максимумов Хаф-карты и подавления не-максимумов в случае, когда на изображении присутствуют несколько объектов интереса. Результаты работы предложенного алгоритма сравнивались с результатами преобразования Хафа с последующим подавлением не-максимумов. Кривые Отклик-Точность для предложенного метода и классического преобразования Хафа с последующим подавлением не-максимумов показаны на рисунке ниже.



**Рисунок 2.** Результаты экспериментов по поиску прямых линий на наборе изображений «Йорк», и. По вертикальной оси отложена характеристика точность, по горизонтальной оси отложена характеристика отклик.

Видно, что предложенный метод работает лучше классического метода по обоим показателям точности и отклика.

**Во втором разделе** рассматривается задача геометрического парсинга для анализа геометрии городских сцен по одной фотографии. Под геометрическим парсингом имеется в виду процесс, когда геометрические элементы на разных уровнях, также как и их группировка внутри одного уровня обнаруживаются явно и одновременно в процессе оптимизации. Отметим, что термин парсинг здесь используется поскольку геометрические примитивы различных уровней обнаруживаются одновременно с их группировкой внутри одного уровня.

Предлагаемая вероятностная модель включает пиксели краев, прямые линии, зенит, горизонтальные точки схода, а также группировку пикселей краев в прямые линии, а прямых линий в семейства параллельных линий. В отличие от существующих методов, большинство которых используют схему решения задачи снизу-вверх, в предложенном методе поиск прямых линий, обнаружение точек схода и высокоуровневый анализ геометрии (поиск зенита и горизонта) интегрированы в единую модель.

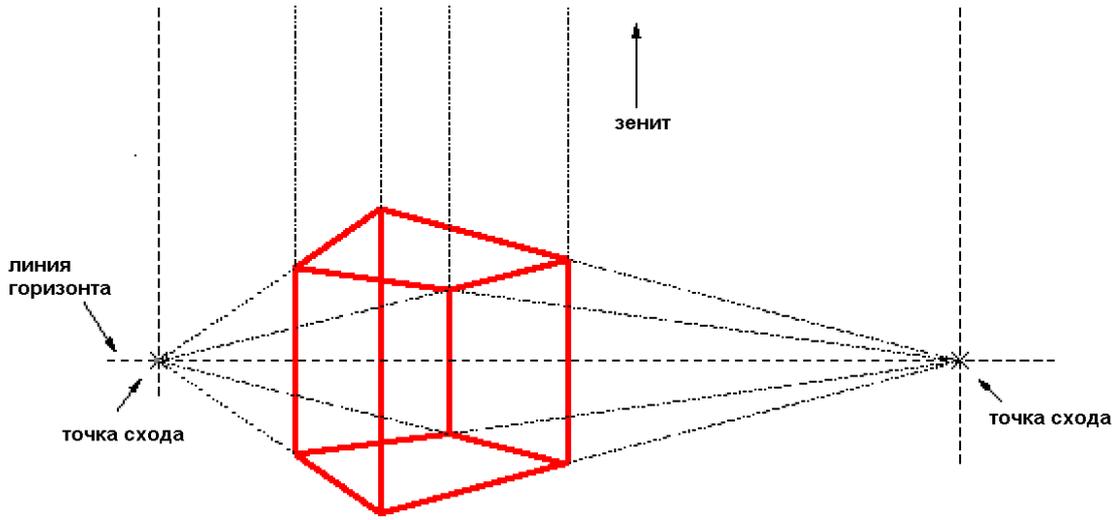


Рисунок 3. Точки схода, зенит и горизонт.

Обозначим пиксели краев за  $\mathbf{p} = \{p_i\}_{i=1}^P$ , где каждый  $p_i$  соответствует позиции пикселя на плоскости изображения. Линии, присутствующие на изображении, обозначаются через  $\mathbf{l} = \{l_i\}_{i=1}^L$ . Обозначим через  $z$  положение зенита (точку схода для семейства вертикальных линий). Через  $\mathbf{h} = \{h_i\}_{i=1}^H$  обозначим точки схода, соответствующие горизонтальным линиям. Таким образом, точки  $h_1, h_2, \dots, h_H$  должны лежать недалеко от одной прямой (это условие будем называть далее ограничением на горизонт).

Общая энергия, таким образом, определяется по формуле:

$$E_{total}(\mathbf{l}, \mathbf{h}, z | \mathbf{p}) = \sum_{i=1}^P E_{edge}(p_i | \mathbf{l}) + \sum_{i=1}^L E_{line}(l_i | \mathbf{h}, z) + \sum_{1 \leq i \leq j \leq H} E_{horizon}(h_i, h_j | z) + E_{prior}(\mathbf{l}, \mathbf{h})$$

где  $E_{prior}(\mathbf{l}, \mathbf{h}) = \lambda_{line} |\mathbf{l}| + \lambda_{vp} |\mathbf{h}|$  - штраф за добавление большого количества линий (согласно принципу минимальной длины описания)  $|\mathbf{l}| = L$  и за добавление большого количества точек схода  $|\mathbf{h}| = H$ , таким образом предпочтение отдается более простым описаниям сцены. Константы  $\lambda_{line}$  и  $\lambda_{vp}$  являются параметрами алгоритма.

Энергия краев задается следующим образом:

$$E_{edge}(p | \mathbf{l}) = \min \left( \theta_{bg}, \min_{i=1, \dots, L} \theta_{dist} \cdot d(p, l_i) \right),$$

где  $d$  обозначает евклидово расстояние на плоскости изображения между пикселем края и прямой линией,  $\theta_{bg}$  - константа, которая соответствует правдоподобию фона, и  $\theta_{dist}$  - константа, которая отвечает за разброс пикселей вокруг одной линии.

Член энергии для прямых линий задается следующим выражением:

$$E_{line}(l | \mathbf{h}, z) = \min \left( \eta_{bg}, \min_{i=1, \dots, H} \eta_{dist} \cdot \phi(l, z)^2 \right)$$

где  $\phi$  обозначает расстояние на Гауссовой сфере между проекцией линии  $l$  и проекцией соответствующей точки схода ( $h_i$  или  $z$ ).  $\eta_{bg}$  - константа, соответствующая правдоподобию линий, которые не являются ни вертикальными, ни горизонтальными,  $\eta_{dist}$  - константа, характеризующая разброс линий в рамках одной группы параллельных линий.

Согласно ограничению на горизонт, описанному выше, все точки схода, кроме одной, лежат рядом с одной прямой на плоскости изображения. Для того, чтобы гарантировать выполнение этого условия, в энергию вводится еще один член:

$$E_{horizon}(u, h | z) = \kappa_{hor} \times tg\phi(u - h, L(z)),$$

где  $\phi$  - модуль угла между вектором  $h_i - h_j$  и перпендикуляром к  $L(z)$ , а  $\kappa_{hor}$  - константа, параметр алгоритма.

Общая энергия, таким образом, объединяет вместе компоненты изображения не-Манхеттенской сцены на разных уровнях, и геометрический парсинг сцены на основе анализа линий происходит посредством минимизации энергии.

В данной работе используется дискретная аппроксимация общей энергии, которую легче минимизировать. Для этого на первом этапе работы алгоритма стандартными методами ищется заведомо большее число линий и точек схода. После выбора кандидатов минимизация полной энергии сводится к минимизации энергии, зависящей от бинарных переменных  $x = \{x_i\}_{i=1, \dots, X}$ ,  $y = \{y_i\}_{i=1, \dots, Y}$  и  $z$ . Значение дискретной функции энергии вычисляется как значение непрерывной функции энергии на соответствующем подмножестве кандидатов в линии и в точки схода.

Фактор-граф, соответствующий получившейся энергии показан на рисунке внизу.

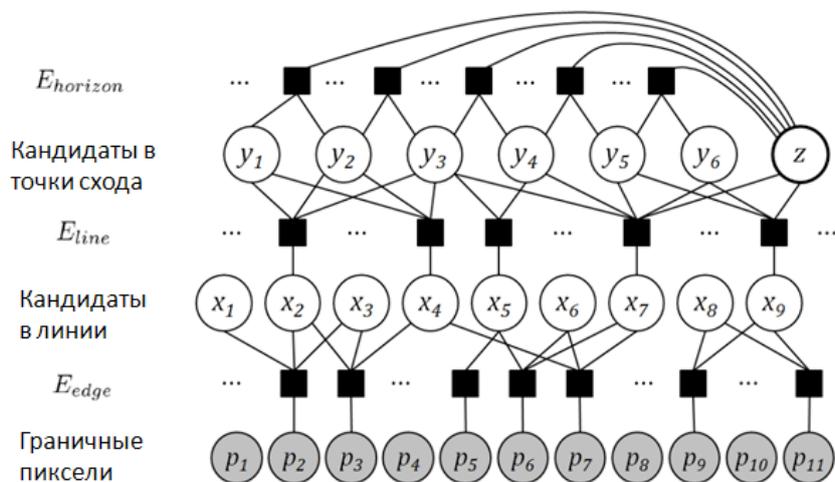
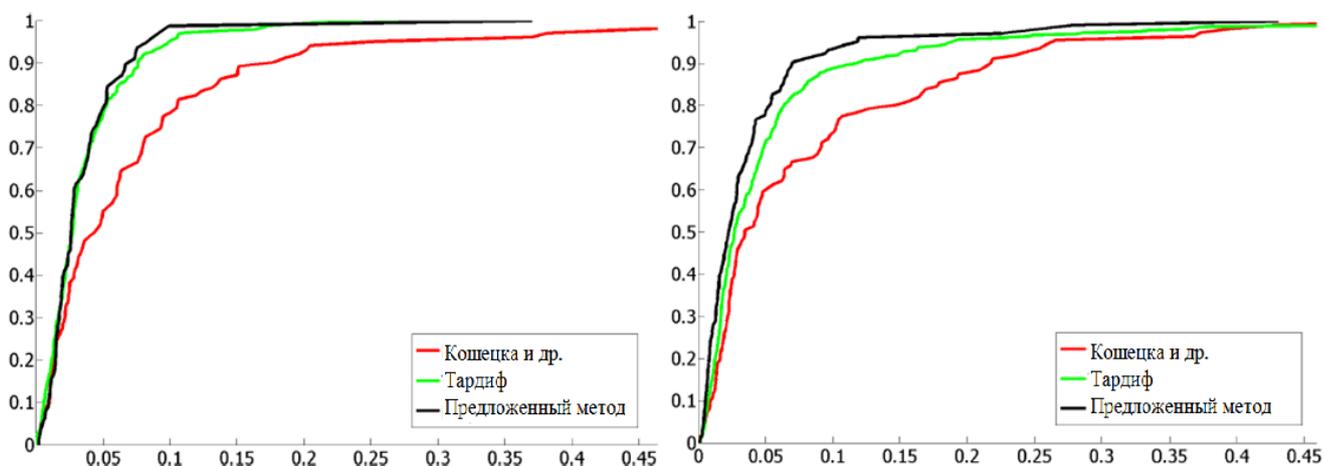


Рисунок 4. Графическая модель, соответствующая дискретной энергии.

Поскольку переменные  $p$  являются наблюдаемыми, можно добиться значительного выигрыша в скорости работы алгоритма за счет суммирования факторов  $E_{edge}$ , которые соединены с одними и теми же (или вложенными) множествами линий. Поскольку члены  $E_{edge}$  представляют собой большинство членов энергии, этот трюк значительно сокращает число членов энергии в модели. Этот факт позволяет использовать простой метод оптимизации, при этом время работы алгоритма составляет всего несколько секунд на одну фотографию. Более подробно, положение зенита определяется перебором всего множества кандидатов (обычно, таких кандидатов не более 10). При фиксированном положении зенита  $z$ , можно провести оптимизацию по бинарным переменным  $x$  и  $y$ , используя метод повторных условных мод со случайным порядком обхода вершин.

В экспериментах, описанных в работе, предложенный метод геометрического парсинга сравнивался с существующими методами поиска точек схода на стандартном наборе изображений «Йорк» и наборе «Евразийские города», подготовленным автором работы. Во время разметки изображений для каждой точки схода было отмечено несколько наиболее отличительных линий. Это позволяет оценить горизонт и зенит с хорошей точностью.

В качестве интегральной характеристики качества работы метода использовать величину отклонения горизонта от истинного положения. Чтобы оценить качество работы методов на всем наборе данных, использовалась эмпирическая функция распределения меры ошибки на всем наборе изображений. Эмпирическая функция распределения изображена на графиках, где по горизонтальной оси отложена величина ошибки  $\tau$ , а по вертикальной оси – доля изображений, для которых ошибка меньше, чем  $\tau$ .



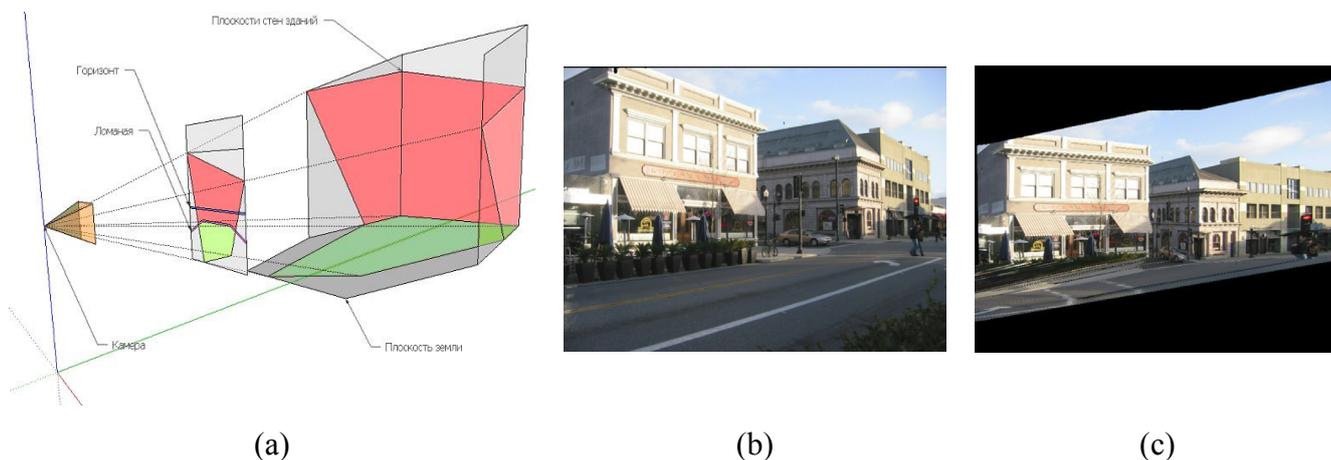
**Рисунок 5.** Результаты сравнения работы предложенного метода с существующими методами Кошечка и др. и Тардиф. По горизонтальной оси отложена ошибка в определении положения горизонта, по вертикальной оси отложена доля изображений в базе, ошибка на которых меньше порога.

Как видно из результатов, предложенный метод дает лучшие результаты, чем существующие методы как на базе изображений Йорк, так и на наборе изображений, собранном автором работы.

**В третьей главе** рассматривается применение предложенных алгоритмов для решения задачи автоматического построения трехмерных моделей антропогенных сцен по одному изображению. Описывается постановка задачи трехмерной реконструкции, проводится анализ методов трехмерной реконструкции по одному изображению. Приводится описание предложенного метода реконструкции, использующего алгоритмы бустинга с вероятностными входами.

Задача трехмерной реконструкции состоит в построении трехмерной модели сцены на основании одной или более фотографии, таким образом, она является обратной задачей. В работе рассматривается задача трехмерной реконструкции по одному изображению, то есть одной перспективной проекции. Такая задача формально имеет бесконечное множество решений, однако для сцен, обладающих регулярной структурой (например, городских сцен), наложение дополнительных ограничений на геометрию сцены и использование семантики сцены позволяет разрешить неоднозначность при восстановлении геометрии сцены.

В диссертационной работе рассматривается задача автоматической трехмерной реконструкции городской сцены по одному изображению. Целью является автоматическое построение визуально качественных трехмерных моделей сцены по одной фотографии, причем существенным требованием к алгоритму является его вычислительная эффективность.



**Рисунок 6** (а) – структура трехмерной модели, используемая для трехмерной реконструкции городских сцен; (б) – входное изображение; (с) – сгенерированный вид.

Во многих существующих работах в области трехмерной реконструкции по одному изображению делались сильные предположения о геометрии сцены. В предложенном в диссертационной работе методе фиксируется структура трехмерной модели, которая задается фиксированным набором параметров, и таким образом задача сводится к нахождению оптимальных параметров трехмерной модели для заданного изображения. Трехмерная модель

сцены состоит из плоскости земли и вертикальных плоскостей, причем граница между стенами и землей представлена ломаной линией (см. Рис. 2). Последнее ограничение позволяет уменьшить число параметров модели, не снижая при этом визуальное качество моделей городских сцен.

В работе предлагается двухэтапный алгоритм поиска параметров модели, то есть задача построения трехмерной модели сцены делится на две подзадачи. Решение каждой из подзадач сводится к нахождению оценки максимума апостериорной вероятности для условного случайного поля, имеющего структуру графа-цепочки. Благодаря простоте структуры графов достигается более высокая вычислительная эффективность предложенного метода трехмерной по сравнению с существующими методами. Ориентация вертикальных плоскостей в работе определяется посредством нахождения точек схода для отрезков горизонтальных прямых. Для настройки условного случайного поля используется метод кусочной настройки, для этого была создана база примеров изображений, для которых параметры модели были заданы вручную. На этом наборе размеченных изображений настраивается несколько классификаторов, вероятностные выходы которых используются при вычислении потенциалов условного случайного поля. Предложенный метод автоматической трехмерной реконструкции городских сцен по одной фотографии реализован в виде программной системы. Метод запатентован в России [14, 15], США [16] и Южной Корее [17].

**В заключении** описываются основные результаты работы.

## **Основные результаты работы**

Результаты, полученные в настоящей диссертационной работе, состоят в следующем:

1. Разработаны алгоритмы бустинга с вероятностными входами, позволяющие в ряде задач получить более высокую точность классификации, чем существующие методы бустинга. Для предложенных методов бустинга впервые получены верхние оценки способности к обобщению, а также доказано, что предложенные методы минимизируют эти оценки, что служит теоретическим обоснованием предложенных алгоритмов.

2. Разработана новая вероятностная модель для поиска многих объектов (в частности, прямых линий). Разработан метод геометрического парсинга антропогенных сцен, позволяющий находить в рамках единой вероятностной модели прямые линии, точки схода, зенит и линию горизонта.

3. Разработан метод автоматической трехмерной реконструкции городских сцен по одной фотографии с применением предложенных методов обучения по прецедентам, позволяющий автоматически строить трехмерные модели городских сцен, обладающие высоким визуальным качеством. Метод реализован в виде программной системы.

## Список публикаций

1. *O. Barinova. Incorporating posterior estimates into AdaBoost. // Pattern Recognition and Image Analysis, МАИК Nauka/Interperiodica, Volume 19, Number 3 / Сентябрь 2009 г., стр. 421-434.*
2. *О. Баринаова. Трехмерная реконструкция городских сцен по одному изображению. // Вестник Компьютерных и информационных технологий, издательство «Машиностроение», Номер 5, 2010 год*
3. *O. Barinova, V. Lempitsky, P. Kohli. On detection of multiple object instances using Hough transform. // In: Computer Vision and Pattern Recognition.; 2010.*
4. *O. Barinova, V. Lempitsky, E. Tretyak, P. Kohli. Geometric Image Parsing in Man-Made Environments. // In: European Conference on Computer Vision.; 2010.*
5. *О. Баринаова, Д. Ветров. Оценки обобщающей способности бустинга с вероятностными входами // труды Конференции ММРО, 2009, стр. 184-188.*
6. *O. Barinova, D. Vetrov. ODDboost: Incorporating posterior estimates into AdaBoost // Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Volume 5632/2009, Machine Learning and Data Mining in Pattern Recognition, pp. 178-190*
7. *О. Баринаова, В. Конушин, А. Соболев, А. Кузьмишкина, А. Якубенко, А. Конушин. Быстрая автоматическая трехмерная реконструкция городских сцен по одному изображению. // в трудах конференции Графикон 2008, том 1., стр. 234-242.*
8. *О. Баринаова, В. Конушин, А. Якубенко, А. Конушин. Быстрый метод семантической сегментации изображений для автоматической трехмерной реконструкции городских сцен по одной фотографии. // в трудах конференции Интеллектуализация обработки информации 2008, стр. 22-24.*
9. *O. Barinova, V. Konushin, A. Yakubenko, K. Lee, H. Lim, A. Konushin. Fast Automatic Single-View 3-d Reconstruction of Urban Scenes. // In: Proc. of European Conference on Computer Vision, 2008 (II). Pp. 100-113.190.*
10. *О. Баринаова, А. Везневец. Повышение обобщающей способности бустинга в задачах с пересекающимися распределениями классов // труды 50-й юбилейной Конференции МФТИ 2007, стр 91-93.*
11. *О. Баринаова, А. Везневец, В. Везневец. Повышение обобщающей способности бустинга в задачах с перекрывающимися классами // труды Конференции ММРО 2007, стр. 82-84.*
12. *O. Barinova, A. Kuzmishkina, A. Vezhnevets and V. Vezhnevets, Learning class specific edges for vanishing point estimation. // In: Proc. of Graphicon'2007, pp. 86-89.*
13. *A. Vezhnevets, O. Barinova.. Avoiding Boosting Overfitting by Removing Confusing Samples. // Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Volume 4701/2007, Machine Learning: ECML 2007, pp. 430-441.*
14. *METHOD FOR DETERMINING GROUND LINE, 2008132273, RU, 17.07.2008, Chae Sunhyuck, Kim Dokyoon, Lee KeeChang, Lim Hwasup, Konushin Anton, Barinova Olga, Konushin Vadim, Yakubenko Anton*
15. *IMAGE PROCESSING METHOD, 2008129149, RU, 17.07.2008, Lim Hwasup, Lee KeeChang, Barinova Olga, Konushin Anton, Konushin Vadim, Yakubenko Anton*
16. *METHOD FOR DETERMINING GROUND LINE, 1911.1137, US, Chae sunhyuck, Kim Dokyoon, Lee KeeChang, Lim Hwasup, Konushin Anton, Barinova Olga, Konushin Vadim, Yakubenko Anton*
17. *METHOD FOR DETERMINING GROUND LINE, 10-2008-0105977, 28.10.2008, Chae Sunhyuck, Kim Dokyoon, Lee KeeChang, Lim Hwasup, Konushin Anton, Barinova Olga, Konushin Vadim, Yakubenko Anton*