

Метод обнаружения дорожных знаков на изображении с обучением на синтетических данных

асп. Чигорин А., асп. Кривовязь Г., к.т.н Велижев А., к.ф.-м.н. Конушин А.

Лаборатория компьютерной графики и мультимедиа, Московский Государственный Университет имени Ломоносова

Введение

Автоматическое обнаружение Рисунок 1 распознавание знаков дорожного движения является актуальной задачей, решение которой может пригодиться при производстве навигационных карт, в системах помощи водителю или в управлении дорожной инфраструктурой. В первом случае знание положения знаков дорожного движения, их ориентации и класса позволит автоматизировать построение дорожного графа, во втором позволит информировать водителя о текущей ситуации на дороге (рекомендуемую скорость, запрещенные повороты, запрещение обгона и так далее). Дорожные знаки сделаны, чтобы быть заметными, и имеют отличительные цвет и форму. Но разнообразие типов знаков и вариаций трансформаций над ними (Рисунок 2), встречаемых при применении в промышленных масштабах, оставляют задачу высокоточной детекции знаков нерешенной до сих пор.

В связи с важностью задачи по данной теме было опубликовано множество работ, но большинство из них работают с выборками ограниченного объема, что не позволяет надежно оценить качество получаемых алгоритмов. Существующие методы обнаружения знаков можно разделить на две категории. В первую попадают те, которые опираются на известные, вручную подобранные, особенности знаков – цвет [1,2] и форму [3,4,5]. Нам неизвестны факты тестирования данных методов на больших объемах данных, поэтому сложно оценивать качество их работы. Ко второй группе относятся методы, основанные на машинном обучении. В основном эти методы основаны на схеме, предложенной в [6]. В [7] авторы предложили многошаговую схему, в которой на первых этапах происходит выделение знаков по форме и цвету с помощью обученных классификаторов, а затем наиболее перспективные кандидаты обрабатываются детектором, аналогичным детектору Viola-Jones [6]. В результате была достигнута точность 97.6% при 2.5 ложных срабатываниях на изображениях размеров в два мегапикселя. В [8] предложены быстрые признаки и укороченная структура классификатора, состоящего из 4-5 этапов, что позволило добиться скорости работы детектора равной 20 мс, что может позволить применять метод в реальном времени. Заявленная точность для знака “въезд запрещен” составила 99.5% при количестве ложных срабатываний равном 10^{-9} на окно детектора.

В данной работе была исследована возможность использования синтетической выборки для обучения детектора дорожных знаков. Также показаны преимущества использования цветов и комбинаций признаков для увеличения точности работы детектора при одновременном уменьшении времени его работы.



Рисунок 1. Пример обнаруженного знака¹



Рисунок 2. Примеры реальных знаков

¹ Данные предоставлены компанией Imajing

Синтез обучающей выборки

Сбор данных для обучения детектора объектов – непростая задача. Например, в России более 200 классов знаков. Некоторые знаки встречаются часто (например, знак “пешеходный переход”), а некоторые очень редко (например, знак “низколетящие самолёты”). Чтобы собрать несколько сотен или даже десятков примеров редко встречающихся знаков нужно просмотреть десятки километров дорог. Поэтому в данной работе мы исследовали вопрос о возможности обучения детектора на основе синтетических данных [8, 14]. В качестве входных данных мы использовали широкодоступные в интернете пиктограммы знаков дорожного движения. Далее эти пиктограммы подвергались разнообразным трансформациям и, в результате, получалась обучающая выборка, размер которой может быть задан заранее. Главная задача, которую мы решали – это найти набор трансформаций, после которого получалось бы такое множество синтетических знаков, подмножеством которого были бы и реальные знаки. В результате мы пришли к следующей схеме генерации синтетических знаков:

1. Сегментация знака на пиктограмме. Получение маски фона
2. Варьирование положения, размера, яркости, насыщенности, размытия, гауссовского шума изображений пиктограмм
3. Вращение знака в трёхмерном пространстве относительно трёх осей и перспективное проецирование получаемых объектов на картинную плоскость
4. Подкладывание фона из реального изображения без знаков, яркость фона приводится к яркости знака

Примеры получаемых синтетических знаков приведены на Рисунок 3. В текущем наборе трансформаций можно отметить отсутствие перекрытия знаков и размытия за счёт движения. Введение данных трансформаций может улучшить текущие результаты ценой значительного увеличения размеров обучающей выборки.



Рисунок 3. Примеры синтетических знаков

Номер этапа	ТР	ФР
1	100%	10^{-1}
2	100%	10^{-2}
3	100%	10^{-3}
4	100%	10^{-4}
5	97.7%	10^{-6}
6	90%	10^{-9}

Таблица 1. Зависимость точности работы детектора от этапа каскада

Обучение детектора

Мы придерживаемся общей схемы построения детектора Viola-Jones на основе каскада классификаторов, обучаемых с помощью бустинга, изложенной в статье [6]. Но, в отличие от этого метода, мы используем разнообразные признаки. Возможность выбора признаков и классификаторов на разных этапах каскада обусловлена как целью повышения точности, так и целью увеличения скорости работы детектора. Повышение точности достигается за счёт комбинирования признаков разной природы (например, текстурных и цветовых), это также повышает и скорость работы, так как требуется добавить в комбинацию бустинга меньше признаков и классификаторов, чтобы добиться того же уровня точности. Также увеличить скорость обнаружения можно за счёт использования “быстрых”, но слабых (с точки зрения точности классификации) признаков на ранних этапах каскада, которые позволят отсеять большинство окон детектора. А на более поздних этапах можно использовать более сильные признаки, на подсчёт которых нужно больше времени.

В нашей реализации на первых четырёх этапах каскада используются быстрые признаки Хаара [6]. Отличие от [6] состоит в том, что они опираются на цвет, так как для знаков цвет является важной характеристикой. Сначала над исходным изображением осуществляются преобразования извлечения цветов, характерных для знаков (красный, синий, желтый). Данные преобразования описаны в [1]. Помимо этого мы добавили два преобразования для извлечения черного и белого цветов:

$$f_w(x) = \frac{(x_r + x_g + x_b)}{3} - \max(x_r, x_g, x_b)$$
$$f_b(x) = 255 - f_w(x)$$

где x – положение пикселя на изображении, x_r, x_g, x_b – значения красного, зелёного и синего каналов в пикселе x .

В результате для каждого из характерных цветов получается карта уверенности в присутствии этого цвета в каждом пикселе. Таким образом, признаки Хаара извлекались по изображению в градациях серого, трём каналам R, G и B, картам уверенности для красного, синего, желтого, белого и черного цветов (в зависимости от знака).

На последующих этапах к признакам Хаара добавляется более сильный признак гистограмм ориентированных градиентов (HOG) [9]. Таким образом, последние этапы нашего каскада похожи на схему, предложенную в [10]. Признак Хаара и HOG дополняют друг друга, так как первый ориентируется на перепады значений яркости между регионами, а второй – на форму. Интерес представляет апробация признаков другой природы, например текстурных признаков LBP [11] и разьединённых диполей [12], способных собирать информацию с разных участков изображения.

Эксперименты

Мы обучали несколько детекторов для разных типов знаков. Типы знаков формируются по признаку цвет-форма. Например, один детектор обучался на круглых синих знаках, другой - на красных треугольных и так далее. Для покрытия большинства классов знаков требуется обучить восемь детекторов. Во всех экспериментах каскад состоял из 5 или 6 этапов. В зависимости от количества пиктограмм количество примеров в обучающей выборке варьировалось от 20000 до 60000. Все обучающие картинки приводились к размеру 30x30 пикселей. Соотношение примеров знаков и примеров фона на каждом этапе составляло 2:1. На первых четырёх этапах использовались только признаки Хаара в комбинации с пороговым классификатором. На выбор подавалось 90000 их вариаций по положению и размеру подокна и цвету. На последних двух этапах к признакам Хаара добавлялось 2000 вариантов признака HOG в комбинации с линейным SVM-классификатором.

Сначала приведём сравнение цветных признаков Хаара с их аналогами, подсчитываемыми по изображениям в градациях серого. Для сравнения мы обучили первый этап каскада на цветных и на серых признаках таким образом, чтобы он верно классифицировал 90% окон с фоном и 99.7% окон со знаками. В случае цветных признаков для этого потребовалось составить комбинацию бустинга из двух признаков, в то время как для серых признаков комбинация состояла из 52 элементов. Это позволило добиться одиннадцатикратного прироста в скорости обнаружения для цветных признаков. При этом выигрышным оказывается использование именно набора разных цветовых признаков, так как на разных шагах алгоритма бустинга выбираются разные цветовые преобразования, которые и попадают в итоговую комбинацию.

Для тестирования предложенного метода нами была выбрана база KUL Belgium Traffic Signs Dataset [13]. Она состоит из 7356 изображений (всего 11219 размеченных на них знаков), на которых отображены 2459 уникальных физических знаков. В ходе наших экспериментов мы считали процент верно обнаруженных физических знаков. Это означает, что знак считается обнаруженным, если он найден хотя бы на одном изображении, содержащем его. В [13] один физический знак в среднем встречается на трёх изображениях.

Для тестирования детектора было выбрано подмножество запрещающих знаков, состоящее из 10 классов. Точность получилась равной 96.5% при 10^{-9} ложных срабатываний на окно детектора, что

соответствует одному ложному срабатыванию на 60 изображений, размером два мегапикселя. В Таблица 1 приведены проценты верных и ложных срабатываний в зависимости от этапа каскада.

Также мы протестировали наш алгоритм на данных, предоставленных компанией Imajing [13]. Данные были сняты во Франции с помощью камеры высокого разрешения с частотой пять кадров в секунду. Мы осуществляли обнаружение знаков ограничения скорости (9 классов знаков). Итоговая точность составила 98% при одном ложном срабатывании на 50 изображений.

Литература

- [1] Andrzej Ruta, Yongmin Li, Xiaohui Liu. Real-time traffic sign recognition from video by class-specific discriminative features. *Journal of pattern recognition*, Volume 43 Issue 1, 2010.
- [2] Andrzej Ruta, Fatih Porikli, Shintaro Watanabe, Yongmin Li. In-vehicle camera traffic sign detection and recognition. *Machine Vision and Applications*, 2009.
- [3] Rachid Belarouss, Jean-Philippe Tarel. Angle Vertex and Bisector Geometric Model for Triangular Road Sign Detection. *Applications of Computer Vision (WACV)*, 2009.
- [4] Rachid Belarouss, Jean-Philippe Tarel. A Real-Time Road Sign Detection Using Bilateral Chinese Transform. *Advances in visual computing. Lecture Notes in Computer Science*, 2009.
- [5] Loy, Barnes. Fast Shape-based Road Sign Detection for a Driver Assistance System. *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings*, 2004.
- [6] Paul Viola and Michael J. Jones. Robust Real-Time Face Detection. *International journal of computer vision*, 2004.
- [7] Radu Timofte, Karel Zimmermann, Luc Van Gool. Multi-view traffic sign detection, recognition, and 3D localization, *Applications of Computer Vision (WACV)*, 2009.
- [8] Gary Overett, Lachlan Tychsen-Smith, Lars Petersson, Niklas Pettersson, Lars Andersson. Creating robust high-throughput traffic sign detectors using centre-surround HOG statistics. *Machine Vision and Applications*, 2011.
- [9] Dalal, Triggs. Histograms of oriented gradients for human detection. *Computer Vision and Pattern Recognition*, 2005.
- [10] Qiang Zhu, Mei-Chen Yeh, Kwang-Ting Cheng. Fast Human Detection Using a Cascade of Histograms of Oriented Gradients. *Computer Vision and Pattern Recognition*, 2006.
- [11] Ojala, Pietikäinen and Harwood. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. *IAPR International Conference on Pattern Recognition*, 1994.
- [12] K. Thoresz and P. Sinha, Qualitative representations for recognition. *J. Vis.*, vol. 1, no. 3, p. 298, Dec. 2001.
- [13] KUL Belgium Traffic Signs Dataset , <http://homes.esat.kuleuven.be/~rtimofte/>
- [14] M. Ozuysal, M. Calonder, V. Lepetit, P. Fua. Fast Keypoint Recognition Using Random Ferns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.