

22. Chang F.-C., Mott H. On the matrix related to the partial fraction expansion of a proper rational function // Proc. of the IEEE. 1974. Vol. 62(8). P. 1162—1163.

Сведения об авторах

- Раис Дахиевич Ахметсафин** — канд. техн. наук, доцент; ООО „Газпромгеоресурс“, Москва; заместитель начальника управления; E-mail: akhmetsafinrd@mail.ru
- Римма Закиевна Ахметсафина** — канд. техн. наук, доцент; Национальный исследовательский университет „Высшая школа экономики“, Москва; E-mail: rakhmetsafina@hse.ru

Рекомендована
НИУ „Высшая школа экономики“

Поступила в редакцию
06.03.13 г.

УДК 004.852, 004.931

П. Н. ДРУЖКОВ

УМЕНЬШЕНИЕ РАЗМЕРНОСТИ ПРИЗНАКОВЫХ ОПИСАНИЙ В ЗАДАЧЕ ДЕТЕКТИРОВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Предлагается модификация алгоритма детектирования объектов на изображениях методом бегущего окна, основанная на выборе подмножества признаков с помощью ансамблей деревьев решений. Приводятся результаты вычислительного эксперимента по сокращению времени детектирования при сохранении качества на примере гистограмм ориентированных градиентов в задаче детектирования пешеходов.

Ключевые слова: детектирование объектов, детектирование пешеходов, гистограммы ориентированных градиентов, выбор признаков, деревья решений.

Введение. Детектирование объектов на изображениях — одна из важнейших задач компьютерного зрения. Алгоритмы, используемые для решения данной задачи, лежат в основе современных интерфейсов взаимодействия с компьютерными системами и применяются, в частности, в робототехнике, следящих системах и т.д.

Перспективные алгоритмы детектирования основаны на извлечении из изображения (или его части) признаков, характеризующих наличие или отсутствие искомого объекта. На этой основе с помощью алгоритма классификации принимается решение о наличии объекта. В работах [1, 2] было показано, что одновременное использование нескольких признаков описаний позволяет улучшить качество детектирования. Однако это приводит к резкому росту размерности решаемых задач, что увеличивает время настройки детектора и его дальнейшей работы. Таким образом, возникает задача понижения размерности, для решения которой используются алгоритмы извлечения (feature extraction) и отбора (feature selection) признаков.

Для автоматического извлечения и отбора признаков используются различные подходы. Среди них отметим методы генерирования новых признаков путем их проецирования на некоторые направления в пространстве признаков, например, найденные с помощью метода главных компонент или частичных наименьших квадратов [3]. Данные методы, успешно используемые при исследовании пространств высокой размерности, не позволяют, однако, сократить время детектирования. Подход, основанный на поиске значимых признаков в многомерных (вплоть до бесконечномерных) пространствах, предложен в работе [4]; алгоритм генерации признаков, описывающих части объектов, рассматривается в работе [5].

Постановка задачи и метод ее решения. Задача детектирования объектов на изображениях заключается в поиске положений всех объектов заданного класса, при этом под положением объекта понимаются координаты обрамляющего его прямоугольника. Входными

данними алгоритма детектирования является изображение I , заданное в виде матрицы интенсивностей пикселей, а выходными данными — совокупность прямоугольников, обрамляющих изображения всех имеющихся объектов исследуемого класса, и соответствующих им числовых показателей надежности детектирования. Таким образом, задача заключается в максимизации количества правильно детектированных объектов при минимизации числа ложных срабатываний. Как правило, баланс между этими двумя показателями выбирается исходя из требований к конкретной системе детектирования.

Одним из наиболее популярных и перспективных подходов к детектированию объектов на изображении является метод бегущего окна (sliding window) [1], основанный на экстенсивном поиске объектов определенного размера $w \times h$, в пикселях. В рамках данного метода задано отображение $\varphi: I_{w \times h} \rightarrow X = F_1 \cdot F_2 \cdot \dots \cdot F_d$ множества всех изображений размером $w \times h$ во множество их признаков описаний X , где F_i — множество допустимых значений i -го признака. Требуется решить задачу обучения с учителем [6], т.е. найти (обучить) классифицирующую функцию $\psi: X \rightarrow \{0,1\}$ из заданного класса K , такую что $L(\psi, T) \rightarrow \min$, где L — функция штрафа, $T = \{(x_i, y_i) : x_i \in X, y_i \in \{0,1\}, i = \overline{1, n}\}$ — выборка векторов признаков для изображений при наличии объекта ($y_i=1$) и его отсутствии ($y_i=0$).

Задачу отбора признаков можно формализовать как нахождение такого подмножества признаков M , при котором целевой признак (класс) y условно не зависит от признаков $\{1, 2, \dots, d\} \setminus M$ при заданном M . На практике точное нахождение такого подмножества сопряжено с существенными трудностями, и задачу можно понимать как нахождение такого подмножества M' признаков, при котором классификатор, построенный только на этом подмножестве, по качеству не уступает классификатору, использующему все признаки.

Алгоритм настройки детектора. В работе [7] предложен алгоритм определения подмножества признаков, использующий ансамбли деревьев решений и итерационно удаляющий лишние признаки из описания X , а также отсеивающий маскированные (зависимые) признаки и устраняющий влияние выбранных признаков на целевой. Таким образом, на каждой итерации алгоритма осуществляется добавление новых признаков в множество выбранных признаков, изначально предполагаемое пустым. На основе результатов экспериментального сравнения, приведенных в работе [7], можно сделать вывод о том, что указанный алгоритм является одним из лучших.

Общая схема предлагаемого в настоящей статье подхода к решению задачи детектирования на основе метода бегущего окна представляет собой последовательность определенных шагов: обучение детектора с использованием всех признаков с помощью итерационной процедуры дополнения выборки примерами изображений фона [1, 8]; запуск алгоритма выбора подмножества признаков по полученной выборке; обучение классификатора по выборке только с отобранными признаками. При этом на первом и третьем этапах могут быть использованы различные алгоритмы классификации.

Вычислительный эксперимент. Рассмотрим результаты экспериментов по обучению детекторов, а также оценке качества и скорости детектирования при использовании подмножеств признаков на примере задачи детектирования пешеходов — одной из ключевых для интеллектуальных систем помощи водителю.

Для обучения детекторов использовалась база TUD-MotionPairs [1], а для тестирования — самая большая на настоящий момент база детектирования CalTech [2]. Тестирование проводилось на каждом тридцатом кадре каждого видеоизображения из наборов set00-set05 в соответствии с методологией, предложенной в работе [2], общее число тестовых изображений 4250. Измерение показателей качества детекторов было реализовано с помощью программных средств базы CalTech.

В качестве признаков описаний использовались гистограммы ориентированных градиентов (HOG — Histograms of Oriented Gradients) [8], являющиеся признанным стандартом в области детектирования пешеходов. Извлечение HOG-признаков осуществлялось с помощью открытой эффективной программной реализации [9]. В качестве классификатора использовались алгоритмы машинного обучения SVM с линейным ядром (Support Vector Machine — метод опорных векторов) [6] и GBT (Gradient Boosting Trees — градиентный бустинг деревьев решений), реализованные в библиотеке OpenCV [10, 11]. Во всех описанных экспериментах использовались детекторы с шагом окна детектирования по вертикали и горизонтали 8 пикселей, и мультипликативным шагом по масштабу 1,1, при этом изначально изображение увеличивалось в 2,5 раза для поиска небольших объектов.

В ходе эксперимента была обучена модель GBT с помощью итерационной процедуры. Данный подход позволяет в результате получить выборку меньшего объема, чем при использовании SVM. На полученной выборке был запущен алгоритм отбора признаков. Затем на выбранном подмножестве признаков и все той же выборке была обучена модель SVM с линейным ядром. Как отмечалось выше, на каждой итерации алгоритма выбора признаков формируется некоторое подмножество. График, представленный на рис. 1, демонстрирует, что рост мощности подмножества (N) постепенно замедляется, и при этом даже после 18 итераций величина N существенно меньше общего количества признаков, равного 3024 (здесь n — количество итераций).

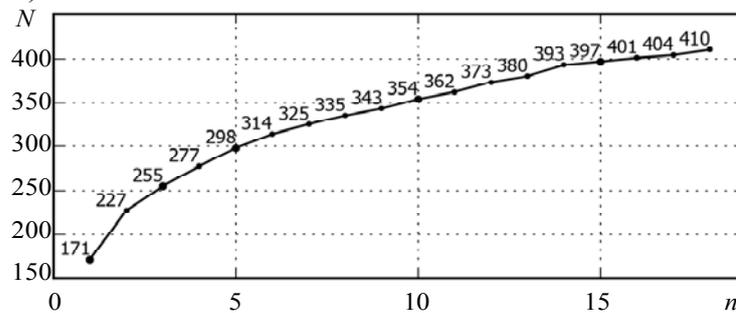


Рис. 1

На подмножествах, полученных после итераций 1, 3, 5, 10 и 15, были обучены линейные модели SVM. Показатели качества детекторов ($Q(e)$, где Q — доля необнаруженных пешеходов, e — количество ложных срабатываний на одно изображение), использующих данные подмножества, а также детекторов, полученных при обучении моделей SVM и GBT без отбора признаков, и HOG-детектора из библиотеки OpenCV, приведены на рис. 2.

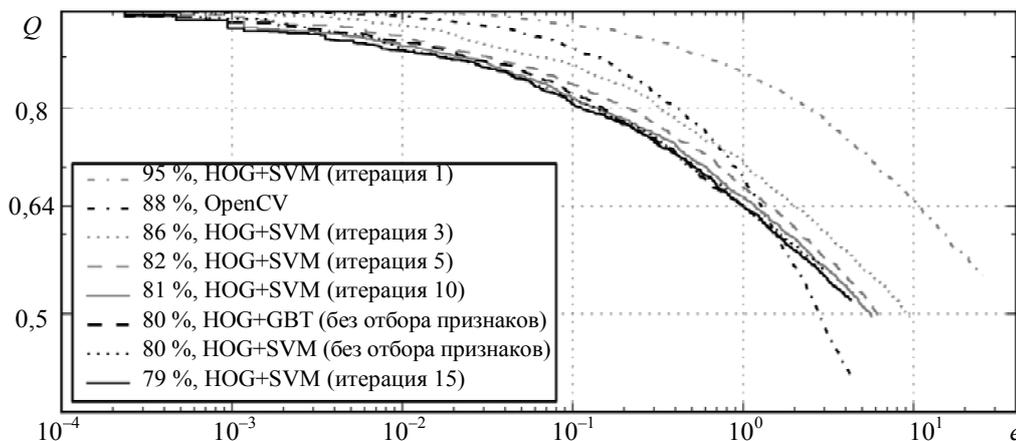


Рис. 2

Кривые получены при варьировании порога по значению надежности детектирования. Анализ графиков показывает, что большинство компонентов HOG-описания являются избыточными и не влияют на качество решения задачи.

HOG-признаки вычисляются блоками по 36 компонентов, в связи с чем при вычислении лишь некоторого подмножества признаков возникают существенные трудности, так как даже блоки для определенного окна детектирования, не содержащие значимых признаков, необходимы при использовании других окон. Однако, во-первых, сокращение размерности может быть использовано для ускорения классификации, а не для вычисления признаков, и, во-вторых, декомпозиция множества вариантов (в частности, принимаемых пешеходом положений/поз) на отдельные случаи может позволить еще больше снизить размерность признаков для каждого из этих случаев.

В ходе экспериментов в качестве конечного классификатора использовалась модель SVM с линейным ядром, позволяющая построить гиперплоскость, разделяющую признаки объектов (пешеходов) и фона. Таким образом, с помощью этого алгоритма требуется определить, с какой стороны от данной гиперплоскости лежит точка в пространстве признаков, что сводится к вычислению скалярного произведения координаты точки и нормали к гиперплоскости. При детектировании объектов методом бегущего окна осуществляется классификация порядка сотни тысяч окон. За счет уменьшения времени вычисления скалярных произведений можно существенно сократить время работы детектора. Так, был проведен эксперимент с использованием 354 признаков, отобранных после 10 итераций работы алгоритма. Описанная модификация позволила уменьшить среднее (по 360 изображениям из набора CalTech set01) время детектирования с 1,54 до 1,15 с (т.е. почти на 25 %). Для экспериментов использовался компилятор Intel C++ Compiler 13.1.0, измерения времени работы детектора проводились с использованием одного ядра процессора Intel Xeon X5570, 12ГБ RAM.

Заключение. Рассмотренный алгоритм отбора признаков для построения детектора объектов на изображениях является удобным инструментом анализа признаковых описаний и позволяет ускорить работу детектора при сохранении качества, что подтверждается результатами экспериментов по выделению HOG-признаков в задаче детектирования пешеходов с использованием алгоритма SVM с линейным ядром.

Статья подготовлена по результатам работы, выполненной в рамках программы „Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007—2013 гг.“, гос. контракт № 11.519.11.4015.

СПИСОК ЛИТЕРАТУРЫ

1. *Wojek C., Walk S., Schiele B.* Multi-cue onboard pedestrian detection // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'09). 2009. P. 794—801.
2. *Dollár P., Wojek C., Schiele B., Perona P.* Pedestrian detection: an evaluation of the state of the art // Pattern Analysis and Machine Intelligence. 2011. Vol. 34, N 4. P. 743—761.
3. *Schwartz W., Kembhavi A., Harwood D., Davis L.* Human detection using partial least squares analysis // Proc. of the IEEE Intern. Conf. on Computer Vision (ICCV). 2009. P. 24—31.
4. *Dollár P., Tu Z., Tao H., Belongie S.* Feature mining for image classification // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'07). 2007. P. 1—8.
5. *Bar-Hillel A., Levi D., Krupka E., Goldberg C.* Part-based feature synthesis for human detection // Proc. of European Conf. on Computer Vision (ECCV'10). 2010. P. 127—142.
6. *Hastie T., Tibshirani R., Friedman J.* The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009.
7. *Tuv E., Borisov A., Runger G.* Feature selection with ensembles, artificial variables, and redundancy elimination // J. of Machine Learning Research. 2009. N 10. P. 1341—1366.
8. *Dalal N., Triggs B.* Histograms of oriented gradients for human detection // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'05). 2005.

9. Piotr's Image & Video MatLab Toolbox [Электронный ресурс]: <<http://vision.ucsd.edu/~pdollar/toolbox/doc/index.html>, free>.
10. OpenCV Library [Электронный ресурс]: <<http://opencv.org>, free>.
11. *Druzhkov P. N., Eruhimov V. L., Kozinov E. A., Kustikova V. D., Meyerov I. B., Polovinkin A. N., Zolotykh N. Yu.* On some new object detection features in OpenCV Library // Pattern Recognition and Image Analysis. 2011. Vol. 21, N 2. P. 377—379.

Сведения об авторе

Павел Николаевич Дружков — аспирант; Нижегородский государственный университет им. Н. И. Лобачевского, кафедра математической логики и высшей алгебры;
E-mail: druzhkov.paul@gmail.com

Рекомендована кафедрой
математического обеспечения ЭВМ

Поступила в редакцию
07.10.13 г.