

基于改进颜色自相似特征的行人检测方法

顾会建*, 陈俊周

(西南交通大学 信息科学与技术学院, 成都 610031)

(* 通信作者电子邮箱 ugigguy@163.com)

摘要:近年来多尺度行人检测在计算机视觉领域受到广泛关注。传统方法需对图像缩放,在不同尺度计算特征,大大降低了行人检测的速度。颜色自相似特征(CSSF)被提出以克服此不足。针对颜色自相似特征具有维度高和分类器训练时间长等问题,提出一种改进的颜色自相似度特征。改进的颜色自相似度特征结合行人结构相似性,首先定义了固定尺寸的窗口,然后在不同的颜色空间滑动固定大小的窗口进行特征提取,最后结合自适应增强(AdaBoost)算法构建行人检测分类器。实验结果显示:相对于传统颜色自相似度特征的千万级维度,新的特征只有几千维,特征提取速度和分类器训练速度显著提高,检测效果略有下降;与梯度方向直方图特征(HOG)相比,特征提取速度提高5倍,检测效果基本不变,新的方法在实时行人检测和监控系统中有好的应用价值。

关键词:结构相似;颜色自相似度特征;多尺度行人检测;尺度不变性;AdaBoost

中图分类号: TP391.4 **文献标志码:** A

Pedestrian detection based on improved color self-similarity feature

GU Huijian*, CHEN Junzhou

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu Sichuan 610031, China)

Abstract: In recent years, multiscale pedestrian detection received extensive attentions in the field of computer vision. In traditional methods, the input image must be resized with different scales to compute the features, which significantly reduces the detection speed. Color Self-Similarity Feature (CSSF) was presented to overcome this problem. An improved CSSF with lower dimension was proposed for the CSSF whose dimension is too high and time-consuming in the training process of the classifiers. Combined with pedestrian structural similarity, a fixed-size window was defined at first, and then the improved CSSF was extracted by sliding the fixed-size window in different color space. Finally, the pedestrian detection classifier was constructed by combining with AdaBoost algorithm. Test shows that compared with the traditional CSSF whose dimension is ten millions, new feature dimension is only a few thousand, and it can be extracted and trained faster, but detection effect decreases slightly; compared with the Histogram of Oriented Gradient (HOG), feature extraction speed improves 5 times, detection effect is essentially the same. The new method has a good application value in real-time pedestrian detection and monitoring systems.

Key words: structural similarity; Color Self-Similarity Feature (CSSF); multiscale pedestrian detection; scale invariance; Adaptive Boosting (AdaBoost)

0 引言

图像中的行人检测是计算机视觉领域的研究热点和难点,在智能交通系统、辅助驾驶以及视频监控等许多方面有着广泛的应用^[1-3]。由于在真实场景中行人的形状、着装多变,且外观易受环境光照、遮挡等诸多噪声影响,如何快速、稳定地获取行人特征,对行人检测的效果极为关键。

国内外很多学者对现有行人检测方法进行了总结^[4-6],目前最常见的行人检测方法为滑动窗口法,该方法选取一个固定大小窗口,通过移动窗口并提取其中特征来判断当前窗口中是否存在行人。Viola等^[7]于2004年提出了一种快速的行人检测方法,该方法采用具有尺度不变性的哈尔(Haar)特征,结合自适应增强(Adaptive Boosting, AdaBoost)算法构建级联分类器对非行人窗口进行快速排除,使得行人检测速度显

著提高,但该算法的准确度并不理想;Dalal^[8]于2005年提出梯度方向直方图特征(Histogram of Oriented Gradient, HOG)与支持向量机(Support Vector Machine, SVM)相结合的行人检测算法,取得了较高的检测准确率。但HOG特征本身不具有尺度不变性,要实现多尺度行人检测,需要对图片进行缩放,每次缩放后都必须重新计算HOG特征,这使得运算量显著增大,难以实时。随后,有部分学者在Dalal工作的基础上对HOG特征进行了改进^[9-11],但是HOG特征缺乏尺度不变性导致其在多尺度行人检测上速度难以显著提升。近年来一些学者提出了其他的行人检测特征,主要有小边特征^[12](Edgelet)、积分通道特征^[13](Integral Channel Feature)、CENTRIST^[14](CENSus Transform hISTogram)等。Edgelet特征是一种描绘行人局部轮廓方向的特征;Integral Channel Feature融合了多种特征,具有参数少且对参数不敏感的优

收稿日期:2014-01-16;修回日期:2014-03-09。 基金项目:国家自然科学基金资助项目(61003143)。

作者简介:顾会建(1989-),男,河南濮阳人,硕士研究生,主要研究方向:计算机视觉、目标检测; 陈俊周(1979-),男,四川成都人,副教授,博士,CCF会员,主要研究方向:计算机视觉、模式识别、机器学习。

点;CENTRIST 是一种刻画图像全局轮廓的特征,具有计算简单快速的优点。这些方法主要依赖行人所在图像区域的纹理和形态信息,缺少对其局部颜色特性的深入分析。

行人衣着多变导致常见的颜色特征难以对其稳定描述,然而行人却有着很强的局部相似结构,例如:人体躯干衣服颜色以及两手臂颜色的相似度很高。基于这一发现,Walk等^[15]提出一种颜色相似度(Color Self-Similarity, CSS)特征,并用于行人检测,新的特征有效提高了检测精度,但检测速度较慢。文献[16]对 CSS 特征进行了改良,提出颜色自相似度特征(Color Self-Similarity Feature, CSSF),但该特征维度达到了千万级,采用 AdaBoost 算法进行训练将耗费大量时间,难以满足实际应用需要。针对这一问题,本文提出了改进的颜色自相似度特征(Enhancement Color Self-Similarity Feature, ECSSF),ECSSF 在固定大小的矩形块中进行构建,不同步进长度下构建的 ECSSF 维度不同,相对于 CSSF 的千万级维度,步进长度为 8 的 ECSSF 只有 4 032 维,这大大缩小了特征提取和分类器训练时间,同时保持了较高的行人检测准确率。

1 局部结构相似性

1.1 颜色自相似度特征

行人具有很强的局部相似性结构,一般情况下,人体上衣的颜色相似度、两手臂的颜色相似度、两个裤腿的颜色相似度均较高,而人体和周围环境的颜色相似度较低。利用该特性可以有效地对行人和非行人目标进行区分。图 1 为行人局部结构相似性示意图,将原始图片转换到 HSV 颜色空间,在 S 通道中,矩形框 1 与矩形框 2 相似度很高,而与矩形框 3 相似度很低,在 H 通道和 V 通道也存在类似的特性。



图1 行人结构相似示意图

基于此观察,文献[16]提出了一套颜色自相似度特征(CSSF),首先对样本图片进行颜色空间转换,例如转换到 HSV 颜色空间,然后在不同的通道上构建特征。文献[16]定义了一个矩形块 R_1 , 提取某个通道内该矩阵块内数值之和 Sum_{R_1} :

$$Sum_{R_1} = \sum_{(x,y) \in R_1} S(x,y) \quad (1)$$

CSSF 的特征为 R_1 与相邻的矩形块 R_2 的比值:

$$F(R_1, R_2) = Sum_{R_1} / Sum_{R_2} \quad (2)$$

矩形块 R_1 和 R_2 大小相等但位置不同。该方法可以衡量每个矩形块之间的相似度。为了增加特征多样性,文献[16]对当前提取的特征进行拓展,获得 TF 特征:

$$TF(R_1, R_2) = \begin{cases} F(R_1, R_2), & F(R_1, R_2) \geq 1 \\ 1/F(R_1, R_2), & F(R_1, R_2) < 1 \end{cases} \quad (3)$$

文献[16]将矩形块 R_1 和其周围邻近的 24 个相同大小的矩形块作比较来提取特征,可获得 48 维的特征, R_1 的大小和位置是多变的,通过遍历不同图像通道、矩形块位置及大小,获取最终特征,实验显示该特征在行人检测上取得较好效果。

1.2 改进的颜色自相似度特征

虽然通过积分图^[17] CSSF 可以快速地提取,但遍历所有图像通道、矩形块位置及大小,提取的 CSSF 维度将达到千万级别,这不仅耗费时间,也使得分类器训练极为困难。本文针对该情况对 CSSF 进行改进,提出了 ECSSF,降低其特征维度,缩短特征提取时间和分类器训练所需时间,并取得了较好的行人检测效果。

本文受 HOG 以及局部二值模型(Local Binary Pattern, LBP)的启发,以 24×24 像素的 Block 以固定步长 w 遍历 64×128 的检测窗口,将每个 Block 划分为 3×3 个 Cell,每个 Cell 的大小为 8×8 ,如图 2 所示。每个 Block 采用与文献[16]相同的方法提取特征,将中心的 Cell 和周围的 Cell 做除法来提取特征,计算公式如下所示:

$$ECSSF[i] = Sum_0 / Sum_i \quad (4)$$

其中: Sum_0 代表某通道中编号为 0 的 Cell 中数值之和, Sum_i 代表某通道中编号为 i 的 Cell 中数值之和。按照文献[16]方法对式(4)提取的特征进行非线性转换,则每个 Block 可以提取 16 维特征,假设固定步长 w 为 4,则 Block 在宽度和高度方向分别步进 10 和 26 次,在 3 个不同的通道构建特征,最终提取的特征维度为 $16 \times (26 + 1) \times (10 + 1) \times 3 = 14256$ 。相对于 CSSF 的千万级维度,ECSSF 的维度显著降低。通过积分图可快速提取 ECSSF,它保留了 CSSF 的尺度不变性,可以通过缩放检测窗口大小的方法实现行人多尺度检测,这样避免了对图像进行金字塔式缩放、重新计算积分图,提高了检测速度。

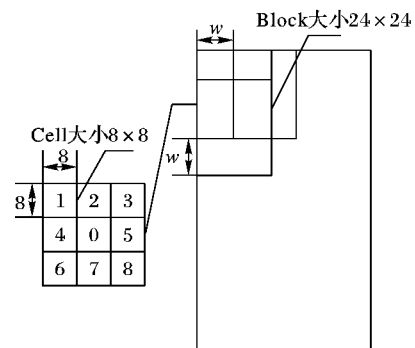


图2 ECSSF 提取矩形块

2 实验

2.1 数据库

INRIA 数据库是行人检测最常用的数据库,它包含了各种光照、环境等情况,训练集有正样本 614 张(包含 2416 个行人),负样本 1218 张;测试集有正样本 288 张(包含 1126 个行人),负样本 453 张。MIT 数据库包含了 924 个行人图片。本文的正样本库由 MIT 的 924 个行人图片和 INRIA 的 2416 个行人图片构成;负样本库是从 INRIA 数据库的 1218 张负样本图片中随机截取的,共 12180 张负样本图片。测试正样本包含 1126 张行人图片,测试负样本包含 4113 张图片。图 3 为部分图片样本实例。

2.2 实验结果

本文采用 AdaBoost 算法训练分类器,图 4 给出迭代次数从 1 到 500 的分类效果,可以看出在迭代次数在 300 左右时分类效果趋于稳定,因此本文选择的分类器迭代次数为 300。

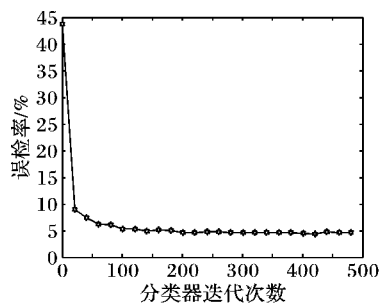
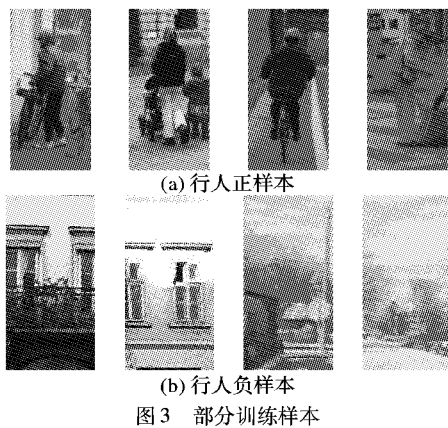


图5给出了ECSSF在INRIA数据库中的部分检测效果。

表1对比了在不同颜色空间、选取不同Block步进长度提取ECSSF与HOG特征以及CSSF的实验效果,带下划线的数字

表1 ECSSF与CSSF、HOG性能比较

特征类型	w	特征维度	准确率/%	召回率/%	特征提取时间/ms	分类器训练时间/h
ECSSF-LUV	4	14 256	<u>97.9</u>	95.4	12.1	32.3
	8	4 032	<u>97.5</u>	95.4	<u>3.7</u>	<u>6.7</u>
ECSSF-RGB	4	14 256	96.1	93.8	12.1	42.3
	8	4 032	95.0	93.4	3.7	6.7
ECSSF-HSV	4	14 256	97.5	95.1	12.1	42.3
	8	4 032	95.9	94.5	3.7	6.7
CSSF		6 912	96.3	94.7	5.6	15.2
HOG		<u>3 780</u>	97.6	<u>95.6</u>	23.4	<u>6.1</u>



3 结语

本文通过改进CSSF,提出了ECSSF,解决了传统CSSF存在维度过高、分类器训练耗时等缺点,同时保存了CSSF的尺度不变性。实验显示,改良的特征可有效描述行人的局部结构相似特性,通过与AdaBoost分类器的结合,行人检测准确率超过了HOG特征,略低于CSSF,但在特征提取速度和分类器训练时间上取得显著提升;由于降低了特征维度,ECSSF的召回率略低于HOG,如何提高其召回率和避免误检是作者下

一步研究的方向。表示不同比较项目的最优结果,其中准确率和召回率的计算分别按照式(5)和(6)。本文提取ECSSF速度明显优于HOG,当步进长度为4时,ECSSF维度为14 256,在LUV颜色空间分类效果最好,特征提取速度相对HOG特征提高了1倍,准确率比HOG特征提高了0.3%,但召回率略有下降(约0.2%),分类器训练时间为32.3 h,相对于HOG特征的6.1 h慢了许多;当步长为8时,ECSSF维度为4 032,本文方法特征提取速度相对HOG特征提高了5倍,其准确率和召回率略有降低(分别为0.1%和0.2%),分类器训练时间为6.7 h,与HOG特征训练时间基本相同;步进长度为8提取的ECSSF相对于步长为4提取的ECSSF,提取的特征维度明显降低,分类器训练速度提高了4倍,但检测效果基本不变(召回率不变,准确率仅下降0.4%);本文采用固定矩形块大小的方式对CSSF进行了仿真,相对文献[16]的CSSF的千万级维度,本文提取的CSSF维度只有6 912,其特征提取速度大大提升,但与步长为8的ECSSF相比,其特征提取速度仍慢了许多,分类器训练时间亦是如此;本文提取的ECSSF检测效果与HOG相当,与文献[16]给出的千万级CSSF相比效果略有下降,但ECSSF在提取速度和分类器训练耗时与两者对比具有显著优势,这对于视频行人检测这类对速度要求较高的应用来说,本文的方法具有良好的实用价值。

$$\text{准确率} = \frac{\text{正确检测行人数目}}{\text{正确检测行人数目} + \text{误判行人数目}} \times 100\% \quad (5)$$

$$\text{召回率} = \frac{\text{正确检测行人数目}}{\text{总行人样本数}} \times 100\% \quad (6)$$

一步研究的方向。

参考文献:

- [1] ENZWEILER M, GAVRILA D M. Monocular pedestrian detection: survey and experiments[J]. Pattern Analysis and Machine Intelligence, 2009, 31(12): 2179-2195.
- [2] MUNDER S, GAVRILA D M. An experimental study on pedestrian classification[J]. Pattern Analysis and Machine Intelligence, 2006, 28(11): 1863-1868.
- [3] JIA H, ZHANG Y. A survey of computer vision based pedestrian detection for driver assistance systems[J]. Acta Automatica Sinica, 2007, 33(1): 84-90. (贾慧星, 章毓晋. 车辆辅助驾驶系统中基于计算机视觉的行人检测研究综述[J]. 自动化学报, 2007, 33(1): 84-90.)
- [4] SU S, LI S, CHEN S, et al. A survey on pedestrian detection[J]. Acta Electronica Sinica, 2012, 40(4): 814-820. (苏松志, 李绍滋, 陈淑媛, 等. 行人检测技术综述[J]. 电子学报, 2012, 40(4): 814-820.)
- [5] DOLLAR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: An evaluation of the state of the art[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(4): 743-761.

(下转第2070页)

- system and its applications[J]. *Mathematical and Computer Modelling*, 2005, 42(3/4): 339–351.
- [3] LIN C-J, HO W-H. An asymmetry-similarity-measure-based neural fuzzy inference system [J]. *Fuzzy Sets and Systems*, 2005, 152(3): 535–551.
- [4] LIN C-F, WANG S-D. Fuzzy support vector machines [J]. *IEEE Transactions on Neural Networks*, 2002, 13(2): 464–471.
- [5] ZHANG X. Introduction to statistical learning theory and support vector machines[J]. *Acta Automatica Sinica*, 2000, 26(1): 32–42. (张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32–42.)
- [6] XIONG H, SWAMY M N S, AHAMAD M O. Optimizing the kernel in the empirical feature space[J]. *IEEE Transactions on Neural Networks*, 2005, 16(2): 460–474.
- [7] SCHOELKOPF B, MIKA S, BURGESS C, *et al.* Input space versus feature space in kernel based methods[J]. *IEEE Transactions on Neural Networks*, 1999, 10(5): 1000–1017.
- [8] YANG Z. Research on support vector machine classification method for imbalanced datasets[D]. Harbin: Harbin Institute of Technology, 2009. (杨智明. 面向不平衡数据的支持向量机分类方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2009.)
- [9] LI J, DENG G, LI H, *et al.* The relationship between similarity measure and entropy of intuitionistic fuzzy sets[J]. *Information Sciences*, 2012, 188: 314–321.
- [10] LE C H. A relevance-based learning model of fuzzy similarity measures[J]. *IEEE Transactions on Fuzzy Systems*, 2012, 20(1): 57–68.
- [11] HWANG C-M, YANG M-S, HUNG W-L, *et al.* A similarity measure of intuitionistic fuzzy sets based on the Sugeno integral with its application to pattern recognition[J]. *Information Sciences*, 2012, 189: 93–109.
- [12] YOU M-Y, MENG G. A generalized similarity measure for similarity-based residual life prediction[J]. *Proceedings of the Institution of Mechanical Engineers Part E – Journal of Process Mechanical Engineering*, 2011, 225(3): 151–160.
- [13] VALENZUELA O, PASADAS M. A new approach to estimate the interpolation error of fuzzy data using similarity measures of fuzzy numbers[J]. *Computers and Mathematics with Applications*, 2011, 61(6): 1633–1645.
- [14] LEE S-H, PARK W-J, JUNG D-Y. Similarity measure design and similarity computation for discrete fuzzy data[J]. *Journal of Central South University of Technology*, 2011, 18(5): 1602–1608.
- [15] PURI M, RALESCU D. Differentials of fuzzy functions[J]. *Journal of Mathematical Analysis and Applications*, 1983, 91(2): 552–558.
- [16] ZADEH L. Fuzzy sets[J]. *Information and Control*, 1965, 8(3): 338–353.
- [17] BEG I, ASHRAF S. Similarity measures for fuzzy sets[J]. *Applied and Computational Mathematics*, 2009, 8(2): 192–202.
- [18] VERT J P, TSUDA K, SCHOELKOPF B. Kernel methods in computational biology[M]. Cambridge: The MIT Press, 2004.
- [19] SHEN J D, PEI Z J, LEE E S. Support vector adaptive network in the modeling of material removal rate in rotary ultrasonic machining [J]. *Journal of Manufacturing Science and Engineering*, 2008, 130(4): 041005.
- [20] HU P, ZHANG J M, PEI Z J, *et al.* Modeling of material removal rate in rotary ultrasonic machining: designed experiments[J]. *Journal of Materials Processing Technology*, 2002, 129(1/2/3): 339–344.
- [21] PEI Z J, PRBHAKAE D, FERREIRA P M, *et al.* A mechanistic approach to the prediction of material removal rates in rotary ultrasonic machining [J]. *Manufacturing Science and Engineering*, 1993, 64(2): 167–172.

(上接第2035页)

- [6] GERONIMO D, LOPEZ A M, SAPP A D, *et al.* Survey of pedestrian detection for advanced driver assistance systems[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(7): 1239–1258.
- [7] VIOLA P, JONES M J, SNOW D. Detecting pedestrians using patterns of motion and appearance[J]. *International Journal of Computer Vision*, 2005, 63(2): 153–161.
- [8] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]// CVPR 2005: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2005, 1: 886–893.
- [9] ZHU Q, YEH M-C, CHENG K-T, *et al.* Fast human detection using a cascade of histograms of oriented gradients [C]// CVPR 2006: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2006, 2: 1491–1498.
- [10] LIN Z, HUA G, DAVIS L S. Multiple instance feature for robust part-based object detection[C]// CVPR 2009: Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2009: 405–412.
- [11] WANG X, HAN T X, YAN S. An HOG-LBP human detector with partial occlusion handling[C]// ICCV 2009: Proceedings of the 12th IEEE International Conference on Computer Vision. Piscataway: IEEE, 2009: 32–39.
- [12] WU B, NEVATIA R. Detection and segmentation of multiple, partially occluded objects by grouping, merging, assigning part detection responses[J]. *International Journal of Computer Vision*, 2009, 82(2): 185–204.
- [13] DOLLÁR P, TU Z, PERONA P, *et al.* Integral channel features [C]// BMVC 2009: Proceedings of British Machine Vision Conference. London: BMVA Press, 2009: 1–11.
- [14] WU J, GEYER C, REHE J M. Real-time human detection using contour cues [C]// ICRA 2011: Proceedings of the 2011 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2011: 860–867.
- [15] WALK S, MAJER N, SCHINDLER K, *et al.* New features and insights for pedestrian detection [C]// CVPR 2010: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2010: 1030–1037.
- [16] ZENG B, WANG G, LIN X. Color self-similarity feature based real-time pedestrian detection [J]. *Tsinghua Science and Technology: Science and Technology*, 2012, 52(4): 571–574. (曾波波, 王贵锦, 林行刚. 基于颜色自相似度特征的实时行人检测[J]. 清华大学学报: 自然科学版, 2012, 52(4): 571–574.)
- [17] VIOLA P, JONES M J. Robust real-time face detection[J]. *International Journal of Computer Vision*, 2004, 57(2): 137–154.