

DOI:10.15923/j.cnki.cn22-1382/t.2017.5.04

基于人体部位行人检测算法

李秀华, 王磊, 宋立明, 喻文举

(长春工业大学 计算机科学与工程学院, 吉林 长春 130012)

摘要: 首先运用 haar-like 特征集, 使用 Adaboost 算法训练全身和头肩、躯干和腿等部位检测器。全身检测器产生行人候选区域, 再划分头肩、躯干和腿的待检测区域, 然后将检测区域的检测结果组合分析, 最后确定是否为行人。移动摄像机拍摄的视频图像结果表明, 该方法在复杂街道上具有良好的检测效果。

关键词: Adaboost 算法; 部位检测器; 行人检测

中图分类号: TP 391.41 **文献标志码:** A **文章编号:** 1674-1374(2017)05-0433-05

Research on component-based pedestrian detection

LI Xiuhua, WANG Lei, SONG Liming, YU Wenju

(School of Computer Science & Engineering, Changchun University of Technology, Changchun 130012, China)

Abstract: With haar-like dataset, the Adaboost algorithm is used to train full-body detector and detectors of part of body such as head-shoulder, torso and leg, and then the pedestrian candidate areas were generated by full-body detector and the pedestrian candidate areas were divided into three parts of head-shoulder, torso and leg. The three body part detectors with the head, torso and leg were detected the corresponding parts of the area and analyze the combination of detection results, to further determine the pedestrian candidate whether people. Experimental results show the method has high performance in detecting pedestrians in cluttered background.

Key words: Adaboost algorithm; part detector; pedestrian detection.

0 引言

图像处理技术应用在我们生活的方方面面, 人脸识别、医学图像处理、图像检索、行人检测等都离不开图像处理技术。每年我们国家发生很多交通事故, 驾驶员负有很大责任, 随着汽车技术和

行人检测技术的发展, 行人检测技术被应用于汽车安全辅助驾驶系统来提醒驾驶员, 减少交通事故。目前行人检测的方法主要有对人的全身检测和对人体的部位检测两大类。Viola 和 Jones^[1]介绍一种在单个人体检测器中集成运动信息和外貌信息的人体检测系统。它的误报率很低, 但是

收稿日期: 2017-07-21

基金项目: 吉林省科技厅基金资助项目(20160312002ZG)

作者简介: 李秀华(1971-), 女, 汉族, 吉林长春人, 长春工业大学副教授, 博士, 主要从事图像处理与智能控制方向研究, E-mail: lixiuhua@ccut.edu.cn.

这种方法检测静止的摄像机所拍摄的图片比较有效,当摄像机运动速度很快时,连续帧的背景会发生变化,它的误报率会非常高。基于整体的检测方法在图片比较简单且行人遮挡不严重的情况下检测效果比较好。Papageorgiou^[2]等的 SVM 检测器、Gavrila^[3]的边缘模板检测器和 Wu^[4]等的 Markov 检测器等。在行人出现严重遮挡时,系统的性能也会明显下降,而基于人体部位的检测方法则表现出了良好性能。Anuj Mohan^[5]等介绍了一种基于部位组合(Voting Combination Classifiers)的人体检测方法,称之为分类器(VCC)的投票组合,从图像中提取 haar-like 特征,然后对其进行训练分类,该方法的实验结果表明,其检测性能明显优于基于整个人体的检测算法。

在文献[5]的基础上,提出了一种在单帧图片上人体部位组合的算法。将人体分为头肩(Head-Shoulder)、躯干(Torso)和腿(Leg)。该算法在满足实时性的同时也提高了检测率。

1 基于部位的行人检测算法

文中采用了 haar-like 特征集,haar-like^[6]特征计算简单,特别是引入积分图后,计算量大大降低,在稀疏行人且遮挡不严重的情况下检测效果较好,实时性较高,用 Adaboost 算法训练行人检测系统的 Full-body 检测器和头肩、躯干和腿等部位检测器。Adaboost 算法是先训练出一系列弱分类器,然后将这些弱分类器组合成强分类器,最后将多个强分类器组合成最终的部位检测器。

文中的行人检测算法只在相应的待检测区域进行检测,检测时间较少;将各个部位检测的结果进行组合分析,尽可能的提高检测率;当行人某个部位被遮挡时也能检测出行人。

1.1 算法步骤

部位检测器分割图如图 1 所示。

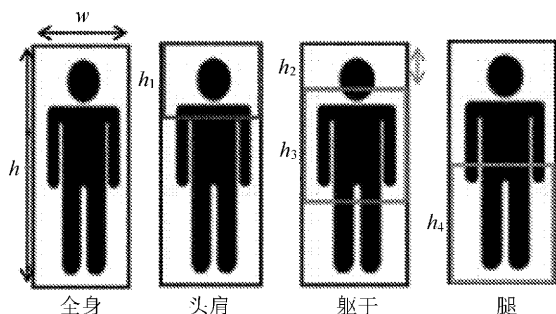


图 1 部位检测器分割图

$$h_1 = \frac{7h}{24}$$

$$h_2 = \frac{5h}{48}$$

$$h_3 = \frac{25h}{48}$$

$$h_4 = \frac{h}{2}$$

文中的行人检测系统分为两个过程。在第一个过程中,用 full-body 检测器在单帧图像上产生行人候选区域。在第二个过程中,参考文献[7]中的方法,根据图 1 所划分的人体部位位置及比例,在每个候选行人区域划分部位的检测区域,因为正常情况下按照自然人体部位几何约束条件,人的头和肩只可能出现在整个人体检出框的上半部分位置,腿只可能出现在整个人体检出框下半部分位置,而躯干只可能出现在它们之间的位置。这样部位检测器只会各自的检测区域内检测而不必检测整个人体,大大提高了检测效率。各个部位检测器可能会检测出多个检出框,依据检出框的几何信息比较检出框响应值的大小,选择出最好的检出框。最后,将各个部位的检测结果进行组合分析,判断候选者是否为人。

算法的具体步骤如下:

- 1)用 full-body 检测器产生候选行人;
- 2)为候选行人设置人体部位检测区域;
- 3)用头肩、躯干和腿等部位检测器在图像中的相应位置进行检测;
- 4)将头肩、躯干和腿等部位检测器的检出结果组合分析判断候选者是否为人。

1.2 部位检测器结合策略

在文献[5]中介绍了一种 VCC(Voting Combination Classifiers)的方法,即投票组合分类器。VCC 是通过一个投票结构来组合分类器。头肩、躯干和腿等部位检测器可以看成是投票者。VCC 系统对投票者进行调查,然后根据结果决定是否为人。只有当所有 3 个组件类都在适当的配置中被检测到时,系统才会对投票结果进行评估,评估按如下所示的逻辑:

$$\text{Person} = \text{Head-shoulder} \ \& \ \text{Torso} \ \& \ \text{Leg} \quad (1)$$

当 Person=true 时,系统则判定为人。

在此基础上增加了部位检测器输出响应值与检测器阈值的差值评估,形成新检测算法,以进一步提高准确率,检测算法有两个递进的约束条件,算法如下:

1)在候选人 X 中,判定候选人头肩、躯干和腿都能检测出的条件为:

$$P = Y_{HS} \& Y_T \& Y_L \quad (2)$$

2)当人体部位 i 被检测出时用 Y_i 表示,人体的各个部位检测区域用 x_i 表示, $S(x_i)$ 表示各个部位检测器输出响应值, $i \in \{HS, T, L\}$ 。则部位检测器与检测器阈值的差值关系式如下:

$$M(x_i) = S(x_i) - T_i \quad (3)$$

T_i 是 i 部位检测器的阈值,则进行如下判

$$w_i = \begin{cases} \frac{e^{-\left\{ \left(\frac{(x_{si}-x_{di})^2}{\sigma_{xi}^2} + \frac{(y_{si}-y_{di})^2}{\sigma_{yi}^2} \right) + \lambda_i \left(\frac{(H_{si}-H_{di})^2}{\sigma_{Hi}^2} + \frac{(W_{si}-W_{di})^2}{\sigma_{Wi}^2} \right) \right\}}}{2}, & Y_i = 1 \\ 0, & Y_i = 0 \\ i \in \{HS, T, L\} \end{cases} \quad (6)$$

候选人实际部位 i 的质心和大小分别为 (x_{di}, y_{di}) 和 (W_{di}, H_{di}) , 候选人检测出的部位 i 的质心和大小分别为 (x_{si}, y_{si}) 和 (W_{si}, H_{si}) , λ_i 为常数,用于权衡部位检出框的位置和大小。

最后,组合检测器的输出为:

$$H(X) = \begin{cases} 1, & F(X) > t \text{ is person} \\ 0, & \text{otherwise is not person} \end{cases} \quad (7)$$

2 仿真实验结果

文中训练和测试所用的样本图片均来自 MIT 和 INRIA 数据库以及 MIT 数据库的行人镜像图片,负样本图片来自 CBCL streetscenes challenge Images 数据库,训练用的正负样本分别为 1 800 个人体图片和 5 400 张不包括人的自然街道和公路图片。采用样本尺寸为 24×48 像素训练 full-body 检测器,将这些正负训练样本图片均按照图 1 所示的部位位置比例裁剪下来组成各个部位检测器的训练样本。这些行人样本是用不同摄像机在不同光照条件下和不同视点拍摄的不同姿势、不同颜色衣服、不同运动方式在街道上行走的人。测试用的正负样本各 300 张,每张图像的大小为 64×128 像素。许多图像的背景非常杂乱,有许多类似人的物体,如树木、邮箱、垃圾桶、路牌、路灯及建筑物窗户等。

将部位检测器单独进行测试并将结果绘制成 ROC 曲线,如图 2 所示。

该曲线说明了不同部位检测器识别图像中人的能力,腿表现最好,其次是头肩部和躯干。腿的优越性能是由于图像中下部分身体的背景通常是

决:

$$Y_i = \begin{cases} 1, & M(x_i) > 0 \\ 0, & M(x_i) < 0 \end{cases} \quad i \in \{HS, T, L\} \quad (4)$$

令 $Y_i \in \{0, 1\}$, $i \in \{HS, T, L\}$, 若人体部位 i 未检出,则 $Y_i = 1$; 若人体部位 i 被检出,则 $Y_i = 0$ 。我们能构造如下一个函数:

$$F(X) = \sum w_i \cdot M(x_i) \quad (5)$$

其中 w_i 按照下式计算:

街道或草地,而头部和手臂的背景相对来说比较复杂。躯干姿态变化较大,一些类似柱状物体容易错认为躯干,但是头肩部有相对于躯干的优势;姿态一般变化不大,人在运动时躯干有时会受到手臂摆动的影响,而头肩的形状相对比较稳定。每个部位的检测结果以及组合后的结果如图 3 所示。

将基于整个人体的检测系统^[8] VCC 和文中提出的行人检测系统在测试集上进行测试,检测结果见表 1。

由于基于整个人体的检测系统结构比较简单,处理一帧大小为 (64×128) 的图像平均耗时仅 11.416 4 ms,但是误检率比较高。文中方法虽然仅比 VCC 系统快几毫秒,可以说基本一样,但能满足实时检测的要求。3 种方法中,文中方法检测率最高,VCC 系统次之,基于整个人体的检测系统检测率最低。

对 ROC 曲线的分析表明,由于对人体部位进行组合,VCC 系统的检测性能相对于基于整个人体的检测系统有所提高,而文中方法对 VCC 系统进行了改进,测试结果表明,文中方法取得了良好的检测效果,系统的检测性能优于 VCC 系统。城市街道两边杂乱的树木、路牌、垃圾箱、消防设备、路灯及建筑屋的窗户与行人非常类似,行人检测的过程中会形成一些误报,文中将人体的几何信息和行人部位检测器的响应值相结合,能减少这种误报,行人检测算法在街道复杂场景中的检测结果如图 4 所示。

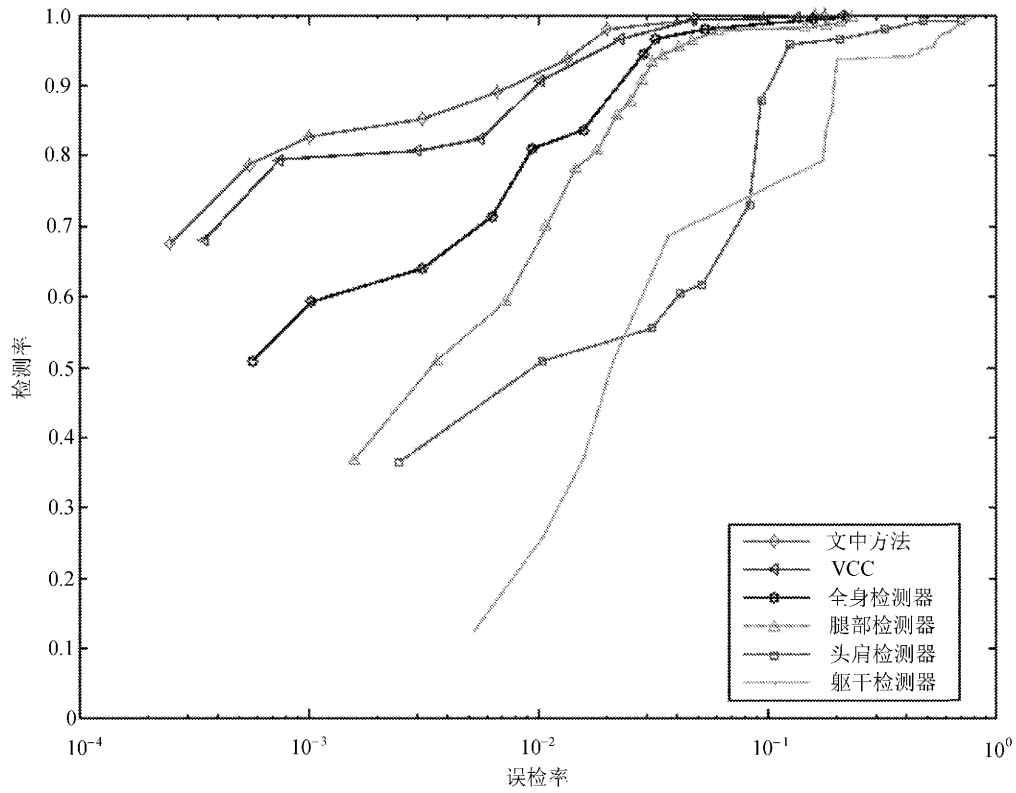


图 2 各种方法的 ROC 曲线

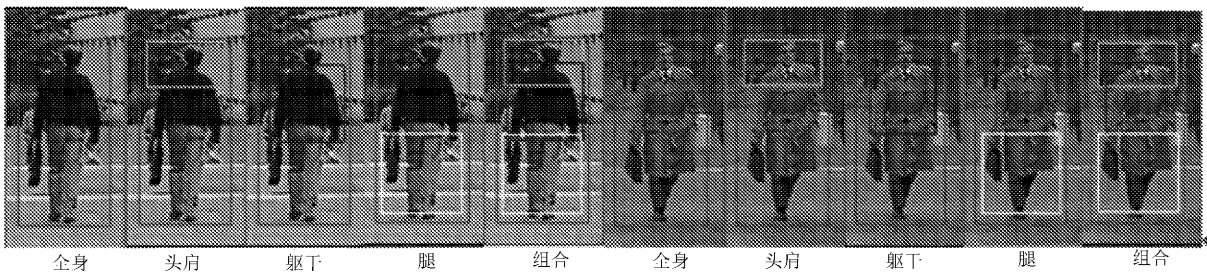


图 3 各个部位检测器的检测结果

表 1 不同方法的性能比较

方法	人体图片	非人体图片	检测率/%	误检率/%	检测时间/ms
全身检测	300	300	94.636	3.030	11.416 40
VCC	300	300	95.439	2.250	45.120 90
文中方法	300	300	96.100	1.365	40.503 35

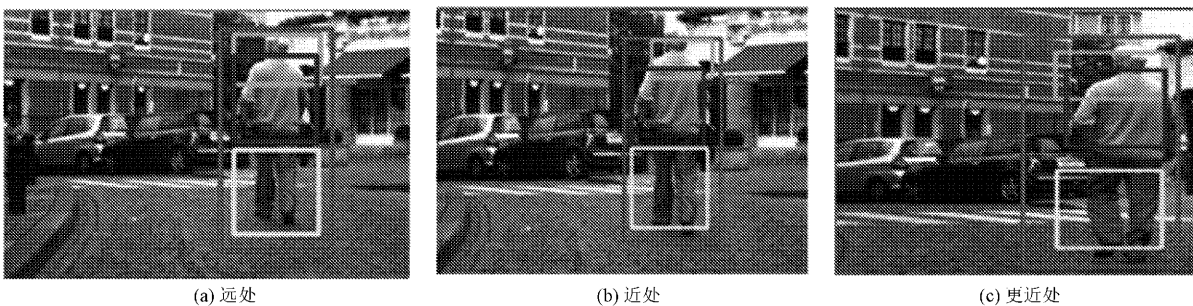


图 4 复杂场景中行人检测结果图像

3 结 语

利用人体几何约束条件,将人体分割成头肩、躯干和腿3个部位,其检测器只在各自限定的部位区域进行检测,能有效解决行人的遮挡问题,检测速度也得到了提高,在数据集上的检测结果表明,该方法能够很好地检测道路上的行人,文中算法表现出的检测性能也是3种方法中最好的。道路上行人检测要求实时性,文中方法只采用了最简单的 haar-like 小波特征,这对行人的检测性能也产生了一定影响,在以后的研究工作中尝试将一些具有更好性能的特征进行组合,来取得更好的检测效果。

参考文献:

- [1] 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.
- [2] Papageorgiou C, Poggio T. A trainable system for object detection [J]. International Journal of Computer Vision, 2000, 38(1): 15-33.
- [3] Gavrila D M. Pedestrian detection from a moving vehicle [J]. Proceedings of European Conference on Computer Vision (EC-CV), 2000, 1843: 37-49.
- [4] Wu Ying, Yu Ting, Hua Gang. A statistical field model for pedestrian detection [J]. IEEE Computer Society Conference on Computer Vision (ECCV), 2004(1): 69-81.
- [5] Mohan A, Papageorgiou C, Poggio T. Example-based object detection in images by components [J]. IEEE Trans on PAMI, 2001, 23(4): 394-361.
- [6] 程思宁,武丹丹,张俊芬,等.基于小波的人脸识别算法[J].长春工业大学学报:自然科学版,2006,27(1):4-7.
- [7] WU B. Detection of multiple partially occluded humans in a single image by bayesian combination of edgelet part detectors [J]. IEEE International Conference on Computer Vision, 2005(1): 90-97.
- [8] Monteiro G, Peixoto P. U nunes vision-based pedestrian detection using haar-like features [J]. Robótica, 2006(5): 45-50.