

基于腿部梯度方向直方图特征优化的行人检测^①郭 烈^{②*} 赵宗艳* 赵一兵* 杨 成** 张广西*

(* 大连理工大学汽车工程学院 大连 116024)

(** 山东英才学院汽车工程学院 济南 250104)

摘 要 针对传统的基于梯度方向直方图(HOG)特征的行人检测耗时较长的问题,提出了基于腿部 HOG 特征优化的行人检测方法。该方法采用加权 Fisher 线性判别(WLFD)代替线性 SVM 来选择最具区分性的 HOG 特征,在保持分类能力的同时减少训练时间和存储空间,而且选择查找表型弱分类器的 Gentle Adaboost 算法来训练优化权重组合 HOG 特征,形成一个强分类器来检测行人。通过对线性 SVM、加权 Fisher 与阈值型以及加权 Fisher 与查找表型三种弱分类器的对比试验表明,基于加权 Fisher 与查找表型 HOG 特征优化后不仅提高了检测精度,而且训练时间和检测时间也能明显降低。

关键词 交通安全, 行人检测, 梯度方向直方图(HOG), Gentle Adaboost

0 引 言

随着现代交通安全性要求的日益提高,汽车安全技术逐渐向智能化方向发展^[1]。为避免碰撞行人采用的基于视频的主动预警是智能化技术之一。主动预警的关键技术之一是如何从传感器信息中分析得到行人信息。目前基于视觉传感器的行人检测方法有多种,其中基于统计分类的方法由于能够较好地克服行人多样性、场景多样性、光照环境多样性等不利条件的影响而被广泛采用。2005 年 Dalal 等^[2]提出了用梯度方向直方图(histograms of oriented gradients, HOG)特征来描述行人的方法,该方法尽管检测速度慢,由于其检测精度高,逐渐在行人检测领域得到应用和完善^[3-5]。Dalal 等^[6]进一步用线性支持向量机(support vector machine, SVM)训练 HOG 和光流场方向直方图特征,实现了运动及变化背景情况下的行人检测。Zhu 等^[7]针对不同尺度大小的块提取 HOG 特征并使用 Adaboost 算法去除冗余的 HOG 特征,为了在保证精度的情况下提高检测速度,该方法随机从 5031 个 HOG 特征块中选择 250 个特征进行训练。显然这种随机挑选模式并不能保证每次循环中所挑选出的特征为全局最优,从而降低了系统的检测性能。本研究对具有典型

HOG 特征的行人腿部进行检测,利用加权 Fisher 线性判别(weighted linear Fisher discriminant, WLFD)配合查找表的 Gentle Adaboost 算法改进利用线性 SVM 作为弱分类器而导致训练时间长的缺陷。由于最大程度上利用了特征的信息,使得最后挑选出的特征个数大大减少,缩短了训练时间,达到了快速检测行人的目的。

1 算法描述

1.1 腿部区域 HOG 特征提取

目前,针对 HOG 特征的研究大都集中在整个人体样本上。事实上,针对人体及其典型部位采取何种分类器或者提取何种特征还有待进一步的研究。通过对样本图像分析发现,直立行走的行人姿态的变化主要表现为四肢处 HOG 特征的变化,头部与躯干部分的 HOG 特征则相对稳定^[8]。因此,本研究针对行人腿部进行 HOG 特征提取并对其优化来检测行人。

HOG 如图 1 所示。对于一幅待检图像,按照图示的分区策略来提取各区域的 HOG 特征。将一副待检图像分成 N 个单元(称为“cell”),由若干个单元可组合成一个块(称为“block”),整幅图像由一定数目的块组成,块与块之间允许重叠。通过统计各

① 国家自然科学基金(61104165)和辽宁省自然科学基金(20102026)资助项目。

② 男,1978 年生,博士,副教授;研究方向:汽车碰撞安全,行人保护技术;联系人, E-mail: guo_lie@163.com

(收稿日期:2012-03-01)

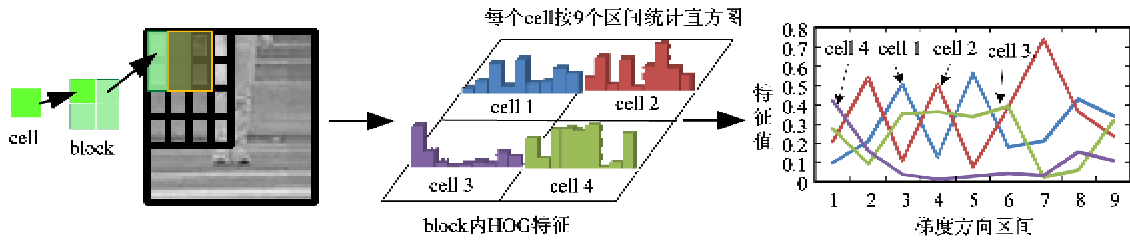


图1 腿部区域 HOG 特征提取示意

个块内各单元在各个方向上的梯度分布,可构成该图像的 HOG 特征^[2]。

通过试验发现,将腿部周围背景像素点数由 16 降至 8 时,检测效果无较明显变化;将腿部周围背景全部去掉时,检测效果则出现了明显的降低。可见背景的存在使得腿部显著特征更为突出,增加了对比度,有助于提高检测效果。因此,选择腿部区域窗口像素大小为 64×64 ,采用 Sobel 算子从样本集中提取所有的 HOG 特征,按照文献[7]所示的分块原则,最终从每个样本中提取出 1386 个 block 的 HOG 特征。

1.2 基于加权 Fisher 线性判别的 HOG 特征优化

采用线性 SVM 作为弱分类器选择典型 HOG 特征时,由于线性 SVM 的训练时间过长,且每个梯度直方图特征是高维的特征,文献[7]在每次循环中随机从 5031 个梯度直方图特征块中选择 250 个特征进行训练,使得其不能保证每次循环中所挑选出的特征为全局最优,降低了系统的检测性能。参考 Laptev^[9]提出用加权 Fisher 线性判别(WLFD)方法来提取有效的 HOG 特征。WLFD 方法是将样本多维特征投影到一维直线上,投影方向的选择则是能将各类样本很好地分开,因此该投影过程具有一定的弱分类功能,同时实现特征降维。投影方向的计算如下:

$$\bar{m} = \frac{1}{n \sum_{i=1}^n w_i} \sum_{i=1}^n w_i f(x_i) \quad (1)$$

$$\bar{S} = \frac{1}{(n-1) \sum_{i=1}^n w_i^2} \sum_{i=1}^n w_i^2 (f(x_i) - \bar{m})(f(x_i) - \bar{m})^T \quad (2)$$

$$a = (\bar{m}^{(1)} - \bar{m}^{(2)}) (\bar{S}^{(1)} + \bar{S}^{(2)})^{-1} \quad (3)$$

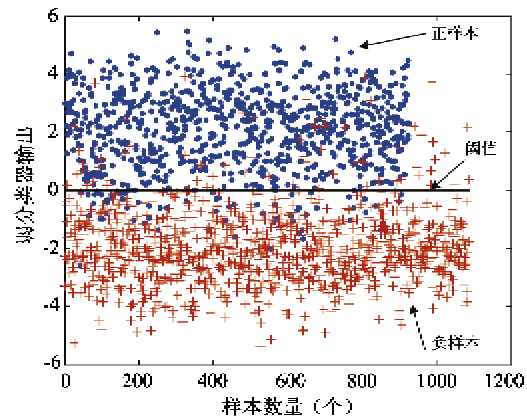
其中 w_i 是第 i 个样本的权重, $f(x_i)$ 是第 i 个样本的特征, $\bar{m}$ 是样本类内均值, $\bar{S}$ 是样本类内加权协方差矩阵, a 是最佳投影方向。

WLFD 能够实现的一个重要前提是每类特征在空间上的分布应该符合高斯分布。然而,在实际应

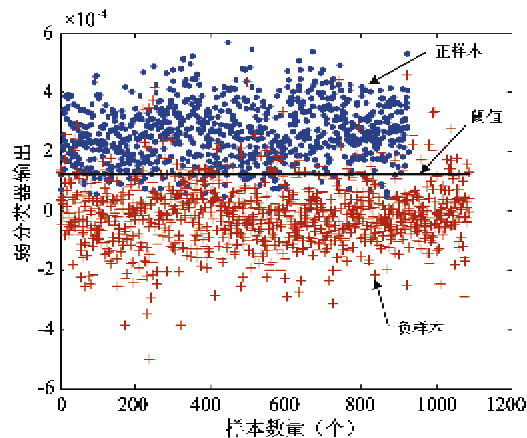
用过程中,每类特征在空间上的分布通常是非高斯的、多模的。但可以设定在大量的特征中,必然存在某些具有良好分类性能的特征,其分布近似于高斯分布,因此,WLFD 应该可以实现投影和分类的作用。假定投影后的样本集分别服从参数为 (u_1, σ_1) 和 (u_2, σ_2) 的高斯分布,则根据贝叶斯分类原理,在两类具有相同的先验概率情况下的最优分类阈值为^[10]

$$Threshold = \frac{\sigma_1 \sigma_2}{\sigma_1 + \sigma_2} \ln \left(\frac{\sigma_1}{\sigma_2} \right) + \frac{\sigma_1 \mu_2 + \sigma_2 \mu_1}{\sigma_1 + \sigma_2} \quad (4)$$

为了比较两种弱分类器分类能力,图 2 比较了线性 SVM 与 WLFD 分别对同一个特征块的分类结



(a) 线性 SVM 对样本集分类结果



(b) WLFD 对样本集分类结果

图2 两种弱分类器对同一特征块的分类结果

果,圆点表示正样本,十字号表示负样本,直线表示分类阈值。从图中可以看出,WLFD 的分类性能稍弱于 SVM,但也表现优异。另一方面,由于 Adaboost 算法本质上就是将弱分类器结合成强分类器的算法,故可以弥补该方法的不足。在计算复杂度上,WLFD 与样本数和特征维数呈线性关系,因而有着速度上的巨大优势。这样便解决了分类器训练时的随机采样问题,在每次循环中可以使用全部的 1386 个特征,弥补了随机抽取样本特征而造成有效特征丢失的缺陷。

1.3 基于查找表型 Gentle Adaboost 算法的分类器训练

Laptev^[9]在 Adaboost 算法中采用了形式简单的阈值型(Stump)弱分类器,分类器输出为两个离散值,对于复杂特征的分类能力相对较弱。Wu 等^[11]提出用查找表(look up table, LUT)来替代阈值型(Stump)弱分类器的方法,实现了实值 Adaboost 算法分类器训练。为了增强弱分类器的表达能力,采用 Gentle Adaboost 算法进行训练^[12],其弱分类器的输出函数为

$$h(x) = P_w(y = 1 | x) - P_w(y = -1 | x) \quad (5)$$

其中 $P_w(y = 1 | x)$, $P_w(y = -1 | x)$ 分别表示特征值为 x 的样本时,其为正样本或负样本的加权概率。

而查找表的方法则是先将整个样本集的特征值划分为对应的 n 个不相交的子集,对于某个特征值为 x 的样本,其弱分类器的输出就表示为它在所属子集中为正样本与负样本的加权概率差。因此,查找表型弱分类器的输出为一个表示置信度的实数值,显然与离散型输出相比,它包含了更多的信息并且更加适用于实际样本的分布情况,能够有效地提高分类器的精度^[3]。采用基于查找表型 Gentle Adaboost 算法来训练行人腿部检测分类器,选择带有权重的 HOG 特征结合形成一个强分类器。由于只是从全部 1386 个特征中挑选某些最具区分性的特征作为目标的特征进行训练和识别,从而大大降低了训练和识别的时间。下面给出查找表型弱分类器的构造过程。

经过加权 Fisher 线性判别(WLFD)后,HOG 特征被降为 1 维,首先将它归一化到 $[0, 1]$ 区间,将该区间均匀分成 n 个子区间:

$$bin_j = [(j-1)/n, j/n], \quad j = 1, 2, \dots, n \quad (6)$$

这样,弱分类器可以定义为:

$$\text{if } f(x) \in bin_j \text{ then } h(x) = W_j^{+1} - W_j^{-1} \quad (7)$$

其中 W_j^{+1} 和 W_j^{-1} 分别表示正、负样本的特征值落在

区间 bin_j 内的条件概率,可表示为

$$W_j^l = P_w(f(x) \in bin_j, y = l) \quad l = \pm 1, j = 1, 2, \dots, n \quad (8)$$

具体地说,令 D_j^{+1} 和 D_j^{-1} 分别表示第 j 个区间 bin_j 内正负样本权值的总和

$$D_j^{+1} = \sum_{f(x_i) \in bin_j, y_i = 1} w_i \quad (9)$$

$$D_j^{-1} = \sum_{f(x_i) \in bin_j, y_i = -1} w_i$$

则条件概率为

$$P_w(f(x) \in bin_j, y = 1) = \frac{D_j^{+1}}{D_j^{+1} + D_j^{-1}} \quad (10)$$

$$P_w(f(x) \in bin_j, y = -1) = \frac{D_j^{-1}}{D_j^{+1} + D_j^{-1}}$$

定义函数

$$Bin_n^j(u) = \begin{cases} 1 & u \in [j-1/n, j/n) \\ 0 & u \notin [j-1/n, j/n) \end{cases} \quad j = 1, 2, \dots, n \quad (11)$$

这样基于查找表型 Gentle Adaboost 弱分类器可表示为

$$h_{LUT}(x) = \frac{D_j^{+1} - D_j^{-1}}{D_j^{+1} + D_j^{-1}} Bin_n^j(f(x)) \quad (12)$$

2 实例研究

本文针对文献[7]在设计弱分类器方面进行了以下改进:一是采用加权 Fisher 线性判别(WLFD)来代替线性 SVM,从而降低 HOG 特征维数;二是使用查找表(LUT)型弱分类器替代传统阈值型(Stump)弱分类器。因此,在实例研究中比较了基于线性 SVM 分类器、加权 Fisher 线性判别的阈值型弱分类器(WLFD + Stump)和加权 Fisher 线性判别的查找表型弱分类器(WLFD + LUT)三种弱分类器的 Gentle Adaboost 算法训练得到的行人检测分类器在使用特征的数量上以及检测精度上的差别。样本由 INRIA 行人库和手工从视频中分割的样本组成并分成两组,一组是含 1000 个行人和 1000 个非行人的训练样本集;另一组是含 1000 个行人和 800 个非行人的测试样本集。

图 3 是经过上述训练后挑选出的前三个特征块,其定位出行人的左腿、右腿以及脚与地面接触的位置。从图中可以看出,由前两种弱分类器挑选出的特征只包含了腿部的部分信息,而第三种弱分类器挑选出的前三个特征就包含了全部腿部信息,我们可以推断该种弱分类器的检测性能最为优异。

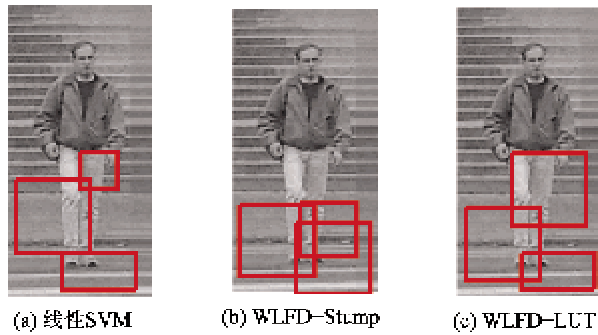


图3 训练得到的前三个特征块

图4是训练得到的强分类器的性能,图中线性SVM所使用的弱分类器达到了49个,而WLFD所使用的弱分类器只有15个,迭代次数大大减少。WLFD不仅解决了每次迭代随机选取部分HOG特征的问题,还能提高其分类速度。在训练过程中可以使用全部1386个HOG特征块进行训练,保证初始选取的特征为全局最优,从而降低了迭代次数。从图中还可看出,LUT型弱分类器的收敛速度相对较快,使用更少的特征就能达到目标检测率。图5

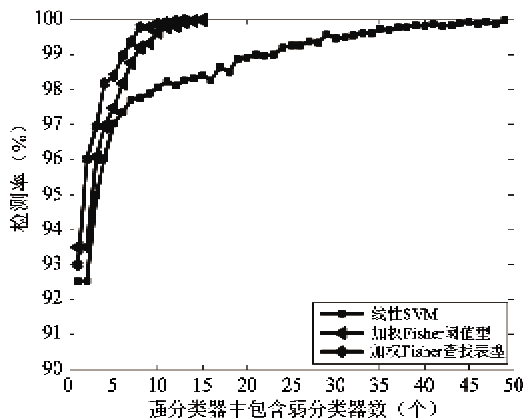


图4 训练得到的强分类器性能比较

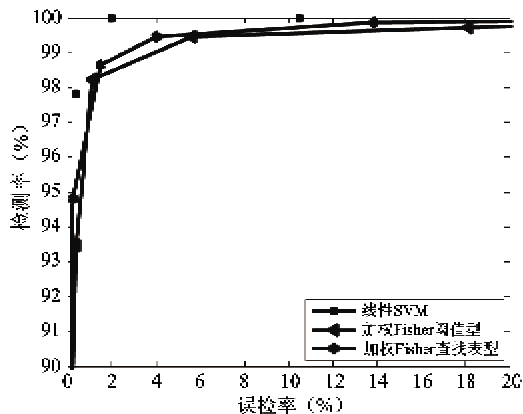


图5 不同弱分类器对腿部区域进行测试的ROC曲线

比较了不同类型弱分类器的接收者操作特征(receiver operating characteristic, ROC)曲线,使用查找表型弱分类器后的Adaboost检测器的精度与阈值型弱分类器相比,精度有所提高,能更进一步优化强分类器,提升检测性能。而使用线性SVM弱分类器方法,由于采用的特征数量较多,其检测率能比其它两种方法高一些,但特征数量的增多会使算法的实时性降低。

表1列出了三种不同分类器的时间消耗及分类性能,计算机为双核2.1GHz的CPU、内存2G。从表中可以看出基于WLFD的查找表型弱分类器分类性能达到94.4%,要优于线性SVM与WLFD的阈值型分类器。在时间消耗方面,与线性SVM相比,使用WLFD的两种弱分类器的训练时间分别减少了88%和77%,验证了其在运算速度上的巨大优势。单独分析查找表型弱分类器,其训练时间是简单阈值型弱分类器的两倍,主要是由于查找表的方法在实际应用过程中,不仅需要记录查找表的每个区间的系数,还需要记录特征的最大值和最小值。这些都会占用存储空间,同时还会增加一些运算量。查找表型弱分类器由于使用了更少的特征,在检测阶段的速率并没有下降。

表1 不同弱分类器时间消耗与分类性能的比较

	线性 SVM	加权 Fisher 与单阈值型	加权 Fisher 与查找表型
检测时间(s)	7.719	0.023	0.025
训练时间(s)	4979.6	592.62	1141.4
检测率(%)	94	93.8	94.4

3 结论

由于在视频图像中,行人腿部区域具有明显的HOG特征。根据该特点,利用Gentle Adaboost算法用于对高维HOG特征进行特征优化,挑选出更具有“语意”的特征块来表征行人。在设计弱分类器时,采用加权Fisher线性判别(WLFD)来代替线性SVM,从而降低HOG特征维数,保证在训练过程中可以使用全部的特征,从而加快分类器收敛速度,大大缩短了训练时间。并用LUT型弱分类器替代传统阈值型弱分类器,提高了分类器的精度。最后通过线性SVM、加权Fisher线性判别的阈值型弱分类器和加权Fisher线性判别的查找表型弱分类器进行对比试验,结果表明,本文提出的基于查找表型的加

权 Fisher 线性判别弱分类器方法具有较高的检测精度和较低检测时间,可以较为准确地定位行人的位置。研究发现,该方法对于遮挡问题的处理还有待进一步提高,下一步的工作拟尝试引入人体其他部位的有效特征进行分类,并结合腿部检测结果来提高检测率和准确度。

参考文献

- [1] 宋健, 王伟玮, 李亮等. 汽车安全技术的研究现状和展望. 汽车安全与节能学报, 2010, 1(2): 98-106
- [2] Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005. 886-893
- [3] 朱文佳, 戚飞虎. 基于 Gentle Adaboost 的行人检测. 中国图象图形学报, 2007, 12(10): 1905-1908
- [4] Watanabe T, Ito S, Yokoi K. Co-occurrence histograms of oriented gradients for pedestrian detection. In: Proceedings of the 3rd Pacific-Rim Symposium on Image and Video Technology, Tokyo, Japan, 2009. 37-47
- [5] Pedersoli M, González J, Bagdanov A D, et al. Efficient discriminative multiresolution cascade for real-time human detection applications. *Pattern Recognition Letters*, 2011, 32(13): 1581-1587
- [6] Dalal N, Triggs B. Human detection using oriented histograms of flow and appearance. In: Proceedings of the 9th European Conference on Computer Vision, Graz, Austria, 2006. 2: 428-441
- [7] Zhu Q, Avidan S. Fast human detection using a cascade of histograms of oriented gradients. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, 2006. 1491-1498
- [8] 曾春, 李晓华, 周激流. 基于感兴趣区梯度方向直方图的行人检测. 计算机工程, 2009, 35(20): 182-184
- [9] Laptev I. Improvements of object detection using boosted histograms. *Journal of Image and Vision Computing*, 2009, 27: 535-544
- [10] Wang H, Li P, Zhang T. Histogram features-based fisher linear discriminant for face detection. *Neural Computing and Applications*, 2008, 17(1): 49-58
- [11] Wu B, Ai H Z, Huang C. LUT-based Adaboost for gender classification. In: Proceedings of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication, Guildford, UK, 2003. 104-110
- [12] Freund Y, Schapire E R. A decision-theoretic generalization of on line learning and an application to boosting. In: Proceedings of the 2nd European Conference on Computational Learning Theory, Barcelona, Spain, 1995. 23-37

Pedestrian detection based on optimized HOG features of legs

Guo Lie*, Zhao Zongyan*, Zhao Yibing*, Yang Cheng**, Zhang Guangxi*

(* School of Automotive Engineering, Dalian University of Technology, Dalian 116024)

(** Department of Automotive Engineering, Shandong Yingcai University, Jinan 250104)

Abstract

To solve the time consuming problem of traditional pedestrian detection methods based on histograms of oriented gradients (HOG) features, a novel pedestrian detection method based on optimization of the HOG features of legs is presented. The method uses weighted linear Fisher discriminant (WLFD) instead of linear support vector machine (SVM) to construct weak classifiers with the aim of selecting high discriminative HOG features, which can significantly decrease the training time and memory while maintaining the comparable classification performance. Moreover, the look up table (LUT) Gentle Adaboost algorithm is selected to optimize the weighted combined HOG features and form a strong classifier to identify the pedestrian. The comparison test shows that the classifier of WLFD with LUT outperforms the weak classifiers of the linear SVM and the WLFD with stump. When the HOG features are optimized by the classifier of WLFD with LUT a higher detection accuracy with lower training and detection time can be achieved.

Key words: traffic safety, pedestrian detection, histograms of oriented gradients (HOG), Gentle Adaboost