

基于 ROI 区域强分辨力 HOG 特征的视频行人检测*

张璐, 陈淑荣

(上海海事大学 信息工程学院, 上海 201306)

摘要: 针对传统梯度方向直方图(HOG)行人检测系统中检测窗扫描区域过大、HOG 特征维度大而引起的检测速度慢问题, 提出了改进的视频行人检测算法。通过运动信息提取感兴趣(ROI)目标区域, 利用 Fisher 准则和多尺度特性选取具有强分辨力的行人 HOG 特征从而降低特征维数, 结合支持向量机(SVM)检测行人。实验结果表明, 本文方法在保证视频行人检测的准确率的同时, 有效地提高了行人检测的速率。

关键词: 行人检测; 感兴趣区域; 梯度方向直方图; Fisher 准则

中图分类号: TP391.4

文献标识码: A

文章编号: 1674-7720(2013)07-0046-03

Pedestrian detection based on ROI and strong distinguished HOG

Zhang Lu, Chen Shurong

(College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China)

Abstract: The traditional pedestrian detection systems using histogram of orient gradient (HOG) exist some defects, e.g., large detection scan area and huge HOG feature dimensions which easily reduce system detection rate. In this paper, we propose an improved pedestrian detection algorithm which utilizes motion information to obtain the ROI (region of interest), then uses Fisher criterion and multi-scale method to select the strongest distinguished HOG features to reduce feature dimension, finally adopts SVM (support vector machine) to detect video pedestrians. Experimental results show that the proposed method effectively improves video pedestrian detection rate and achieves higher accuracy rate.

Key words: pedestrian detection; ROI; HOG; Fisher criterion

行人检测是计算机视觉和模式识别领域中的研究热点, 在视频监控中有广泛的应用。目前, 用于行人检测的方法主要分为 3 类: (1) 基于运动信息的方法^[1-2], 根据行人运动信息(如步态)实现行人检测, 实时性好但只限于运动行人; (2) 基于先验知识建模的方法^[3], 利用已知先验知识构建人体模板和待测目标进行匹配, 但模板的初始参数难以获取; (3) 基于特征统计分类的方法^[4-9], 根据训练数据获取有效分类器对输入检测窗口进行检测, 实现简单且检测效果较好, 是目前静态图像行人检测的主流方法, 但对视频行人检测速度较慢。因此, 如何保证视频行人检测的准确率, 并提高检测速度是一个急需解决的问题。

1 本文行人检测方法

视频录像中行人的不同姿势、复杂的背景变化对算

法的检测准确率和检测速度都提出了严格要求。本文提出利用视频中的运动信息获取 ROI 区域, 结合 Fisher 准则挑选强分辨力行人 HOG 特征, 并结合 SVM 分类器检测行人。其中, 训练部分采用正负样本库提取强分辨力 HOG 行人特征, 并通过 SVM 训练得到强分辨力 HOG 特征的行人分类器。检测时利用帧间信息获得 ROI 区域, 通过多层次缩放 ROI 区域获取所有检测窗, 并利用强分辨力行人 HOG 特征来表征检测窗中行人, 通过训练好的分类器实现对视频中行人的检测。该算法框架如图 1 所示。

1.1 视频图像 ROI 目标区域提取

视频帧图像中存在大量的非人区域, 全局扫描需要很多冗余计算。本文采用三帧差分法获取运动目标前景, 进行二值化处理、形态学处理和前景块扩展, 获取 ROI 目标区域。具体实现步骤如下。

* 基金项目: 上海海事大学校基金(20110043)

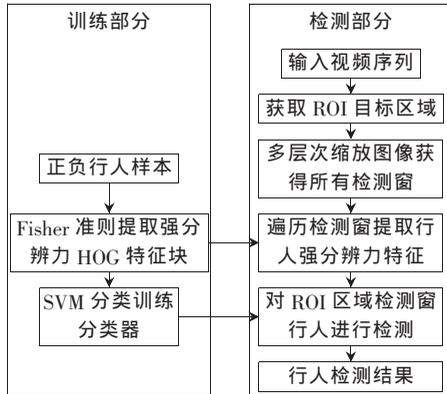


图1 行人检测系统框架

(1) 采用三帧差分法获取运动目标前景。当前帧的前景帧差图为:

$$d = \left| \frac{I_L(x, y, i-2) + I_L(x, y, i-1) - I_L(x, y, i)}{2} \right| \quad (1)$$

其中, $I_L(x, y, i)$ 、 $I_L(x, y, i-1)$ 、 $I_L(x, y, i-2)$ 分别为第 i 帧、第 $i-1$ 帧和第 $i-2$ 帧的亮度分量, d 为帧间图像对应像素点亮度差的绝对值。

(2) 对目标前景图进行二值化处理和形态学处理。二值化处理为:

$$D(x, y) = \begin{cases} 255, & d \geq D \\ 0, & d < D \end{cases} \quad (2)$$

其中, D 为初始阈值, $D(x, y)$ 为二值图对应像素点的亮度值。经过膨胀、腐蚀二值图像, 获得消除了噪声、微小运动区域的目标前景块。

(3) 根据行人体型特征参数, 对比目标前景块, 获取有效的行人目标前景块。通过设置行人宽高比 ($w:h$) 特征阈值, 除去过窄、过小等不符合行人特征的目标前景块。

(4) 扩展行人目标前景块, 以确保行人不出现在 ROI 目标区域之外。矩形扩展长度为:

$$\text{Len} = u \times \min(w, h) \quad (3)$$

其中, u 是 0~1 的一个值, w 、 h 分别表示前景块的宽、高。

(5) 标记扩展后的行人目标前景块集合, 构建 ROI 区域以进行后续的特征提取和检测。ROI 目标区域提取如图 2 所示。



图2 ROI 目标区域提取

1.2 强分辨力行人特征提取

1.2.1 多尺度 HOG 特征

HOG 特征是通过计算局部区域的梯度方向直方图描述目标形状特征, 对光照变化和小量的偏移不敏感。DALAL N 提出的基本 HOG 是在 64×128 的检测窗内提

取 105 个 16×16 等大小的 Block 特征块, 利用这些特征块构成描述符来训练和检测行人。

等大小的基本 HOG 特征块对描述行人较大的部位 (如头部、身体等) 的整体信息能力有限, 因此本文构建 HOG 特征时采用了多尺度^[6]方法, 对于 64×128 检测窗, Block 块的大小选取依次从 16×16 到 64×128 , 宽高比 ($w:h$) 分别取 1:1、1:2 和 2:1, 共获得 452 个不同尺度的 Block 块。

在特征块提取过程中, 为避免特征块交叠而引起的重复计算, 引入“积分图^[6]”思想, 每个像素点的特征由 9 维向量表示, 即:

$$p(x, y) = \lambda G \quad (4)$$

其中, G 为像素点梯度幅度值; $\lambda = (0, \dots, j_k, \dots, 0)$ 表示像素点梯度方向, $j_k = 1$, k 为梯度方向落在的区间。HOG 特征简化为:

$$s(x, y) = s(x, y-1) + p(x, y) \quad (5)$$

$$h(x, y) = h(x-1, y) + s(x, y) \quad (6)$$

其中, $s(x, y)$ 表示点 $(0, 0)$ 到 (x, y) 直线上的像素向量累加, $h(x, y)$ 表示点 $(0, 0)$ 到 (x, y) 连接矩形内像素向量累加。对于矩形区域内的直方图, 统计可通过计算矩形 4 个顶点的值获得。矩形区域像素值可表示为:

$$N = h(x_1, y_1) + h(x_4, y_4) - h(x_2, y_2) - h(x_3, y_3) \quad (7)$$

其中, $h(x_1, y_1)$ 、 $h(x_2, y_2)$ 、 $h(x_3, y_3)$ 、 $h(x_4, y_4)$ 分别为矩形的 4 个顶点。

为了减少光照变化和阴影的影响, 在计算特征时采用了 L2-norm 归一化因子, 即:

$$x = \frac{x_i}{\sqrt{\sum_{i=1}^i x_i^2 + e^2}} \quad (8)$$

其中, x_i 为归一化前特征值, i 为 Block 块中特征数目, e 为略大于 0 的常数。

1.2.2 基于 Fisher 准则的行人强分辨力特征选取

由于高维 HOG 特征向量存在大量冗余信息, 不仅影响提取、分类速度, 某些维度向量还可能降低检测的精度。本文利用 Fisher 准则^[5]分类特性挑选分类贡献较大的强分辨力行人特征块。通过计算两类样本特征的类间离散矩阵 s_b 与总类内离散度矩阵 s_w 的比值 $F(j)$ 作为输出权值, 利用直线将行人特征、非行人特征区分开来, 将分类的过程转化为特征点在直线投影的过程, 根据行人 Block 块特征点距离直线远近来判断 Block 块的描述能力。Fisher 准则提取强分辨力特征的实现步骤如下。

(1) 遍历检测窗中每个 Block 块 (f_i), 根据式 (9)~式 (12), 利用 Fisher 准则计算每个块的输出权值 $F(j)$ 。

$$m_i = \frac{1}{N_i} \sum_{x \in m_i} x, i=1, 2 \quad (9)$$

其中, m_i 是两类样本均值的均值向量, N_i 是样本类的样本数目。

$$S_b = (m_1 - m_2)(m_1 - m_2)^T \quad (10)$$

$$S_w = \sum_{x \in \Omega_1} (X - m_1)^2 + \sum_{x \in \Omega_2} (X - m_2)^2 \quad (11)$$

$$F(j) = (S_b / S_w) \quad (12)$$

其中, S_b 、 S_w 分别为样品类间离散度矩阵值、总类内离散度矩阵值; $F(j)$ 为块输出权值, $F(j)$ 越大, Block 特征块区分能力越强。

(2) 依次对各个 Block 块的 $F(j)$ 从大到小降序排列。

(3) 从降序序列中挑选新特征集 f_{select} :

$$f_{select} = \sum_1^i f_i, i = i + 1 \quad (13)$$

逐次添加一个 f_i 直到 f_{select} 分类结果达到预期目标。利用 SVM 验证分类效果, 获得最优强分辨力行人特征 f'_{select} 。

(4) 对强分辨力特征 f'_{select} 进行 SVM 训练获得分类器, 用于后续 ROI 目标区域行人检测。

1.3 ROI 目标区域行人检测

通常 ROI 目标区域采用的检测窗大小固定, 而待检测视频图像中的目标位置、大小是随机变化的。为避免行人漏检, 本文采用多层次穷尽搜索获得检测窗。将 ROI 区域按照一定的步长逐步缩放, 直到达到预先设定的尺度。在不同尺度下分别进行行人检测, 并将所有结果融合处理、标记。行人检测结果如图 3 所示。



(a) 融合前检测结果 (b) 融合后检测结果

图 3 行人检测结果

2 实验及结果分析

本文实验在 MATLAB2008a 上实现, 计算机环境配置为 3 GHz CPU 和 2 GB 内存。采用的数据来源于 INRIA 行人数据库和 PETS2006 视频库。在提取强分辨力行人特征时, 分别将 INRIA 样本库中的 1 000 个正、负样本用于训练和检测, 并利用强分辨力 HOG 行人特征分类器对 PETS2006 视频库的视频行人进行检测。实验采用 libsvm 工具包, 选用线性 SVM 训练分类器。

文中通过对单尺度基本 HOG 扩展, 得到 Block 数为 452 的多尺度 HOG, 并利用 Fisher 准则分别对不同尺度 HOG 进行降维处理, 获得不同维数的强分辨力特征, 结合 SVM 分类器验证不同尺度、不同维度的特征检测效果。实验结果如表 1 所示。

表 1 不同 HOG 特征块数的检测对比

检测结果 (1 000 个样本)	单尺度下 block 数			多尺度下 block 数			
	105	30	10	452	105	30	10
SVM 分类准确率/%	92.5	90.5	86.1	95.2	92.6	90.8	87.7
提取特征时间/s	38.1	35.6	33.6	50.9	35.8	33.2	30.7
检测分类时间/s	9.78	2.5	1.04	23.1	9.6	2.2	1.01

实验表明, 扩展后的多尺度 HOG 特征, SVM 分类准确率显著提高; 相同特征维数下 (表中只列出 Block 数为 105、30、10 的特征), 多尺度特征比单尺度特征具有更高的检测准确率且降维幅度越大, 多尺度特征优势越明显。

文中采用设定最低检测率来提高检测速率, 利用 SVM 分类的效果来选择最佳分辨力特征。如表 1 所示, 在多尺度下, 当选取特征块数为 10 时, 保证了行人检测的准确率, 提取、检测的速率比基本 HOG 都有所提高, 检测速率提高了 9.7 倍。本文取多尺度下 Block 数为 10 的特征集作为 f'_{select} , 训练分类器进行行人检测和分类。

实验通过 ROI 目标区域提取, 大大降低了检测窗口数量, 并传统 HOG 全局扫描算法、Adaboost^[3,8] 算法作比较。由于本文算法 ROI 目标区域大小直接受行人数量、大小影响, 实验选取统一的单行人视频序列进行测试, 帧像素大小为 720×480。实验结果如表 2 所示。

表 2 几种算法性能比较

检测结果	算法		
	全局扫描+单尺度 HOG	ROI 提取+单尺度 HOG	ROI 提取+多尺度降维 HOG (本文算法)
训练 SVM 分类器	30min	30min	10min
检测窗数/帧	4 000	10	10
检测时间/(s/帧)	1 532	3.97	0.52
检测准确率/%	91.3	91.3	91.1

参考文献[3,8]提到的 Adaboost 算法是利用一系列弱分类器组合强分类器, 根据每个样本分类结果修改权值进行下一层分类训练, 训练分类器时依次对每个特征集单独训练, 耗费时间很长。本文算法采用 Fisher 准则提取的强分辨力行人特征训练分类器, 特征整体一次投影实现一次分类, 具有优势。

实验证明, 本文算法比全局扫描减少了检测窗数量, 显著提高了检测的速率; 在保证准确率的前提下改进单尺度 HOG, 利用多尺度降维 HOG 进一步提高行人的检测速率。本文算法不仅大幅度提高了行人检测的速率, 而且保证了视频行人检测准确率在 90% 以上。

本文针对传统 HOG 特征在行人检测中存在高维度计算的问题, 提出一种利用视频运动信息和强分辨力行人 HOG 特征相结合的视频行人检测算法。利用 INRIA 行人库和 PETS2006 视频库验证了算法的有效性。后续将针对夜间、雨天等复杂背景, 进一步验证和改进算法。

参考文献

- [1] 刘鑫, 刘辉, 强振平, 等. 混合高斯模型和帧间差分相融合的自适应背景模型[J]. 中国图象图形学报, 2008, 4: 729-734.
- [2] 王成亮, 周佳, 黄晟. 基于高斯混合模型与 PCA_HOG 的快速运动人体检测[J]. 计算机应用研究, 2012, 29(6):

- 2156-2160.
- [3] 周晨卉,王生进,丁晓青.基于局部特征级联分类器和模板匹配的行人检测 [J]. 中国图象图形学报, 2010, 15 (5): 824-829.
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR2005, 2005, 1: 886-893.
- [5] 孙响,刘富强,李志鹏.基于稳定区域梯度直方图的行人检测方法 [J]. 计算机辅助设计图形学报, 2012, 24 (3): 372-377.
- [6] 黄茜,顾杰锋,杨文亮.基于梯度向量直方图的行人检测[J]. 科学技术与工程, 2009 (13): 3646-3651.
- [7] 姚雪琴,李晓华,周激流.基于边缘对称行和 HOG 的行人检测方法 [J]. 人工智能及识别技术, 2012, 38 (5): 179-182.
- [8] 种衍文,匡湖林,李清泉.一种基于多特征和机器学习的分级行人检测方法 [J]. 自动化学报. 2012, 38 (3): 375-381.

(收稿日期: 2013-01-17)

作者简介:

张璐,男,1989年生,硕士,主要研究方向:视频分析与处理。

陈淑荣,女,1972年生,副教授,主要研究方向:视频分析与处理,现代通信网络及控制技术。

