

多特征融合的行人检测算法

康长青

(湖北文理学院数学与计算机科学学院,湖北 襄阳 441053)

摘要:针对站立行人的外观特性,提出多特征优选的行人检测算法。首先在特征提取阶段,分别利用梯度直方图,灰度共生矩阵、HSV颜色来提取行人的边缘、纹理和颜色频率特征,构成丰富的特征集。接着在分类器创建阶段,使用偏最小二乘法降维算法,优选出权重较大的特征,形成二次判别分类器。最后利用训练好的分类器,对行人进行监测。实验结果表明,本算法在 FPPW = 0.0001 时漏检率为 3% 左右,检测精度高于 HOG 和 PID 算法。

关键词:行人检测;梯度直方图;偏最小二乘法;漏检率

中图分类号:TP751.2 文献标识码:A DOI:10.3969/j.issn.1001-5078.2013.09.23

Pedestrian detection algorithm using multiple features fusing

KANG Chang-qing

(School of Mathematical & Computer Science, Hubei University of Arts and Science, Xiangyang 441053, China)

Abstract: A multiple features fusing algorithm is presented to detect pedestrian based on their appearance characteristics. Firstly, the edge, texture and color frequency features are extracted via HOG, Co-occurrence Matrices and HSV color in features-extracting stage, and a richer feature set is formed. Then, some important features are selected via partial least squares for dimension reduction and a quadratic discriminant analysis classifier is formed in classifier-creating stage. Lastly, the pedestrian is monitored by a well trained classifier. Experimental results show that the proposed algorithm outperforms both the HOG and PID algorithm, with a 3% miss rate at 10^{-4} FPPW.

Key words: pedestrian detection; HOG; PLS; miss rate

1 引言

行人检测是研究在视频中分割行人并精确定位的技术,在智能视频监控、智能交通等领域应用广泛。由于人体受姿态、外表、干扰等因素的影响,使得行人检测技术成为一个非常有挑战性的课题。已有的检测算法可以分为人体模型法^[1]、模板匹配法^[2-4]和特征分类法^[5-9]三类。第一类方法对行人外观特征和运动特征进行数学建模,缺点是模型求解比较复杂。第二类方法采用人为经验与实验验证的模板匹配方式,然而人形模板的多态性和复杂性很难构造出合适的模板。第三类方法通过行人统计特征提取,利用模式分类进行行人和非行人的分类,得到了广泛的认可和关注。典型的方法有 DALAL 提出的基于梯度直方图(HOG)的行人检测算法^[5],

表现出了优良的检测性能。后来学者们对 HOG 方法进行了各种算法改进^[6-10]。如文献[6]使用的综合密度,梯度和空间位置等特征信息的协方差矩阵法,文献[7]提出的融合颜色信息的局部二进制模式法和文献[8]提出的融合 edgelets, HOG 和协方差矩阵的级联法等。分析这些方法可知,它们通过增加不同的低级特征而取得了各自的分类效果,而对纹理这一重要分类特征却较少考虑。因此本文考虑行人纹理特征,在文献[5]的基础上提出融合边缘,颜色,纹理等丰富特征的行人检测算法。

基金项目:国家自然科学基金科研项目(No. 61272296)资助。

作者简介:康长青(1979-),男,博士,研究方向为图像处理和计算机视觉。E-mail:kangchangqing@163.com

收稿日期:2012-11-24

2 行人外观特征

行人的外观特征通过形成分块特征向量的方式进行。将检测窗口 d_i 分解为一些相互重叠的块,通过提取每一块的边缘,纹理和颜色频率特征来形成特征向量 v_i 。下面将分别介绍这三个类别特征的提取方法。

2.1 边缘特征的提取

边缘特征的提取采用文献[5]提出的 HOG 描述特征,HOG 特征是针对矩形区域的梯度方向与强度统计信息而定义的一种特征。其计算过程如下:

(1) 计算所有像素点的梯度 $G(x, y)$ 和方向 $Ori(x, y)$, 即:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (1)$$

$$Ori(x, y) = \arctan\left(\frac{I(x, y+1) - I(x, y-1)}{I(x+1, y) - I(x-1, y)}\right) \quad (2)$$

(2) 按空间位置将图像分成均匀的小块 (cell), 相邻的 cell 组成一个大块 (Block)。

(3) 在 Block 内进行归一化直方图来消除光照的影响, 形成 HOG 特征向量。

HOG 特征描述的是边缘的结构特征, 位置和方向空间的量化可以抑制平移和旋转带来的影响; 采取在局部区域归一化直方图, 可以部分抵消光照变化带来的影响。

2.2 纹理特征的提取

纹理特征的提取采用灰度共生矩阵法^[11]进行。灰度共生矩阵通过图像灰度级之间的二阶联合条件概率密度函数来表示纹理, 其中 $P(i, j | d, \theta)$ 表示在给定的距离 d 和方向 θ 上, 以灰度 i 为起点, 出现灰度 j 的概率, 表示如下:

$$P(i, j) = \frac{\#\{[(x_1, y_1), (x_2, y_2)] \in S | I(x_1, y_1) = i \& I(x_2, y_2) = j\}}{\#S} \quad (3)$$

其中, # 表示集合中元素数; 右边的分子项表示具有某种空间关系; 灰度级分别为 i, j 的像素对的个数; 而分母项为像素对的总个数。

利用共生矩阵可以计算出大量的纹理特征, 利用文献[11]提出的 14 个纹理特征, 选取其中的角二阶矩, 对比度, 相关性, 方差, 差分矩, 熵, 和平均, 和方差, 和熵, 差分方差, 差分熵, 方向性等 12 特征作为特征向量, 详细的公式参考文献[11]。

基于灰度共生矩阵的纹理特征可以利用衣服纹理的同质性和方向性, 来区分衣服纹理和背景纹理, 进而为行人和非行人分类提供帮助。

2.3 颜色特征的提取

由于行人的身体部位 (如头/脸) 可以提供的颜色特征信息有对放大、缩小、微小变形不敏感的优点, 有助于改善检测效果, 因此可以将这些信息用颜色频率来表示的特征。在 HSV 空间选取颜色空间中梯度最大的幅值来计算每个像素点的梯度值和方向, 用每个颜色空间被选择的次数来表示颜色特征。

计算方法为: 每个 Block 用一个 3 维的直方图分别统计 H、S、V 3 个颜色空间被选择的次数, 得到一个 3 维的颜色频率向量, 并进行归一化。这样每个 Block 只用一个 3 维的向量就可以表示颜色信息。

3 行人检测分类器的创建

3.1 偏最小二乘法分类器

利用偏最小二乘法^[12] (PLS) 的主成分分析功能进行分类器的创建。

PLS 模型可以看作由外部关系 (即独立的 X 块和 Y 块) 和内部块间关系构成。 X, Y 块的外部关系分别表示为:

$$X = TP + E = \sum t_k p_k + E \quad (4)$$

$$Y = UQ + F = \sum u_k q_k + F \quad (5)$$

其中, T, U 称为 X, Y 的成分矩阵; t_k, u_k 为相应的成分向量; P, Q 分别是 X, Y 载荷矩阵; p_k, q_k 相应的载荷向量。 E, F 是残差矩阵, 分别是 X, Y 中无法用个潜在变量 T, U 反应的部分。

PLS 模型的计算采用 NIPALS 算法, 通过对成分 t, u 和权重向量 w, c 进行迭代求解, 最大化以下公式:

$$\|\text{cov}(t, u)\|^2 = \max_{|w_i|=1} [\text{cov}(Xw_i, Yc)]^2 \quad (6)$$

其中, $\text{cov}(t, u)$ 为核心向量 t, u 的样本协方差。

通过 NIPALS 算法^[12] 提取一个成分后, 对矩阵 X, Y 进行退化, 并继续提取成分。如此往复, 直到提取的成分进行回归能达到一个较为满意的精度为止。 NIPALS 的算法描述如下:

(1) 随机初始化向量 u

(2) $w = X^T u$

(3) $t = Xw, t \leftarrow t / \|t\|$

(4) $c = Y^T t$

(5) $u = Yc, u \leftarrow u / \|u\|$

(6) 重复 (2) ~ (5), 直到收敛。

(7) 退化矩阵 $X, Y: X \leftarrow X - t t^T X, Y \leftarrow Y - t t^T Y$

用 NIPALS 算法对通过将检测窗口 d_i 中提取特征向量 v_i 进行预处理, 计算得到主成分数、各主成分数的权重和主成分间的关系矩阵。

利用计算出的各主成分权重,丢掉包含行人概率较低的检测窗口,对概率较高的检测窗口,使用全部特征进行分类。经过这样处理,不但可以减少特征计算的数目,还可以减少需要提取的特征块,进而提高行人检测的速度。

3.2 二次判别分析判别器

经过以上 PLS 特征优选后,利用二次判别分析建立判别器。二次判别分析的原理是在已知特征模式的类别数及其概率分布的前提下进行 Bayes 决策,达到将特征模式分配到后验概率最大的类别中。

假设各个类别的先验概率相同,条件概率密度采用正态分布,得到二次判别函数如下:

$$g_i(x) = \frac{1}{2} \log(|\Sigma_i|) - \frac{1}{2} (x - u_i)^T \Sigma_i^{-1} (x - u_i) \quad (7)$$

对所有的 $j \neq i$, 若 $g_i > g_j$, 将 x 归于 ω_i 。其中 u_i 为各类的均值向量, Σ_i 为各类的协方差矩阵。 u_i 和 Σ_i 的估计采用极大似然估计。设模式类别有 C 类, n_i 表示第 i 类训练样本数, x_{ij} 为第 i 类第 j 个训练样本, 则各类均值向量和各类协方差矩阵的极大似然估计表示为:

$$u_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \quad (8)$$

$$\Sigma_i = \frac{1}{n_i} \sum_{j=1}^{n_i} (x_{ij} - u_i)(x_{ij} - u_i)^T \quad (9)$$

4 行人检测算法流程

在前文特征提取和创建的分类器的基础上,建立多特征优选的行人检测系统,系统分为离线训练阶段和在线监测阶段。如图 1 所示。

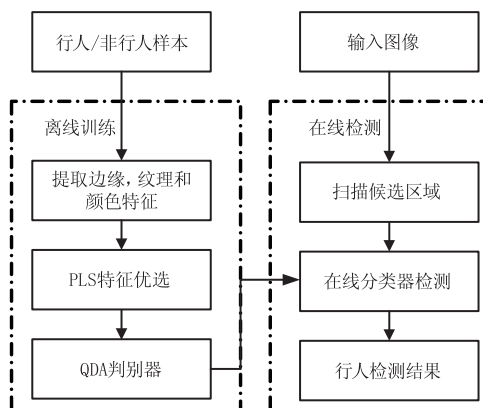


图 1 行人检测方块图

(1) 离线训练阶段。首先,收集人脸样本和非人脸样本,并利用上文提到的特征提取技术,提取出样本中行人的边缘、纹理和颜色特征,建立丰富的特征集,然后利用 PLS 降维算法从这些特征集中优选出权重较大的特征及权重系数,从而形成 QDA 分类

判别器。

(2) 在线监测阶段。对于新的输入图像,利用离线阶段训练好的分类器,扫描图像中行人的候选区域,对行人进行监测,得到检测结果。

5 实验结果和讨论

为了测试本文提出算法的性能,并与其他文献中的方法进行对比,在 INRIA 数据库上进行实验。

INRIA 数据库包含了训练集与测试集,训练集中含有 64×128 像素的正训练样本 2416 个,以及 1218 张不含人的背景图片,测试集中含有同样大小的正测试样本 1132 个,以及 453 张不含人的背景图片。利用本文的特征提取方法对 INRIA 数据库进行特征提取。

灰度共生矩阵的特征提取,采用 16×16 (步进 8 个像素) 和 32×32 (步进 16 个像素) 的块,在 HSV 颜色空间里,对每一个颜色通道, $d = 1, \theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ 创建 4 个分别对应 θ 灰度共生矩阵,并量化为 16 bins。从共生矩阵里提取前文提到的 12 个特征,产生 63648 个特征。HOG 特征和颜色频率特征提取,采用块的大小 12×12 到 64×128 的区间之间。对于本实验中的 2478 个块,对一个块提取 36 个 HOG 特征,得到 98928 个特征。同时对于每一块,利用颜色频率提取方法,提取的颜色特征 8274 个。对于大小为 64×128 的检测窗口,共得到 170820 个向量特征。为了避免特征向量太多,导致的检测效率降低。

可见无论是在特征向量的提取、分类器的训练与分类中,其计算量都很大。

在实验中记录正测试样本与负测试样本在分类器下的输出,并通过改变分类器阈值得到 DET (Detection Error Tradeoff) 曲线,即漏检率 (Miss Rate) 对误检率 (False Positives Per Window, FPPW) 曲线,如图 2 所示。

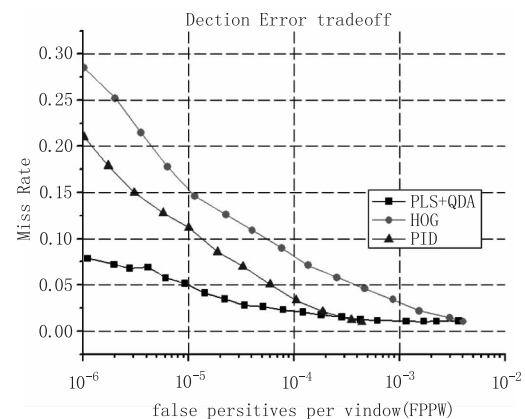


图 2 三种算法的漏检率/误检率

从图中可以看到,本文算法(PLS + QDA)性能优于文献[5]中的 HOG 和文献[13]中的 PID 算法,在 $FPPW = 0.0001$ 时本算法漏检率为 3% 左右,在 $FPPW \in [0.000001, 0.0001]$ 本算法的漏检率均小于 HOG 和 PID 算法。图 3 为算法的跟踪效果图。在算法的计算时间上,本文算法 1 秒钟可以处理 2929 个检测窗口。

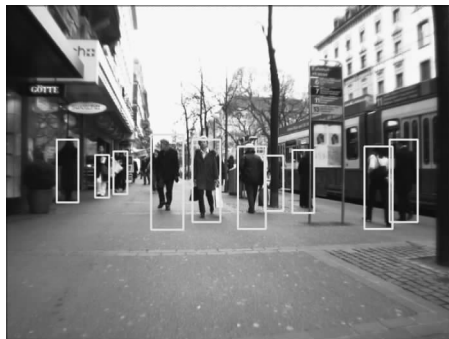


图 3 算法效果图

6 结论

针对站立行人的外观特性,提出融合边缘,纹理和颜色频率的多特征融合的行人检测算法。算法利用偏最小二乘法进行重要特征优选,形成二次判别分类器。实验结果表明,算法漏检率小,取得较好的检测精度;但算法也存在速度不够快的缺点,下一步研究将使用更好的降维方法来提高检测的速度。

参考文献:

- [1] Oren M, Papageorgiou C, Sinha P. Pedestrian detection using wavelet templates [C]// Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 1997, 193 - 199.
- [2] Viola P, Jones M, Snow D. Detecting pedestrians using patterns of motion and appearance [C]// Proc of International Conference on Computer Vision, 2003, 734 - 741.
- [3] Mikolajczyk K, Schmid C, Zisserman A. Human detection based on a probabilistic assembly of robust part detectors [C]// Proc of European Conference on Computer Vision, 2004, 69 - 82.
- [4] Su Xiaoqian, Sun Shaoyuan, Ge Man, et al. Pedestrian detection and tracking of vehicle infrared images [J]. Laser & Infrared, 2012, 42(8): 949 - 953. (in Chinese)
苏晓倩, 孙韶媛, 戈谯, 等. 车载红外图像的行人检测与跟踪技术 [J]. 激光与红外, 2012, 42(8): 949 - 953.
- [5] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]// Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 2005, 886 - 893.
- [6] Tuzel O, Porikli F, Meer P. Human detection via classification on riemannian manifolds [C]// Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 2007, 1 - 8.
- [7] Mu S, Yan Y, Liu T, et al. Discriminative local binary patterns for human detection in personal album [C]// Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 2008, 1 - 8.
- [8] Zhu Q, Yeh M C, Cheng K T, et al. Fast human detection using a cascade of histograms of oriented gradients [C]// Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 2006, 1491 - 1498.
- [9] Wu B, Nevatia R. Optimizing discrimination-efficiency tradeoff in integrating heterogeneous local features for object detection [C]// Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 2008, 1 - 8.
- [10] Chen Y T, Chen C S. Fast human detection using a novel boosted cascading structure with meta stages [J]. IEEE Trans. On Image Processing, 2008, 17(8): 1452 - 1464.
- [11] Haralick R, Shanmugam K, Dinstein I. Texture features for image classification [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1973, 3(6): 610 - 621.
- [12] Rosipal, N Kramer. Overview and recent advances in partial least squares [J]. Lecture Notes in Computer Science, 2006, 3940: 34 - 51.
- [13] Lin Z, Davis L S. A pose-invariant descriptor for human detection and segmentation [C]// Proc of European Conference on Computer Vision, 2008, 421 - 436.