

# 基于 Gentle Adaboost 的行人检测

朱文佳 戚飞虎

(上海交通大学计算机科学与工程系, 上海 200240)

**摘要** 行人检测是物体检测领域的一大难点。为了更加快速地检测行人,将人脸检测中 Boosted Cascade 算法应用到行人检测中,并对其进行了改进,即先利用加权平均矢量投影的方法将高维的梯度直方图特征化为1维,再利用 lookup table 进行概率密度估计,从而将 Gentle Adaboost 成功地应用于行人检测。实验表明,该方法不仅训练时间短、检测速度快,而且检测精度接近目前的最佳水平。

**关键词** 行人检测 Gentle Adaboost

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2007)10-1905-04

## Gentle Adaboost Based Pedestrian Detection

ZHU Wen-jia, QI Fei-hu

(Department of Computer Science and Engineering, Jiaotong University, Shanghai 200240)

**Abstract** Pedestrian detection is a difficult problem in the field of object detection. This paper applies the boosted cascade algorithm which originates from face detection to pedestrian detection. We improve the algorithm by first project the high dimensional hog (histogram of gradient) feature to 1 dimension, then estimate the posterior probability with lookup table so as to apply the gentle adaboost algorithm. Experiments indicate our algorithm is both efficient in training and detection. The detection rate is close to the state of the arts.

**Keywords** pedestrian detection, Gentle Adaboost

## 1 引言

行人检测是目前人们开展的一大研究热点。从实用角度来看,它在诸如智能车辆、自动监控、人机交互、虚拟现实等领域都具有广泛的应用前景;而从研究角度来看,由于人体无论是在身材、姿势、视角还是在衣着、光照方面都有着极大的类内变化,因而行人检测也有相当的研究价值。

目前的主流检测方法大致可以分为以下两类:

(1) 基于部位的方法<sup>[1]</sup>,首先检测人体的不同部位,如腿部、躯干、头肩等,再利用部位之间的几何关系即可得出检测结果。

(2) 基于扫描窗口的方法,通常是用一个分类

器扫描图片中的所有可能的子窗口。如 Dalal 提出了基于梯度分布的梯度直方图特征<sup>[2]</sup>,配合 SVM 分类器,达到了近年来行人检测领域最高的精度,然而该方法的缺点在于检测速度比较慢,即使是使用线性 SVM,每个子窗口的判断也需要几千次的计算量。Zhu 等人在检测速度上做了改进<sup>[3]</sup>,他们利用人脸检测领域的 Boosted Cascade 算法<sup>[4]</sup>,在基本不影响精度的情况下将检测速度提高了近 70 倍,在图片尺寸较小的情况下(320×240)可以达到实时。

然而文献[3]的方法也有以下局限性:一是训练时间过长,二是训练图片的尺寸较大(64×128)时,无法检测图片中较小的行人,三是对于较大的图片检测速度偏慢。

本文对文献[3]的方法做了进一步改进,利用

收稿日期:2007-05-29;改回日期:2007-07-05

第一作者简介:朱文佳(1983~),男,2005获上海交通大学学士学位。现为上海交通大学计算机科学与技术系在读硕士研究生。主要研究方向为计算机视觉与模式识别。E-mail: zhuwenjia@hotmail.com.cn

加权平均矢量投影算法配合 Gentle Adaboost, 将训练时间从几天降到 8 个小时; 同时减小了训练尺寸 (32 × 64), 并将梯度直方图特征做了一定简化, 使之能够以更快的速度检测更小的行人。而本文方法的检测精度则依然接近于当前的最佳水平。

## 2 方法概述

### 2.1 梯度直方图特征

如上所述, 行人的外观受身材、姿势、衣着、光照等多方面的影响, 如何才能有效地描述行人呢?

Dalal 提出的梯度直方图特征比较好地解决了这个问题<sup>[2]</sup>。具体来说, 就是首先把图片分成许多块, 再把每个块分成几个格, 然后在每个格中, 计算梯度的方向分布直方图作为特征向量; 最后把每个块的特征向量串联起来构成整个图片的特征向量。梯度直方图的优点在于它是基于梯度的方向分布直方图, 所以既可以描述人体的轮廓特征, 同时又对光照不敏感。梯度直方图的示意图如图 1 所示<sup>[5]</sup>。梯度直方图的实现涉及许多细节, 这里就不再赘述, 有兴趣的读者可以参考文献<sup>[5]</sup>。

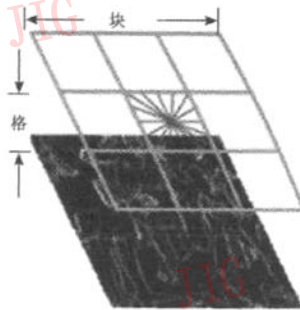


图 1 梯度直方图特征示意图  
Fig. 1 Illustration of HOG feature

值得一提的是, 由于本文的训练图片尺寸比文献<sup>[2]</sup>中的图片小了一倍, 如果将每个块分成许多格的话, 那么由于每个格将没有足够的像素来稳定地估计梯度的方向分布直方图, 所以在本文方法的实现中每个块对应一个格, 不再细分。

### 2.2 Boosted Cascade

基于扫描窗口的算法从应用角度看, 其面临的最大问题就是计算量。以一张 640 × 480 大小的图片为例, 所有可能的窗口将在 100 000 个以上。如何才能降低计算量呢? 直观上分析, 由于一张图片中

的绝大部分区域是不包含行人的, 而且相当比例的区域与行人的外观相去甚远, 因此用少量的计算量就能做到将其排除。只有那些与行人最相似的区域才值得花费较大的计算量。Viola 和 Jones 提出的 Boosted Cascade 算法<sup>[4]</sup>正是符合以上标准的经典算法。这个算法最初应用于人脸检测。在文献<sup>[3]</sup>中, 这个算法又被成功地应用于行人检测。概括地说, 该方法主要有以下几个特点:

(1) 级联结构的分类器。所谓级联分类器是指分类器分成很多级, 只有通过了前一级, 才能进入后面一级。这样一来, 许多简单的图片窗口可以在前几级被快速排除, 从而节约出时间用于对那些更像目标的区域进行检测。该级联的检测器的结构如图 2 所示。

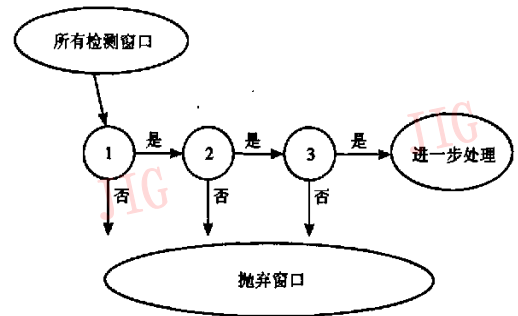


图 2 级联分类器的结构示意图  
Fig. 2 Illustration of cascade classifier

(2) 基于 Adaboost 的学习算法。Adaboost 算法最初是用于将一组弱分类器组合成一个强分类器。假如对每个特征定义一个弱分类器, 那么它就能从一个庞大的特征集中选择很小的一部分关键的特征, 从而使其成了一种非常有效的特征选择算法。在 Boosted Cascade 的算法中, Adaboost 算法被用来训练级联分类器中的每一级分类器。

Adaboost 算法的流程如下:

- (1) 给定训练样本和初始权重
- (2) For  $t = 1, \dots, T$ 
  - ① 归一化样本权重,
  - ② 对每一个特征, 训练一个弱分类器,
  - ③ 选择最优的弱分类器  $f_t$ ,
  - ④ 更新样本权重。

(3) 得到强分类器  $H = \sum_{t=1}^T f_t$

在文献<sup>[3]</sup>中, Zhu Qiang 先定义了各种大小和

尺寸的梯度直方图特征块,然后利用 Adaboost 算法选出最有效的块。这样既避免了文献[2]中所有块都是相同大小,特征不够丰富的局限性,又排除了冗余特征,从而提高了效率。

然而由于文献[3]中是使用线性 SVM 作为 Adaboost 算法中的弱分类器,从而大大影响了训练速度,为了提高训练速度,文献[3]仅仅从 5 031 个梯度直方图特征块中选择了 250 个来训练每一级分类器,这对特征来说无疑是一种浪费。

### 3 Gentle Adaboost

目前最常使用的 Adaboost 算法分别有 Discrete Adaboost, Real Adaboost 和 Gentle Adaboost<sup>[6]</sup>。它们的主要区别在于 Discrete Adaboost 要求弱分类器的输出是二值,后两者则将要求放宽到实值。文献[6]中的结果表明, Gentle Adaboost 通常可以取得更好的效果。

实现 Gentle Adaboost 的关键是要能估计出后验概率,即给定特征向量  $x$  的情况下,估计出在当前权值分布下的该样本为正样本的概率  $P_+(y=1|x)$ 。

在 Gentle Adaboost 中,弱分类器定义为

$$f(x) = P_+(y = 1 | x) - P_-(y = -1 | x) \quad (1)$$

在特征向量是 1 维的情况下,可以先利用 lookup table<sup>[7]</sup> 将特征的取值范围分为多个区间,然后通过统计每个区间中的权值总和来估计后验概率。具体地说,令  $W_i^{+1}$  和  $W_i^{-1}$  分别表示第  $i$  个区间  $B(i)$  内正负样本权值的总和。

$$W_i^{+1} = \sum_{x_j \in B(i) \& y_j = 1} w_j \quad (2)$$

$$W_i^{-1} = \sum_{x_j \in B(i) \& y_j = -1} w_j$$

那么后验概率为

$$P_+(y = 1 | x \in B(i)) = \frac{W_i^{+1}}{W_i^{+1} + W_i^{-1}} \quad (3)$$

$$P_-(y = -1 | x \in B(i)) = \frac{W_i^{-1}}{W_i^{+1} + W_i^{-1}}$$

将之代入式(1)即可得到弱分类器。

然而难点在于每个梯度直方图特征是高维的特征。对于高维特征,可以采取先投影到 1 维,然后再用 lookup table 估计后验概率。

接下来的问题就是如何确定投影方向? 本文采用了一种非常快速而有效的算法——加权平均矢量

投影算法。投影方向的计算如下:

$$d = \sum_{i=1}^n y_i w_i x_i \quad (4)$$

其中,  $y_i = 1$  表示第  $i$  个样本为正样本,  $y_i = -1$  表示负样本。  $w_i$  表示第  $i$  个样本的权重。  $x_i$  表示第  $i$  个样本的特征向量。

如果对上式稍作变形,即

$$d = \sum_{y_i=1} w_i x_i - \sum_{y_i=-1} w_i x_i \quad (5)$$

则  $d$  也可以看成是正样本加权中心到负样本加权中心的连线。

应该说,这样的分类方法是非常简单的,从单个弱分类器来看,其效果肯定不如线性 SVM。然而由于它在计算复杂度上,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。在文献[2]中,由于线性 SVM 非常耗时,每一级训练时只能使用 250 个特征;而本文方法训练时,每一级可以用到 3 000 个特征。另一方面,由于各种 Boosting 算法本身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。

### 4 实验结果

为验证本文方法的有效性,本文采用 Inria 行人库<sup>[2]</sup>进行了实验。该行人库分为测试集与训练集。训练集中含有 2 416 张行人图片和 1 218 张不含行人的背景图片;而测试集中则含有 1 127 张行人图片和 454 张背景图片。该行人库中行人的外观各异,是相当复杂的行人库。该库中的一些实例如图 3 所示。

训练前,先将所有正样本缩放到  $32 \times 64$  pixels,然后随机选择 3 000 个不同大小尺寸的梯度直方图特征进行训练。级联分类器每一级的目标检测率为 99.5%,目标误报率为 50%,共训练 25 级。



图 3 Inria 行人库中的一些样本  
Fig. 3 Some examples from Inria data set

整个训练过程在一台 Pentium 4, 2GHz CPU 以及 2G 内存的 PC 上耗时约 8 小时, 大大短于文献 [3] 中的几天。对于 12 800 个测试窗口, 文献 [3] 的方法需要 250ms, 本文的方法仅需要 200ms。

最终在 Inria 测试集上的测试效果 (ROC (receiver operating characteristic) 曲线) 如图 4 所示。

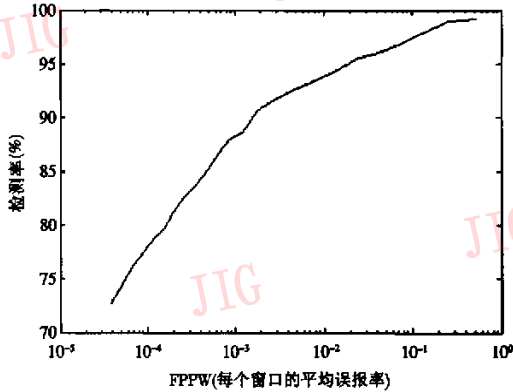


图 4 在 Inria 测试集上的 ROC 曲线

Fig. 4 ROC curve on Inria test set

需要一提的是, 本文的检测率在误报率为  $1/10\ 000$  时为 78% 左右, 这比文献 [2]、[3] 的结果都要明显地低一些。这是因为本文实验用的训练图片大小仅为  $32 \times 64$ , 是他们训练用的图片的一半。在文献 [2] 中, 训练图片尺寸为  $64 \times 128$  时, 检测率为 88%, 但是当训练尺寸缩小为  $48 \times 112$  时, 检测率就下降到了 82%。由此可见, 考虑到训练尺寸因素, 本文算法的精度还是比较令人满意的。

## 5 结论

本文将 Boosted Cascade 算法框架应用于行人

检测中, 提出了加权平均矢量投影的方法, 由于将高维的梯度直方图特征化为 1 维, 简化了后验概率的估计, 从而将 Gentle Adaboost 成功地应用于行人检测中。在 Inria 行人库上取得了令人满意的检测效果。

但本文的工作仍有不足。未来的工作包括进一步研究训练尺度和精度的关系, 比较各种 Boosting 算法的优劣, 以及将其他算法 (如 logistic 回归), 融合到 Gentle Adaboost 中。

## 参考文献 (References)

- 1 Leibe B, Seemann E, Schiele B. Pedestrian detection in crowded scenes [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], San Diego, CA, USA, 2005, 1: 878 ~ 885.
- 2 Dalal Navneet, Triggs Bill. Histograms of oriented gradients for human detection [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], San Diego, CA, USA, 2005, 1: 886 ~ 893.
- 3 Zhu Qiang, Avidan Shai, Yeh Mei-Chen. Fast human detection using a cascade of histograms of oriented gradients [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], New York, 2006, 2: 1491 ~ 1498.
- 4 Viola Paul, Michael Jones. Robust real-time face detection [J]. International Journal of Computer Vision, 2004, 57(2): 137 ~ 154.
- 5 Dalal Navneet. Finding People in Images and Videos [D]. Phd Thesis. Inria Rhone-Alpes. 2006 [EB/OL]. <http://lear.inrialpes.fr/pubs/2006/Dal06>.
- 6 Friedman Jerome, Hastie Trevor, Tibshirani Robert. Additive logistic regression: a statistical view of boosting [J]. The Annals of Statistics, 2000, 28(2): 337 ~ 407.
- 7 Wu Bo, Ai Hai-zhou, Huang Chang. Fast rotation invariant multi-view face detection based on real adaboost [A]. In: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition [C], Seoul, Korea, 2004: 79 ~ 84.