

# 基于局部特征级联分类器和模板匹配的行人检测

周晨卉 王生进 丁晓青

(清华大学电子工程系, 北京 100084)

**摘要** 在智能视频监控领域,行人检测正受到广泛关注。为了提高检测率,将基于局部特征的 Adaboost 级联分类器和模板匹配相结合。首先通过分析和比较提出了一种行人局部特征的选取方法,然后对人体局部进行建模,将模板匹配的思想融入行人检测。最后通过实验证明,该方法在行人检测上可以取得较好的效果。

**关键词** 行人检测 局部特征 Adaboost 模板匹配

中图分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2010)05-824-06

## Pedestrian Detection Based on Partial Feature and Model Matching

ZHOU Chenhui, WANG Shengjin, DING Xiaoping

(Department of Electronic Engineering, Tsinghua University, Beijing 100084)

**Abstract** In the field of intelligence video surveillance, pedestrian detection is paid great attention to the security of daily life. In order to raise the detection rate, we combine partial features based on Adaboost with model matching. First, a new way of selecting proper partial feature is shown through analysis and comparison. Then, partial body is used in modeling which shows that model matching is combined in pedestrian detection. At last, experiments indicate that our method achieves a good result in pedestrian detection.

**Keywords** pedestrian detection, partial feature, Adaboost, model matching

## 0 引言

在当今社会中,智能化的监控正受到广泛关注。其中,行人检测由于需求的迫切性成为了一个研究的热点。在目标检测中,常用的一类方法是基于统计学习的分类器设计算法。这类方法能够对任意图像中的目标进行检测,而不要求了解相关的背景知识,这使得检测过程的适用范围更广。在机器学习的方法中,Adaboost 能够将大量容易得到的弱分类器组合起来构成强分类器,并且不需要任何关于弱分类器性能的知识,其在理论上能够达到任意好的效果<sup>[1]</sup>。

Adaboost 在人脸检测方面<sup>[2]</sup>的成功应用使它在行人检测方面被推广。但是由于人体的非刚性,

使得 Adaboost 在提取有效特征时的鲁棒性受到影响。在人脸检测中,五官相对固定的位置使得提取的 Haar 特征有很好的普适性;而在行人检测中,人的姿势和轮廓的改变可能会使训练得到的特征失效,因此在行人检测上简单地使用 Adaboost 很难达到其在人脸检测上取得的优异效果。针对 Adaboost 对行人整体进行检测的缺点,通过特征分析和比较选取恰当的局部来训练分类器,并且运用模板的方法辅助检测,使得在相同的虚警率下检测率有了明显提高。

## 1 Adaboost 算法简介

在分类器设计的领域, Freund 和 Schapire 在 1995 年提出了 Adaboost 算法<sup>[3]</sup>。Adaboost 算法通

基金项目:国家高科技研究发展计划(863)项目(2006AA01Z115);国家重点基础研究发展计划(973)项目(2007CB311004)

收稿日期:2009-05-01;改回日期:2009-08-17

第一作者简介:周晨卉(1986—),男,清华大学电子工程系硕士研究生。主要研究方向为计算机视觉和模式识别。E-mail: zch04@mails.tsinghua.edu.cn

过在训练数据上的学习得到弱分类器,通过每次更新权值,将不同的弱分类器融合起来,形成强分类器来做决策。如果样本数足够大,强分类器最后判断的错误率将趋近于零。但是在实际中由于过拟合<sup>[4]</sup>等问题,最后的效果会差于理论值。

常用的改进型 Adaboost 有 discrete Adaboost, real Adaboost 和 gentle Adaboost<sup>[5]</sup>。其中 real Adaboost 算法的过程如下<sup>[2]</sup>:

- 输入: $m$  个样本  $(x_1, y_1), \dots, (x_m, y_m)$ , 其中  $y_i \in Y = \{1, 2\}$ , 训练轮数为  $T$ ;
- 初始化样本权重  $w_1(i, y) = 1/|B|, (i, y) \in B$ ;
- $Fort = 1, \dots, T$ ;

1) 对每个弱分类器  $h$ :

(1) 把样本集  $X$  分成几个互不相交的块  $X_1, \dots, X_n$ ;

(2) 在权重  $w_i$  下,计算:  $W_l^i = P(x_i \in X_j, y_i = l)$ , 其中  $l = \pm 1$ ;

(3) 对每个块  $X_j$  设置弱分类器  $h$  的输出:

$$\forall x \in X_j, h(x) = \frac{1}{2} \ln \left( \frac{W_{+1}^j + \varepsilon}{W_{-1}^j + \varepsilon} \right)$$

其中,  $\varepsilon$  是一个小的大于零的常数;

(4) 计算归一化因子:  $Z = 2 \sum_j \sqrt{W_{+1}^j W_{-1}^j}$ ;

2) 选择拥有最小的  $Z$  的  $h$  作为最佳弱分类器;

3) 更新样本权重  $w_{i+1}(i) = w_i(i) \exp[-y_i \times h_i(x_i)]$ , 并将其归一化为概率密度函数。

- 得到强分类器:  $C(x) = \text{sign} \left[ \sum_{i=1}^T h_i(x) - b \right]$ ,

$b$  的默认值为 0。

其中块的划分可以根据 Haar 特征的值来进行,即将特征值归一化到  $[0, 1]$ ,再将区域均分成  $n$  个子区域:  $\text{bin}_j = [(j-1)/n, j/n], j = 1, \dots, n$ 。

在实际应用中,常常用级联的 Adaboost 来减少样本训练的时间,对每一级要求有很高的检测率,而对虚警率的要求则可以适当放宽。

## 2 组合特征检测

### 2.1 特征选择

在进行行人检测时样本常常是针对整个人的,但由于人姿态的变化特别是腿部动作的复杂性,如果直接用直立的人作为样本则往往会无法检测出行走的人,因此用全身来做样本是不合适的。为了找到鲁棒的特征,应该对人的局部进行分类器训练。

文中所选的训练样本来自 MIT 的行人数据库。将人体分为以下几个部分:全身、肩头、上半身、三分之二全身、躯干和腿部。对这些部分分别用基于 Haar 特征的 Adaboost 算法来做训练,并对训练后的结果取出第 1 个 Haar 特征。

从上面的结果可以看出,各个局部最主要的 Haar 特征是人体与周围背景的不同,即人的轮廓和周围背景存在较明显的梯度,这和直观的感觉是一致的。而人的竖直轮廓线是最长的,因此对一个好的样本提取的 Haar 特征应以水平特征居多(图 1 中的特征都是水平特征),如果存在过多的竖直特征或中心特征,则说明由于样本的特殊性和类内方差较大等原因导致提取的 Haar 特征偏离了人原有的特点,此时的检测效果会差于预期效果。

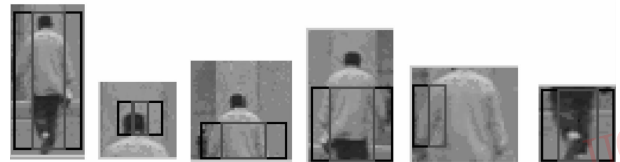


图 1 局部特征

Fig. 1 Partial feature

对上述样本采用级联方式进行训练后得到了特征分布的统计数据。在表 1 中,全身、三分之二全身和躯干中水平特征的比例较大,这是由于在这 3 部分的样本中都包括了人体最长的竖直边缘——躯干。对于肩头和上半身来说,头部在整个样本中占有很大的比例,而头部的轮廓是圆形的,因此水平特征和垂直特征都会大量出现,从数据上看几乎各占一半。虽然理论上正立人的腿部水平 Haar 特征会较多,但衣服的遮挡、腿部的各种动作等会使腿部的轮廓位置发生偏移,这种特点使得水平特征不易被准确提取,因此占的比例不多(48.60%);相反,竖直特征占有很大的比例,这说明训练出来的分类器中

表 1 局部特征的比较

Tab. 1 Comparison of partial feature

	水平特征	其他特征	水平特征比例/%
全身	155	65	70.45
肩头	257	249	50.79
上半身	215	205	51.19
三分之二全身	191	85	69.20
躯干	269	149	64.35
腿部	208	220	48.60

有很多无效的和不对应的特征,因此人体的腿部姿势不利于用 Adaboost 训练。虽然躯干的水平特征占有的比例较高(64.35%),但从图 1 中躯干部分的第 1 个 Haar 特征注意到,该特征受到手臂摆放位置的影响非常大,因此它对于非正立的样本会失效。基于躯干部分的训练还有一个缺点:将柱状物从虚警中排除的能力较差。经过实验发现无论是柱子还是正面或背面的腿部区域都容易被错认成是躯干。对于肩头而言,由于样本尺度相对较小,因此要达到相同的性能必然需要更多的弱分类器(从表 1 中可以发现在训练同级数时它用的特征个数是最多的),这样就可能出现过拟合的情况,导致之后几级的分类器性能退化,因此会存在较多的虚警。虽然基于肩头和躯干的分类器存在相似的问题——虚警高,但是肩头部位有相对于躯干的优势:姿态变化少。躯干部位存在受到手臂动作影响大的问题,而肩头的形状相对比较稳定(在没有装饰物的情况下,一般而言头部近似于圆)。因此,虽然肩头的检测会有许多虚警,但是由于其区域中目标不变性较好,因此可以用较少级数的肩头检测器取得高的虚警率和较高的检测率,然后再进一步作检测。

经过比较,可以看出全身、上半身和三分之二全身的特征是较好的。为了进一步地加以区分,将它们在测试集上进行检测。取 500 张正样本,取 100 张不存在行人的大小为  $640 \times 480$  的图作为负样本。图 2 是实验结果,在虚警率较低( $< 0.5 \times 10^{-5}$ )的情况下,全身样本的表现最好,然而此时的检测率低于 60%,没有实际应用的意义;而在虚警率略高的情况下,三分之二全身样本的表现最好,这一段的检测

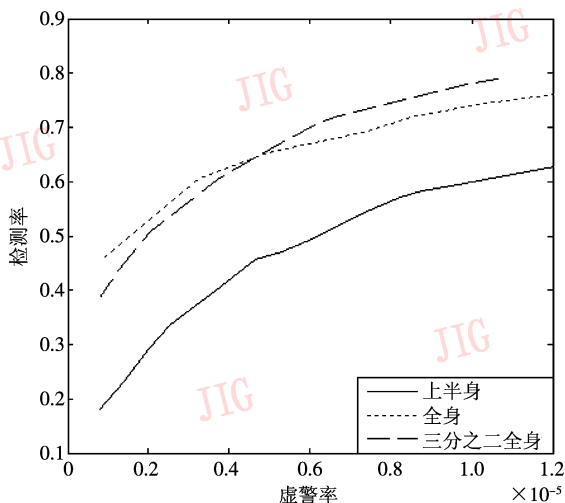


图 2 局部特征的 ROC 曲线

Fig. 2 ROC curve of partial feature

率较高。比较以上 3 条曲线,可以看出三分之二全身的效果最优,这是因为与全身样本相比,三分之二全身样本受到腿部姿势的影响较小;与上半身样本相比,由于人的腰部至臀部处的活动关节少,轮廓变化相对较小,因此三分之二全身样本有更多的有效特征可以提取。综上所述,三分之二全身样本是用于行人检测最好的局部特征。

在行人检测中,如果直接采用上述方法得到的分类器进行检测,则虚警率仍然偏高。根据级联 Adaboost 检测率和虚警率的关系可知,高的检测率和低的虚警率是互相矛盾的,不可能同时达到。因此,需要采取其他辅助的判断方法,在检测率保持基本不变的情况下降低虚警率。

## 2.2 基于模板匹配的头部检测

在行人这个被检测对象中,头部的轮廓有很好的不变性。因为无论从哪个角度看,头部的轮廓都接近于圆形,而行人的其他部分则没有这种特点。在这种情况下,模板匹配的方法能较好地描述目标,因此可以利用它来排除虚警。

Hausdorff 距离<sup>[6]</sup>是一种适合度量模板匹配优劣的方法,被广泛应用于基于边缘特征的匹配。Hausdorff 距离是针对两个有限闭集间距离度量的数学定义。假设空间中存在 2 个点集:

$$A = \{a_1, \dots, a_n\}, B = \{b_1, \dots, b_n\}$$

则集合  $A, B$  之间的 Hausdorff 距离定义为

$$H(A, B) = \max(h(A, B), h(B, A))$$

$H(A, B)$  也称为集合  $A$  和集合  $B$  的双向 Hausdorff 距离。

$h(A, B)$  和  $h(B, A)$  分别被称为集合  $A$  到集合  $B$ , 集合  $B$  到集合  $A$  的单向 Hausdorff 距离,定义为

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

$$h(B, A) = \max_{b \in B} \min_{a \in A} \|a - b\|$$

其中,距离的定义可以是  $L_1$  (和范数),  $L_2$  (欧氏范数) 或  $L_\infty$  (极大范数) 等。当  $a \in A$  为集合  $A$  所有的点中离集合  $B$  任一点最远的点时,就以  $a$  和其在  $B$  中的最近邻点间的距离作为  $A$  到  $B$  的单向 Hausdorff 距离  $h(A, B)$ 。而  $A$  与  $B$  间的 Hausdorff 距离则是取  $h(A, B)$  和  $h(B, A)$  中的较大值,它度量了 2 个点集之间的最大不匹配度。为了提高匹配的抗噪性能,采用了改进型 Hausdorff 距离 (MHD)<sup>[7]</sup>。

首先对待测图像进行轮廓提取,并将待匹配图像和模板二值化,然后将建立好的模板在匹配图像

中遍历,找出 Hausdorff 距离最小的位置,认为它是头部的候选区域,如图 3 所示。其中所使用的 Hausdorff 距离是单向的,这是由于匹配的目的是找到和模板相似的轮廓,而反过来的方向却是不必要的。这样做可以避免提取边缘后待匹配图像中噪声的干扰,同时匹配的速度提高了一倍,效果却和原来一样,能正确框住头部。

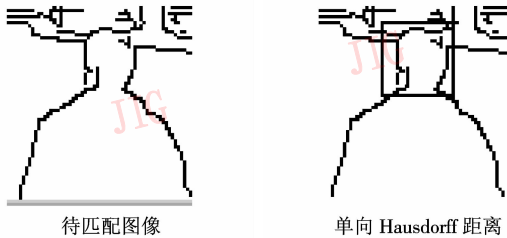


图 3 模板匹配的结果  
Fig. 3 Result of model matching

由于头的大小在待测图像中占的比例不同,所以应该用不同大小的模板去匹配。此时距离要对模板的大小进行归一化,然后选取各种不同大小模板中 Hausdorff 距离最小的作为最优框的大小和位置。如果待匹配的图像中没有头部,则求得的 Hausdorff 距离通常会远大于在图像中有头部的 Hausdorff 距离。对一组通过 Adaboost 方法检测后认为存在目标的图像进行基于头部的模板匹配,图 4 中的横坐标是 Hausdorff 距离,可以看出,虚警的 Hausdorff 距离的分布区域(红色的圆)与头部的 Hausdorff 距离的分布区域(蓝色的圆)有较大的差异:头部的 Hausdorff 距离较小,主要集中在原点附近;而虚警的 Hausdorff 距离通常要大得多,且分布较分散。经过实验测试,在相同检测率的情况下,使用模板匹配可以使得虚警率下降 20%~30%;在相同虚警率的情况下,使用模板匹配可以使检测率提高 5%左右。

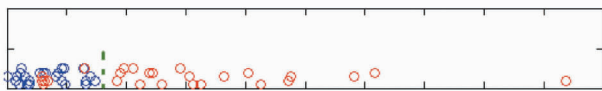


图 4 模板匹配中的 Hausdorff 距离  
Fig. 4 Hausdorff distance in model matching

因为人的头部轮廓近似于圆,因而头部的模板取圆形。然而由于存在人的下颌与颈部之间可能有衣物遮挡导致脸部轮廓变形、没有衣物遮挡时下巴和颈部之间轮廓不明显等因素,用一个完整的圆形

做模板是不太适合的。为了获得适当的模板,对圆形轮廓做一些改进:定义张角  $\theta$ ,如图 5(a)所示,用有缺口的圆作为模板。用不同大小的张角对一组正负样本进行测试,定义:

模板样本区分度 = 负样本的平均 Hausdorff 距离 - 正样本的平均 Hausdorff 距离

从图 5(b)的曲线上可以看出,在张角接近  $180^\circ$ ,即模板取半圆时,正负样本的区分度是最大的。虽然随着张角的增大样本的类内方差有所增加,但是样本的区分度(可以看作是一种分辨率)起了主导作用,因此实际检测的结果仍然优于取整个圆作为模板时的结果(在相同检测率下,取半圆作为模板时能够排除的虚警个数约为取整个圆时的一倍)。

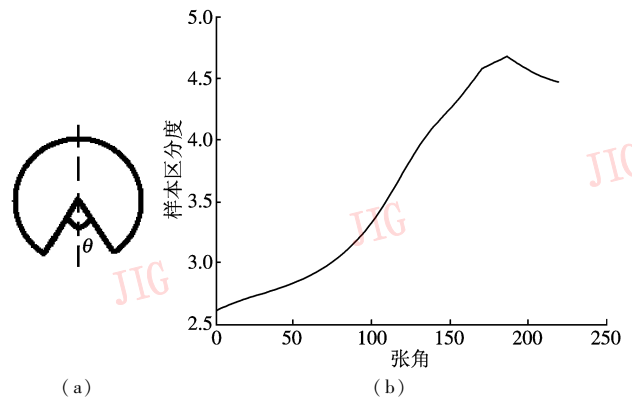


图 5 模板的选择  
Fig. 5 Choice of model

### 3 实验结果

对于行人检测的过程如下:首先在整幅图像中用基于肩头部位 Haar 特征的 Adaboost 算法提取可能存在目标的 ROI,然后在 ROI 周围用基于三分之二全身的 Adaboost 方法确认目标,最后将候选目标用模板的方法匹配头部轮廓进行虚警排除。前两步的 Adaboost 分类器的级数可以较少,它的效果要优于只有一步的基于三分之二全身的 Adaboost 分类器。这是因为只有一个 Adaboost 分类器时,随着级数的增加后面的弱分类器性能变得不可靠,而分为两次检测后第 2 步中基于不同区域重新训练的弱分类器性能必定好于训练了许多级后的弱分类器性能,它的样本区分度更好,且不易产生由于训练级数太多而导致过拟合的情况。另一种方法是首先用三分之二全身的 Adaboost 算法提取可能存在目标的

ROI,然后在 ROI 的上半区域用基于肩部部位 Haar 特征的 Adaboost 方法确认目标,最后将候选目标用模板的方法匹配头部轮廓进行虚警排除。但这种方法中模板匹配的效果并不明显,经过实验测试,在相同检测率的情况下这种组合方法的虚警率比前一种要高出 20%以上。因此,采用如图 6 所示的组合方法进行行人检测。



图 6 检测过程

Fig. 6 Process of detection

为了证明本文方法的有效性,进行如下实验:训练的正样本来自于 MIT 的行人数据库,左右镜像后共 1 848 张,样本大小取为  $32 \times 64$ 。在 Inria 数据库上进行实验,分别得到使用本文方法和 gentle Adaboost<sup>[8]</sup> 的 ROC 曲线(如图 7 所示)。

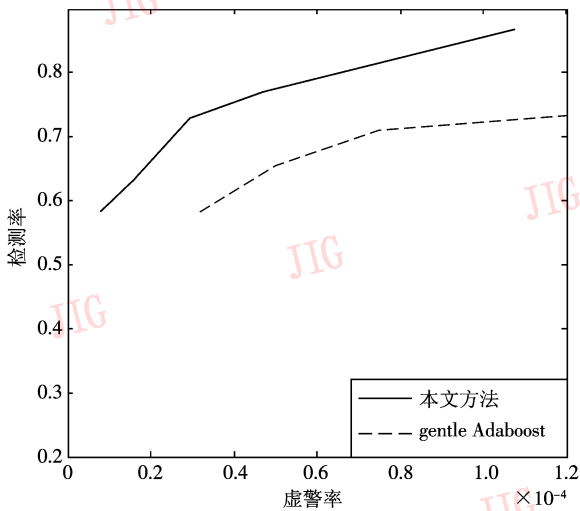


图 7 新老方法的比较

Fig. 7 Comparison of different approaches

从图 7 可以看出:在虚警率相同的情况下,采用本文方法的检测率要比 gentle Adaboost 方法的检测率高 15%左右;在检测率相同的情况下,本文方法的虚警个数不到用普通 Adaboost 方法产生的虚警个数的一半。上述实验结果说明,提出的基于 Adaboost 的组合特征的检测方法比传统的 Adaboost 方法有更好的性能,同时也优于基于广义类 Haar 小波的 Adaboost<sup>[9]</sup> 等方法。图 8 是一些相关的实验结果,可以看出本文对于下半身被部分遮挡的目标检测有较好的鲁棒性。



图 8 检测结果

Fig. 8 Detection results

### 4 结 论

传统的 Adaboost 行人检测只是对整个样本进行训练,并没有关注人体各个部分特征重要性的不同。本文首先分析了人体局部特征的不变性,并通过实验获得重要的局部特征,然后结合模板匹配的方法排除虚警,形成新的检测方法,最后通过实验证明了这种方法的有效性。但是仍有不足之处,由于模板匹配部分的耗时较多,检测一幅  $480 \times 640$  大小的图像需要数秒。未来的工作包括进一步研究模板匹配和 Adaboost 之间的联系,探索更优的特征检测方法。

### 参考文献 (References)

[1] Viola P, Michael J. Robust real-time face detection [J]. International Journal of Computer Vision, 2004, 57 (2): 137-154.

[2] Huang C, Wu B, Ai H Z, et al. Omni-directional face detection based on real adaboost [C]//Proceedings of 2004 International Conference on Image Processing. Washington DC, USA: IEEE Computer Society, 2004: 593-596.

[3] Freund Y, Schapire R E. A decision-theoretic generalization of

- on-line[J]. Learning and an Application to Boosting. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
- [4] Mitchell T M. Machine Learning [M]. Translated by Zeng Huajun, et al. Beijing: China Machine Press, 2003: 53-70. [Tom M. 机器学习[M]. 曾华军等译. 北京: 机械工业出版社, 2003: 53-70.]
- [5] Jerome F, Trevor H, Robert T. Additive logistic regression: A statistical view of boosting[J]. The Annals of Statistics, 2000, 28(2): 337-407.
- [6] Daniel P H, Gregory A K, William J R. Comparing images using the Hausdorff distance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(9): 850-863.
- [7] Dubuisson M P, Jain A K. A modified Hausdorff distance for object matching [C]//Proceedings of the 12th International Conference on Pattern Recognition. Washington DC, USA: IEEE Computer Society, 1994: 566-568.
- [8] Zhu Wenjia, Qi Feihu. Gentle Adaboost based pedestrian detection[J]. Journal of Image and Graphics, 2007, 12(10): 1905-1908. [朱文佳, 戚飞虎. 基于 Gentle Adaboost 的行人检测[J]. 中国图象图形学报, 2007, 12(10): 1905-1908.]
- [9] Navneet D, Bill T. Histograms of oriented gradients for human detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC, USA: IEEE Computer Society, 2005: 886-893.