

# 基于 Bandelet 变换与人体部位的人体检测方法

韩 红, 伍 星

(智能感知与图像理解教育部重点实验室, 陕西西安 710071)

**摘 要:** 本文由图像的几何流特性出发, 提出了一种基于第二代 Bandelet 变换与部位的人体检测方法. 首先, 利用优化后 Bandelet 变换的 Bandelet 系数及其统计特征作为人体图像的特征, 通过相关试验确立了相关的最优参数和统计特征. 然后再利用 AdaBoost 算法训练人体及各部位分类器. 最后通过计算人体各部位的似然度, 联合部位结合策略进行人体检测. 试验结果表明, 本文提出的特征提取方法能够更好地表征人体, 并能有效地改善分类器性能, 相应的部位检测方法可显著提高静态图像中人体目标检测的鲁棒性.

**关键词:** 人体检测; 特征提取; Bandelet; AdaBoost; 人体部位

**中图分类号:** TP391      **文献标识码:** A      **文章编号:** 0372-2112 (2011) 08-1785-08

## Human Detection Based on Bandelet Transform and Body Parts

HAN Hong, WU Xing

(Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xi'an, Shaanxi 710071, China)

**Abstract:** A new method is proposed in this paper for feature extraction based on geometric flow of images and the second generation of Bandelet transformation, where Bandelet coefficients and their statistical values were extracted as the feature of human images. Afterwards the full body and body parts classifier were trained on AdaBoost algorithm. At last, likelihoods of each body parts were computed combined with Bayesian decision-based approach to perform human detection. The results of human detection experiments indicate our proposed feature extraction method's better capability in describing human characteristics while effectively improving the performance of classifier. Combined with body parts detection, our proposed human detection method well enhanced the robustness of human detection task in both static and moving images.

**Key words:** human detection; feature extraction; Bandelet; AdaBoost; body parts

## 1 引言

人体检测在计算机视觉中有许多重要的应用, 如视频监控、智能汽车及智能交通、机器人和高级人机交互等. 然而, 由于人体自身姿态的变化、衣服多样性和光照等因素的影响, 人体外观变化非常大, 导致人体检测成为一个非常困难的问题.

目前的人体检测方法主要有基于人体模型的方法、基于模板匹配的方法和基于统计分类的方法. 基于统计分类的方法的比较鲁棒, 但需要很多训练数据, 而且很难解决姿态和遮挡的问题, 主要包括两个步骤: 特征提取和分类器设计. 其中所选特征的表征能力将直接影响分类器的性能. 目前所选的特征包括: 原始灰度特征空间、Haar-like 小波特征、形状描述子特征、Gabor 特征、有向梯度直方图 (HOG) 特征<sup>[1]</sup>, SIFT 特征等. 而根据分类器的设计方法, 现有的基于统计分类器可分为基于神经

网络 (NN) 的方法, 基于支持向量机 (SVM) 的方法和基于 Adaboost 的方法. 其中 Dalal 和 Triggs 在 2005 年提出了 HOG 特征提取方法, 根据图像像素的梯度方向来统计其梯度大小, 具有算法复杂度低, 同时分类正确率较高的优点, 在近几年的人体检测研究中, 得到了广泛的应用. 如何进一步提高检测的正确率, 是人体检测研究领域的核心问题之一. 传统的基于边缘的方法用于人体检测的问题在于对于几何特性的表述不严格, 很难刻画图像. 为此, 法国学者 Pennec 和 Mallat<sup>[2]</sup> 等引入了几何流来刻画图像的几何性质, 在 2004 年提出了第二代 Bandelet 变换, 避免了第一代 Bandelet 变换的重采样和弯曲等繁杂操作. 目前, 由于其能够更好的表示图像, 基于第二代 Bandelet 变换的研究主要集中于图像的压缩和降噪中. 条带波的核心思想是把图像中的几何特征定义为矢量场, 而不是简单地看成普通边缘的几何. 条带波对图像处理而言, 有良好的多尺度特性, 可以保证连续地

细化图像,具有时频局部化特性和多方向性、各向异性,以及可根据图像自身的几何结构特性,自适应获取所需的最优基函数的自适应性等.本文即从图像的几何流特性出发,提出了一种基于第二代 Bandelet 变换的人体检测特征提取新方法.

另一方面,无论使用何种特征,基于整个人体的检测方法,都不能很好地处理遮挡问题,而基于人体部位的检测方法则可以用来处理被部分遮挡人体检测问题.基于人体部位的检测方法,将人体分成若干部位,分别检测各个部位,然后按人体各个部位的几何关系进行判定,从而得出最终的检测结果.由于基于部位的人体检测方法可以增强人体检测的鲁棒性,有许多有关这方面的研究<sup>[3]</sup>.

在文献[4]中,Anuj Mohan 等提出了一个基于样本的人体检测框架,通过用部件在单帧图片中检测人体.这个人体检测系统由 4 个部位检测器和一个装配检测器组成.4 个部位检测器分别为头,腿,左臂,和右臂检测器.结果显示这个系统的检测性能比单个 full-body 检测器明显要好.它能够检测出部分被遮挡的人和某些部位与背景对比度较低的人体.然而,这个系统检测一张 720 × 480 的图片需花将近 20 分钟,且检测性能依赖高分辨率的人体形状.

Bo Wu 等提出了一个用 Edgelet 特征描述轮廓模式的基于部位的静止人体检测方法<sup>[5]</sup>.在部位检测器的响应和显式为人之间建模的基础上,其方法为多个人体定义了一个联合似然率.对单个人体的检测效果很好,而且能够处理人群拥挤的场景.然而,该方法只检测近似正面或背面的人体(左右旋转角度限制在  $[-30^\circ, +30^\circ]$ ),且人体需直立或行走,摄像机的向下倾角不得超过  $45^\circ$ .文献[4,6]也有类似的限制.在本文中利用 Bandelet 变换中的 Bandelet 系数及其统计特征作为图像的特征,结合 Adaboost 分类器,使用分块提取特征的多级部位进行检测,即先提取整幅图像的 Bandelet 特征,然后将特征矢量细分成若干小块;按整个人体的比例提取运动区域输入分类器进行分类,得到候选人;然后依次检测各部分,同时依据各个部位检测器的检测似然度综合判定.结果表明,此方法可显著提高图像中人体目标检测的鲁棒性.

## 2 Bandelet 变换

本文采用了 Peyré 和 Mallat 2004 年提出的第二代 Bandelet 变换<sup>[4]</sup>,通过多尺度分析和几何方向分析共同完成图像的分解.

第二代 Bandelet 变换算法步骤

**Step 1** 二维离散正交小波多尺度变换;

**Step 2** 结合二叉树分割法和 CART(自底向上融

合)算法建立多尺度图各子带的最佳二叉树分解,得到 Bandelet 块;

**Step 3** 对各 Bandelet 块,根据 Lagrange 罚函数法求取一个最优方向,即最佳几何流方向;

**Step 4** 根据每个 Bandelet 块中的最佳几何流方向,借助正交投影和小波系数重排,得到一个一维离散信号,并对实施一维离散小波变换.

### 2.1 二叉树分割优化算法

子带的分割同第一代 Bandelet 一样,采用二进分割方法.先将各子带等分成四个子带,每一子带在下一层的分割中又被分成四个子带,依次下去直到底层子带达到预先设定的最小尺度.分割的结果可以用二叉树表示,分割过程中产生的 Bandelet 块与二叉树的节点一一对应.

为了得到各子带最佳的二叉分割方式,Peyré 采用自底向上的全局优化算法<sup>[6]</sup>,其算法流程如下:

**Step 1** 对每个  $L \times L$  的方块  $S$ ,计算最佳几何流方向,及最小的 Lagrange 函数值  $L_0(S)$ ;

**Step 2** 令  $L = 2L$ ,对每个  $L \times L$  方块(仍可记为  $S$ ),计算最佳几何流方向和相应的 Lagrange 函数值  $L(S)$ ;

**Step 3** 对每个尺寸为  $L \times L$  的方块  $S$ ,它的四个孩子记为  $(S_1, S_2, S_3, S_4)$ ,计算这四个孩子作为叶节点联合在一起的 Lagrange 函数值:

$$\tilde{L}(S) = L_0(S_1) + L_0(S_2) + L_0(S_3) + L_0(S_4) + L_0(S) + \lambda \cdot T^2 \quad (1)$$

**Step 4** 令  $L_0(S) = \min\{L(S), \tilde{L}(S)\}$ ;

**Step 5** 若  $L$  小于最大的分割尺度,则重复 Step2、Step3、Step4.

$L_0(S)$  即为最终的二叉树分割结果,同时还可以得到各 Bandelet 块的最佳几何流方向.将具有基本相同几何流方向的相邻 Bandelet 块合并在一起,逐步合并一些较小的 Bandelet 块,可以得到最终 Bandelet 块划分结果.

### 2.2 最佳几何流搜索

如上所述,在二叉树分割的优化算法中,需要同时进行最佳几何流方向的搜索,最后可以得出各 Bandelet 的最佳几何流方向.其中各子块的最佳几何流方向通过以下步骤确定<sup>[5-8]</sup>.

对于尺寸为  $L \times L$  的子块  $S$ ,将圆周角  $[0, \pi)$  等角度离散为  $L^2 - 1$  个:

$$\theta = \frac{k\pi}{L^2 - 1}, \quad k = 0, 1, 2, \dots, L^2 - 1 \quad (2)$$

对于无几何流的情形,标记  $\theta = \text{inf}$ ,表示不实施 Bandelet 化.这样  $\theta$  的取值共有  $L^2$  个.

对于每个  $\theta(\theta \neq \text{inf})$ , 构造一个  $L \times L$  大小的网格点, 计算每个网格点(坐标为  $(x(i), y(j))$ )在采样角度上的正交投影误差:

$$t = -\sin(\theta) \cdot x(i) + \cos(\theta) \cdot y(j) \quad (3)$$

将网格点按误差值从小到大排序, 可以得到一个  $L^2 \times 1$  的排序索引. 然后将这个小块内的二维离散小波变换后的系数按照排序索引进行重排序, 得到一个一维信号  $f_d$ .

对此一维信号  $f_d$  进行一维小波变换得到  $f_\theta$ , 同时计算其量化值  $\tilde{f}_\theta$ , 量化公式为:

$$Q(x) = \begin{cases} 0, & |x| \leq T \\ \text{sign}(x) \cdot (q + 0.5) \cdot T, & qT \leq |x| \leq (q + 1)T \end{cases} \quad (4)$$

$T$  为量化阈值,  $q \in \mathbb{Z}$ .

区域  $S$  内的最佳几何流方向, 应使得下面的 Lagrange 函数最小:

$$L(f_\theta, R) = \|f_\theta - \tilde{f}_\theta\|^2 + \lambda \cdot T^2(R_g + R_b) \quad (5)$$

$R_g$  表示编码几何流所需比特数,  $R_b$  表示编码量化后的 Bandelet 系数所需比特数.  $\lambda$  是 Lagrange 乘子.

## 3 特征提取

### 3.1 数据库与性能评价方法

本文用于人体检测的样本来自 INRIA 人体数据库<sup>[9]</sup>, 从中选取了 6656 幅样本图像. 其中 2416(正) + 1877(负)幅作为训练集, 1132(正) + 1213(负)幅作为测试集, 样本大小均为  $128 \times 64$  像素. 图 1 是部分样本.



(a) 部分正样本



(b) 部分负样本

图1 部分样本

为了确定特征提取的最优参数, 我们使用相同的性能评价方法来进行分类性能的比较. 即对比不同参数情况下假阳率(False Positive Rate)对不同的检测正确率(Detection Rate)的 ROC 曲线, ROC 曲线越靠上, 表示对应的系统性能越好.

### 3.2 最佳参数选择

Bandelet 变换用于图像压缩时以尽量减少非零系数个数为目的, 此处则借助 Bandelet 变换进行特征提取, 所以进行 Bandelet 变换时的参数会有所不同, 需要

通过试验来确定. 为此, 针对影响 Bandelet 变换的主要参数, 包括二进剖分的最小尺度  $j_{\min}$ 、向上建立四叉树的最大尺度  $j_{\max}$ 、二维小波变换的层数 level、以及量化阈值  $T$ , 使用训练集进行了参数选择试验.

#### 3.2.1 二维小波变换层数 level

本文选取了 1 到 5 层的二维小波变换, 以及不进行二维小波变换, 共 6 种情况进行比较. 其中,  $j_{\min} = 2$ ,  $j_{\max} = 2$ ,  $T = 15$ , 结果如图 2 所示.

如图 2 所示, 仅使用一层的二维小波变换, 效果最好. 分析其原因为: 分解层数越大, 丢失的上一层低频近似系数的特征表征能力, 要大于所得到的高频细节系数的特征表征能力, 所以仅使用一层的二维小波变换也有利于最后检测时提取检测图像上所有扫描窗口的特征值.

#### 3.2.2 $j_{\min}$ 与 $j_{\max}$

就 Bandelet 变换而言, 理论上  $j_{\min}$  越小, 同时  $j_{\max}$  越大越好, 这样建立的四叉树最合理. 但较小的  $j_{\min}$  以及较大的  $j_{\max}$  会增加时间复杂度. 如果  $j_{\min}$  与  $j_{\max}$  越接近, 甚至相同并且都很小时仍具有很好的特征表达效果, 那将大大降低特征提取所需的时间. 为此, 针对 level = 1,  $T = 15$ , 结果如图 3 所示.

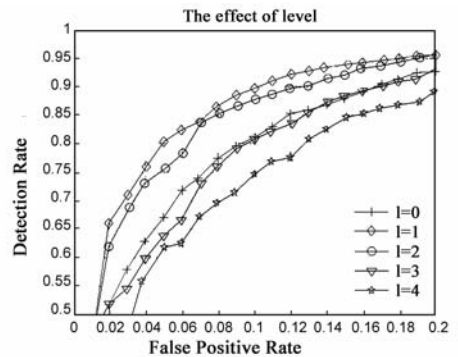


图2 参数Level选择结果

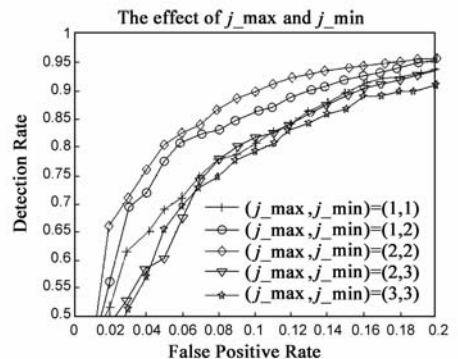


图3 参数j\_min与j\_max选择结果

由此可见, 仅使用  $4 \times 4$  大小的块进行 Bandelet 变换提取特征, 表征能力相对最好, 同时耗时也相对较少.

#### 3.2.3 量化阈值 $T$

在图像编码中用来控制压缩比,  $T$  越大, 压缩比越

高,失真也越大.另一方面,又影响了沿某方向上的一维小波变换时大于  $T$  的系数值的大小,所以  $T$  选取过大过小都不利于找到最佳几何流方向.针对不同的处理对象, $T$  值也会有所不同,需要通过试验找到最佳的  $T$  值.取  $level = 1, j_{-min} = 2, j_{-max} = 2$  时,试验结果如图 4 所示.

从图 4 中可以看出, $T$  的选取对训练误差率和测试正确率影响不是很大,这应该是由图像的多样性导致的.本文中,将其取为 15.

### 3.3 统计特征选择

特征的选择对人体检测的准确性有非常重要影响,Bandelets 系数能够体现图像在不同尺度上的目标边缘与轮廓的信息,为了充分利用这些信息,本文算法中还采用了 Bandelets 系数的如下统计特征:能量:

$$e = \sum_{i,j} c_{i,j}^2, \text{熵: } - \sum_{i,j} c_{i,j}^2 / e \times \log(c_{i,j}^2 / e), \text{均值: } \frac{1}{N} \sum_{i,j} c_{i,j}^2,$$

最大值:  $\max(c_{i,j})$ . 其中,  $c_{i,j}$  即为各个 Bandelet 块中  $(i, j)$  坐标上的系数,  $N$  为各个 Bandelet 块中所含元素的个数.这些统计量反映了不同尺度上的方向信息.从图 5 可以看到,加入了统计特征后,ROC 曲线要高于采用 HOG 和单纯使用 Bandelet 特征的曲线.

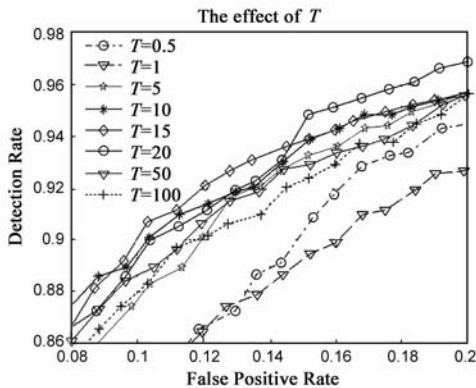


图4 参数T选择结果

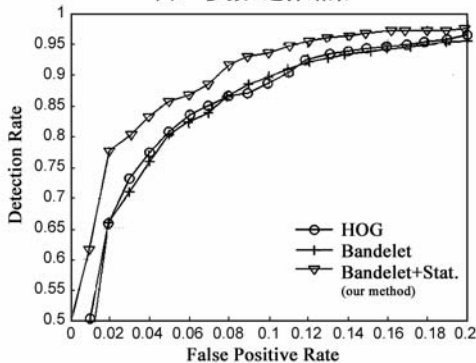


图5 使用统计特征的结果比较

### 3.4 人体部位的取样

所使用的人体样本取自 INRIA 静态人体数据库,部分人体样本如图 1 所示.人体的部位样本则是基于此人

体样本集,按照一定的位置及大小截取下来的.考虑到人体头部的位置变化较大,首先对头肩部进行手动截取,截取大小均为  $48 \times 48$  像素.舍弃一些头肩部偏离较严重的样本,共得到 2277 幅头肩部样本,记为 HS.将对应的人体样本记为 FB.部分样本如图 6 所示.



图6 部分头肩部样本

另外,按人体上部 32-80 像素进行截取,可得到人体躯干样本,大小均为  $64 \times 48$  像素,记为 TR.取人体的下部  $64 \times 64$  像素大小的部分作为人体下肢样本,大小均为  $64 \times 64$  像素,记为 L.部分躯干及下肢样本如图 7、8 所示.各部位在整个人体中的位置及大小如图 9 所示.

按图 9 所示的人体部位位置和大的定义,可从非人体样本图像上截取取出各人体部位的负样本图像.

为了评价各个部位及整个人体的分类性能,首先,使用第二、三部分提出的联合 Bandelet 特征提取方法,



图7 部分躯干样本



图8 部分下肢样本

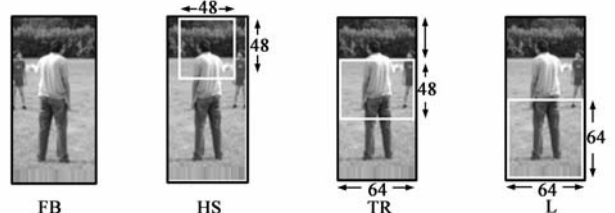


图9 人体部位定义

提取训练样本(FB)及其各部位 HS、TR、L 的特征值,分别用于训练 Linear-SVM 分类器.然后,提取测试样本(FB)及其各部位 HS、TR、L 的特征值,并使用对应训练好的 Linear-SVM 分类器进行分类性能比较.比较使用了 Detection Rate 对 False Positive Rate 的 Roc 曲线来表示,如图 10 所示.

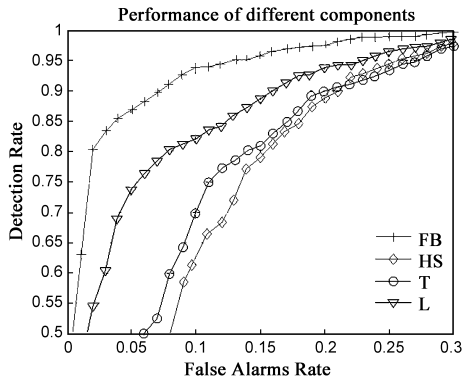


图10 人体及其各部位分类性能比较

试验结果表明:(1)单独任一分类器的分类性能都要远远弱于使用整个人体时的分类性能;(2)头部的分类性能是三者中最弱的,这是由于头部的变化最大.

### 3.5 特征提取方法

全人体图像的 Bandedet 特征提取方法:

(1)对每个全人体图像进行二维离散正交小波变换,小波变换的层数为 1;

(2)将二维离散正交小波变换后的图像进行  $4 \times 4$  像素的二进剖分,并将每个  $4 \times 4$  像素的小块作为一个 Bandedet 块;

(3)计算每个 Bandedet 块在各方向上的投影误差,并按最小 Lagrange 函数法,求取一个最优的投影误差排序索引;

(4)根据每个 Bandedet 块的最优排序索引,将其上的二维离散小波变换系数进行重排序,得到对应的一维信号;

(5)将每个一维信号作一维小波变换,并将对应的一维小波变换系数,按 Mallet 法则重新组合为二维形式,作为对应 Bandedet 块的 Bandedet 系数,将所有 Bandedet 块的 Bandedet 系数构成该全人体图像的 Bandedet 系数矩阵;

(6)按每个 Bandedet 块,提取 Bandedet 系数的统计特征,联合 Bandedet 系数作为该全人体图像的 Bandedet 特征.

在部位的特征提取上,使用分小块提取方法.图 11 给出了分块大小为  $8 \times 8$  像素时,特征提取的方法.从待检图像的特征矩阵中提取各部位特征的方法如图 12 所示.先按一定的缩放尺度,将待检图像进行缩放,再提取整幅图像的特征 Bandedet 特征,然后将特征矢量细分成若

干小块.该方法的一个优点是只需对整幅图像提取一次特征,避免大量重复的特征提取操作,而且各个部位大小可以自适应,不再受到检测窗口大小限制.

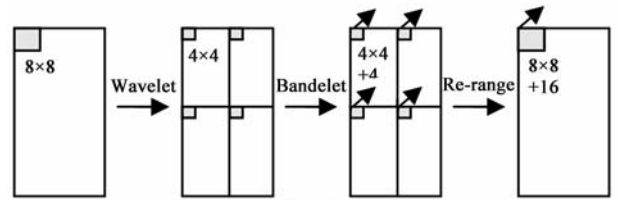


图11 特征提取方法

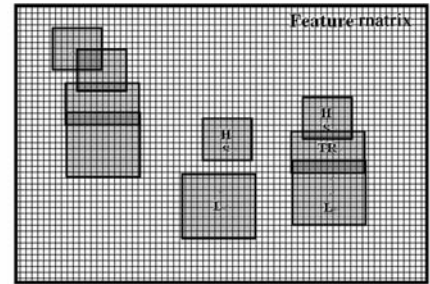


图12 分块特征提取

我们设立了一个多阶段基于部位的人体检测方法:

第 1 阶段,从图像的特征矢量中提取整个人体大小区域上的特征,其中包含了一定数量的小块特征,然后输入全人体分类器进行分类,得到候选人体.这个阶段要求在假阳率不过高的情况下,检出率尽量高;

第 2 阶段,用各部位分类器检测候选各部位所在的部位,并计算其联合似然率.头肩(HS)、躯干(TR)及下肢(L)的联合似然率分别记为  $M_1$ 、 $M_2$  和  $M_3$ .

第 3 阶段,若  $M_1 + M_2$ 、 $M_2 + M_3$ 、 $M_1 + M_3$  中任意一个大于某一阈值  $t_1$ ,则判定该候选人体为真人体,否则进入第 4 阶段;

第 4 阶段,各个部位检测器对某个候选人体的检测结果被结合起来形成该候选者的联合似然率  $M = M_1 + M_2 + M_3$ .若  $M$  大于某一阈值  $t_2$ ,仍判定该候选人体为真人体,否则判定该候选人体为非人体.

## 4 部位结合策略

改进方法的部位检测是建立在候选人体区域之上的,所以在候选人体区域  $X$  中,头肩、左躯干、右躯干和下肢被检测出的概率分别为<sup>[5]</sup>:

$$P\{Y_i = 1 | Y_{FB} = 1, X_i\} \quad (6)$$

其中  $Y_i = 1$  表示人体的部位  $i$  被检测出,  $X_i$  表示人体的各部位区域,  $i \in \{HS, TR, L\}$ .由于人体的各个部位的检测是独立的,所以可推出下式:

$$P\{Y_i = 1 | Y_{FB} = 1, X_i\} = \prod_{i \in \{HS, TR, L\}} P\{Y_i = 1 | X_i\} \quad (7)$$

设  $Z_i$  是部位  $i$  检测器的最后一段的响应值,  $Y_i = 0$  表示人体的部位  $i$  未被检测出,  $i \in \{HS, TR, L\}$ .根据

文献[6],各部位的检出和未被检出的后验概率为:

$$P\{Y_i = 1 | X_i\} = \frac{e^{z_i}}{e^{-z_i} + e^{z_i}} \quad (8)$$

$$P\{Y_i = 0 | X_i\} = \frac{1}{e^{-z_i} + e^{z_i}} \quad (9)$$

由式(7)、(8)可得:

$$P\{Y_i = 1 | Y_{FB} = 1, X_i\} = \frac{e^{\sum z_i}}{\prod (e^{-z_i} + e^{z_i})} \quad (10)$$

设各部位的实际检测的正确率为  $T_i \in \{0, 1\}$ ,  $i \in \{HS, TR, L\}$ , 则有:

$$P\{Y_i = T_i | Y_{FB} = 1, X_i\} = \frac{e^{\sum w_i z_i}}{\prod (e^{-z_i} + e^{z_i})} \quad (11)$$

$w_i$  按下式计算:

$$w_i = \begin{cases} e^{\lambda_1 \left( \frac{(x_{si} - x_{di})^2}{\sigma_x^2} + \frac{(y_{si} - y_{di})^2}{\sigma_y^2} \right) + \lambda_2 \left( \frac{(H_{si} - H_{di})^2}{\sigma_H^2} + \frac{(W_{si} - W_{di})^2}{\sigma_W^2} \right)}, & T_i = 1 \\ 0, & T_i = 0 \end{cases} \quad (12)$$

其中,  $(x_{di}, y_{di})$  是候选人体实际部位  $i$  的质心,  $(x_{si}, y_{si})$  是候选人体检测出的部位  $i$  的质心,  $(W_{di}, H_{di})$  是候选人体实际部位  $i$  的大小,  $(W_{si}, H_{si})$  是候选人体检测出的部位  $i$  的大小,  $\lambda_1$  为一常数, 用于表示部位检出框位置与大小的重要性.

设  $c_i$  为检出的概率, 其中  $c_i \in (0, 1)$ ,  $i \in \{FB, HS, TR, L\}$ , 根据本文中提出的算法, 候选的人体即为人体的概率为:

$$\begin{aligned} p\{Y = 1 | X\} &= p\{Y_{FB} = 1, Y_i = c_i | X\} \\ &= p\{Y_i = c_i | Y_{FB} = 1, X\} * p\{Y_{FB} = 1 | X\} \\ &= \frac{e^{\sum w_i z_i + z_{FB}}}{\prod (e^{-z_i} + e^{z_i})} \end{aligned} \quad (13)$$

候选人体为非人体的概率为:

$$\begin{aligned} p\{Y = 0 | X\} &= p\{Y_{FB} = 0, Y_i = 0 | X\} \\ &= p\{Y_i = 0 | Y_{FB} = 0, X\} * p\{Y_{FB} = 0 | X\} \\ &= \frac{1}{\prod (e^{-z_i} + e^{z_i})} \end{aligned} \quad (14)$$

使用贝叶斯决策来确定由全人体(FB)检测器检出的候选人体是人还是非人. 候选人体的置信度可表示为如下式的似然比函数:

$$M_i = \log \frac{p\{Y_{FB} = 1, Y_i = T_i | X_i\}}{p\{Y_{FB} = 0, Y_i = 0 | X_i\}} = \sum w_i Z_i + Z_{FB} \quad (15)$$

根据第三部分中设立的多阶段的人体检测方法, 则组合检测器的输出如下:

$$H(X) = \begin{cases} 1, & M_1 + M_2 > t_1 \parallel M_1 + M_3 > t_1 \parallel M_2 + M_3 \\ & > t_1 \parallel M_1 + M_2 + M_3 > t_2 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

## 5 结果分析

首先, 利用 Bandedet 特征提取方法, 提取 INRIA 人体数据库中的全人体及其对应的头肩、躯干、下肢各部位的特征分别训练对应的 Adaboost 分类器.

图 13 为使用全人体分类器的分类性能, 其中使用了两种特征, 即 HOG 特征及 Bandedet 特征. 从中可以看出本文中提出的基于人体部位的分类器的性能都要高于其他分类器的性能, 同时使用 Bandedet 特征的性能要大大高于 HOG 特征的性能.

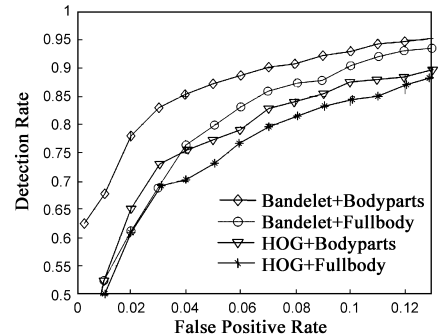


图13 各分类器的分类性能对比

试验选用的测试图像均来自 MIT 人体数据库, 图像大小在 500 像素左右. 使用的缩放因子为  $[0.5, 0.6, 0.7, 0.8, 0.9]$ . 对静态图像, 按照本文中提出的多阶段人体部位检测方法, 可以非常准确地检测出静态图像中的人体, 其中分块提取特征时使用的小块大小为  $8 \times 8$  像素的方块, 根据文献[12], 选择分类器阈值为  $t_1 = 8$ ,  $t_2 = 10$ . 扫描图像时, 使用浓密扫描, X 方向平移 8 个像素, Y 方向平移 16 个像素. 各部位的扫描窗口缩放大约 16 个像素, 上下左右平移尺度为 8 个像素.

人体检测的试验结果如图 14~18 所示. 图 14 给出了各个部位及全人体的检测结果, 其中细线框区域为所检测出的部位, 粗线框为最终检测出的人体区域. 图 14~18 分别给出了包含单人和多人, 背景混杂或者光照有变化的图像的本文算法以及 HOG 的检测结果.

比较图 14(e) 和 (f), 可以看到, 在单人的情况下, 本文的算法只出现了一个位于左上方的虚警框, 而 HOG 的检测结果不仅未将人体全部检出, 而且还存在多个虚警框. 而比较图 15(b) 和 (c), 虽然 HOG 也将图像中人体检出, 但本文算法最终给出的人体区域与 HOG 方法相比, 与实际区域更为符合. 而在图 16 中, 图片中存在的多人, 而且有相互遮挡的状况发生, 本文算法能够检测到被部分遮挡的人体, 而 HOG 则完全不能检出该人体.

图 17 则给出了混杂背景中存在有多人的情况下的检测结果. 比较图 17(b) 和 (c), 可以看到, 本文算法虽未将有严重遮挡的人体检出, 但并不存在虚警框. 而

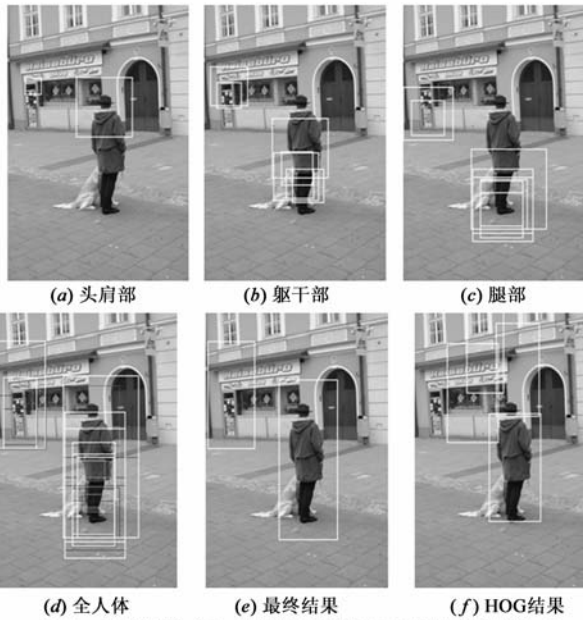


图14 基于部位的人体检测结果与HOG的结果比较(一)

HOG 不仅未将图片中存在的人体都检出,而且还存在多个虚警框,这主要是因为混杂的背景所引起的.图 18 给出了对有光照变化,且存在多人的图片的检测结果.比较图 18(b)和(c),本文算法除了未将存在严重遮挡的人体检出外,图片中存在的其他人体均被检出.HOG 算法则漏检了多个人体.

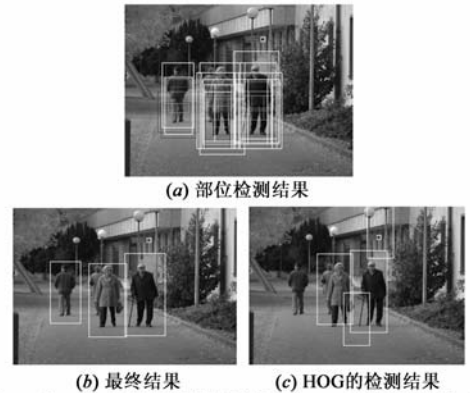


图17 基于部位的人体检测结果与HOG的结果比较(四)

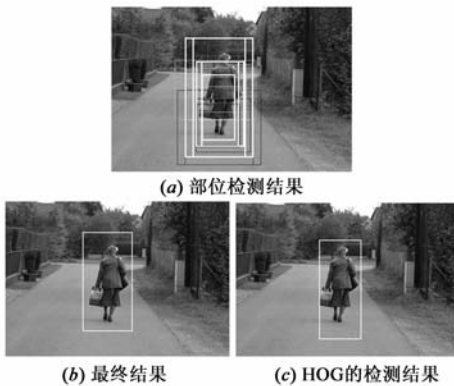


图15 基于部位的人体检测结果与HOG的结果比较(二)

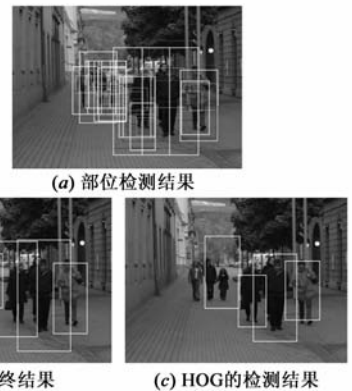


图18 基于部位的人体检测结果与HOG的结果比较(五)

综上所述,从对不同类型的图片的检测结果可以看到,即使图像中存在多个人体,且有互遮挡的状况发生,即使存在有混杂背景,本文方法也能取得良好的效果.且与文献[11]中的检测结果相比较,对于互有遮挡的多个人体的检测,检测结果更好.

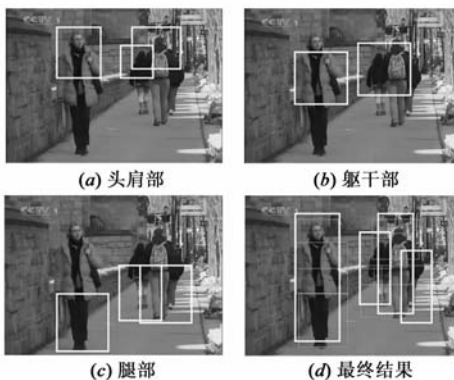


图16 基于部位的人体检测结果与HOG的结果比较(三)

### 6 结论

基于第二代 Bandelet 变换,本文提出了一种新的人体特征提取及部位检测方法.首先,通过参数选择试验确定 Bandelet 变换应用于特征提取时的最优参数,并通过特征选择试验筛选了部分统计特征作为附加特征来增强特征的分类能力.利用 Bandelet 变换中的 Bandelet 系数及其统计特征作为图像的特征.在进行特征提取时,将特征矢量细分成若干小块,使用分块提取特征的方式来提取样本及各扫描窗口上图像的特征值.与 HOG 利用图像本身的梯度方向特征不同,Bandelet 变换的中心思想则是定义图像中的几何特征为矢量场,而

不是看成普通的边缘的集合,矢量场则表示了图像空间结构的灰度值变化的局部正则方向.这也是本文算法所提取的特征要优于 HOG 方法的根本所在.在训练分类器时,使用 AdaBoost 算法来训练全人体及各部位分类器.最后,对于被测图像,先使用全人体分类器来获取可能包含人体的区域,然后计算对应人体区域各可能位置上部位的检测似然度,最后使用部位结合策略来进行人体判定.

通过对比使用不同特征提取方法训练的具有相同假阳率的分类器对测试样本的分类结果可以看出,本文所提出的方法的分类正确率要高于 HOG 特征提取方法.这表明本文所提出的特征提取方法能有效的表征人体并提高分类器的分类性能.文中还对比了全人体分类器及本文提出的基于人体部位的分类器的分类性能,结果表明本文提出的基于人体部位的分类器的性能都要高于其他分类器的性能.

#### 参考文献

- [1] N Dalal, B Triggs. Histograms of oriented gradients for human detection[A]. CVPR 2005[C]. San Diego: American Press, 2005. 886 – 893.
- [2] G Peyré, S Mallat. Second generation bandelets and their application to image and 3D meshes compression[A]. Mathematics and Image Analysis[C]. Paris, 2004. 6 – 9.
- [3] Mikolajczyk K, Schmid C, et al. Human detection based on a probabilistic assembly of robust part detector[A]. ECCV 2004[C]. Prague: Czech Republic Press, 2004. 69 – 82.
- [4] Mohan A, Papageorgiou C, et al. Example-based object detection in images by components[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2001, 23(4): 349 – 361.
- [5] WU B. Detection of multiple partially occluded humans in a single image by bayesian combination of edgelet part detectors [A]. ICCV 2005[C]. Beijing: China Press, 2005, (1): 90 – 97.
- [6] 田广, 戚飞虎, 等. 单目移动拍摄下基于人体部位的行人检测[J]. 系统仿真学报, 2006, 18(10): 2906 – 2910.  
Tian Guang, Qi Feihu, et al. Component-based pedestrian detection with monocular moving camera [J]. Journal of System Simulation, 2006, 18(10): 2906 – 2910. (in Chinese)
- [7] G Peyré, S Mallat. Surface compression with geometric bandelets[J]. ACM Transactions on Graphics (TOG), 2005, 24(3): 601 – 608.
- [8] Shuyuan Yang, Fan Liu, et al. Multiscale bandelet image compression[A]. ISPADS 2007[C]. Xiamen: China Press, 2007. 412 – 415.
- [9] INIRIA human database[DB/OL]. <http://lear.inrialpes.fr/data>, 2005.

- [10] 焦李成, 侯彪, 等. 图像多尺度几何分析理论与应用——后小波分析理论与应用[M]. 西安: 西安电子科技大学, 2008.
- [11] 张文革, 刘芳, 等. 基于 Bandelet 域逐子块阈值的图像去噪[J]. 电子学报, 2010, 38(2): 290 – 294.  
Zhang Wenge, Liu Fang, et al. Image denoising using sub-squarewise threshold in bandelet demain[J]. Acta Electronica Sinica, 2010, 38(2): 290 – 294. (in Chinese)
- [12] E LE Pennec, S Mallat. Sparse geometrical image representation with bandelets[J]. IEEE Trans Image Processing, 2005, 14(4): 423 – 438.
- [13] Saad Ali, Mubarak Shah. A supervised learning framework for generic object detection in image[A]. ICCV 2005[C]. Beijing: IEEE, 2005. 2. 1347 – 1354.
- [14] Paclik P, Novovicova J, et al. A trainable similarity measure for image classification[A]. ICPR 2006[C]. Hong Kong: IEEE, 2006. 3. 391 – 394.
- [15] Friedman J, Hastie T, et al. Additive logistic regression: A statistical view of boosting[J]. The Annals of Statistics, 2000, 28(2): 337 – 407.
- [16] Q zhu, S Avidan, M Yeh, et al. Fast human detection using a cascade of histograms of oriented gradients[A]. CVPR 2006[C]. New York: American Press, 2006. 1491 – 1498.
- [17] P Felzenszwalb, D McAllester, et al. A discriminatively trained, multiscale, deformable part model[A]. CVPR 2008[C]. Anchorage, AK: American Press, 2008. 1 – 8.
- [18] Lin L, Wu T F, et al. A stochastic graph grammar for compositional object representation and recognition [J]. Pattern Recognition, 2009, 42(7): 1297 – 1307.
- [19] TaiPeng Tian, Stan Sclaroff, Fast globally optimal 2D human detection with loopy graph models[A]. CVPR 2010[C]. San Francisco: American Press, 2010. 81 – 88.

#### 作者简介



韩红女, 1974 年生于陕西西安. 中国电子学会高级会员, 副教授, 硕士生导师. 研究方向为机器视觉与多源信息融合.  
E-mail: hanh@mail.xidian.edu.cn



伍星男, 1986 年生于江西南昌. 硕士研究生, 研究方向为机器视觉.