

# 基于后验 HOG 特征的多姿态行人检测

刘 威,段成伟,遇 冰,柴丽颖,袁 淮,赵 宏

(东北大学研究院,辽宁沈阳 110179)

**摘 要:** 行人检测是当前计算机视觉领域的挑战性课题之一.本文提出一种基于后验 HOG 特征的多姿态行人检测方法.首先,统计全部行人样本的梯度特征能量共性信息,对单个行人样本的 HOG 特征进行加权获得能够表现行人边缘轮廓的后验 HOG 特征,有效减少复杂背景的影响.其次,利用 S-Isomap 特征降维方法和 K-means 聚类方法对不同姿态和视角的行人做子类划分,并针对每一个子类训练子类分类器.最后,根据多个不同姿态的子类分类器输出值,训练等权重加和方式的多姿态-视角集成分类器.不同数据集上的测试结果表明,本文所提利用共性信息获得的后验特征超过了经典 HOG 和其它典型特征的描述能力.与现有方法相比,通过将所提出的特征与多姿态-视角集成分类器结合,有效地提高了检测精度.

**关键词:** 后验 HOG 特征; 梯度能量图; S-Isomap; 支持向量机; 行人检测

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 0372-2112 (2015)02-0217-08

**电子学报 URL:** <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2015.02.002

## Multi-Pose Pedestrian Detection Based on Posterior HOG Feature

LIU Wei, DUAN Cheng-wei, YU Bing, CHAI Li-ying, YUAN Huai, ZHAO Hong

(Research Academy, Northeastern University, Shenyang, Liaoning 110179, China)

**Abstract:** Pedestrian detection remains one of the challenging tasks in the area of computer vision. A multi-pose pedestrian detection method based on posterior HOG feature is proposed. Firstly, the generality information of gradient feature energy is computed with all pedestrian samples. The posterior HOG feature is obtained by weighting the HOG feature of individual pedestrian sample with the computed gradient feature energy. The posterior HOG feature can capture the contours and edges of pedestrians, and significantly reduce the influence of complex and cluttered background. Secondly, pedestrians of different poses and views are divided into subclasses with S-Isomap and K-means algorithm. A classifier is trained for each subclass. Finally, a multi-pose-view ensemble classifier is trained to combine the output values of different subclass classifiers with an equally weighted sum rule. Experimental results on different datasets suggest that the proposed posterior feature outperforms the classic HOG feature and other typical features. Compared with the existing methods, by combining the posterior feature and the multi-pose-view ensemble classifier, the proposed method boosts the detection accuracy effectively.

**Key words:** posterior HOG feature; gradient energy map; S-Isomap; support vector machine; pedestrian detection

## 1 引言

行人检测在智能车、机器人、视频监控等方向有着重要的应用<sup>[1~3]</sup>.由于行人姿态可变,以及受光照、背景、服装、遮挡等因素影响,使得行人检测仍然是计算机视觉领域的一个挑战性课题.

目前基于计算机视觉的行人检测大多是基于特征提取和机器学习的方法<sup>[4~7]</sup>.在特征提取方面,轮廓、纹理、频域信息以及颜色区域等特征常用于描述行人与背景的区别.如 HOG<sup>[8]</sup>,EOH<sup>[9]</sup>,Edgelet<sup>[10]</sup>,Shapelet<sup>[11]</sup>等特

征描述了行人的轮廓特征,LBP 特征<sup>[12]</sup>描述了行人的纹理特征,CSS 特征<sup>[13]</sup>利用局部部位间的颜色相似性刻画了人体的结构区域特征,Haar 小波特征<sup>[14]</sup>描述行人的频域信息.在众多特征中,Dalal 等提出的 HOG 特征描述了图像局部区域的梯度强度以及梯度方向分布,对行人外观和形状可以进行很好的表征,且对光照和小量的偏移不敏感,在行人检测中表现出优良性能,成为目前行人检测的主流方法.一些学者对经典 HOG 特征也做了很多改进,如 LMGO-HOG<sup>[15]</sup>,与熵结合的 EHO<sup>[16]</sup>等,进一步提高了 HOG 特征的描述能力.不过,单个特征仅

从单个角度(如轮廓、颜色、频域或纹理)描述行人,描述能力有限.为此,一些文献提出了以 HOG 特征为基础的组合特征,如 HOG-LBP<sup>[17]</sup>, HOG-CSS<sup>[13]</sup>.无论是经典的 HOG 特征、改进的 HOG 特征,还是以 HOG 为基础的组合特征,都仅利用了行人样本的个体信息,仍存在易受背景噪声影响的问题,对具有较强竖直纹理或者混乱边缘的背景造成较高的误识别<sup>[18]</sup>.

在行人检测算法中, SVM 和 boosting 是两种最常见的机器学习分类器.近年来,为了解决行人检测中由于视角、姿态不同导致类内外观差异较大的问题,一些文献按照姿态和视角的规律,把行人分成多个子类,降低行人姿态的类内差异,取得了良好的效果.如文献[19]按照行人不同姿态和视角,将行人分成 9 个子类,针对每个子类训练分类器并利用子分类器输出训练一个 Adaboost 行人分类器.文献[20]将行人样本划分为前、后、左、右及混合五个视角,每个视角都训练一个两阶段的 Adaboost 级联分类器,并采用投票方法决定最终分类.文献[21]将行人按距离(远、中、近)和视角(前、后、左、右)分成十二类样本,对每类样本分别训练一个级联分类器检测相应距离和视角的行人.文献[22]提出一种基于 LLE 和 K-means 聚类算法的子类自动划分方法,并采用纠错输出编码(ECOC)方法检测多姿态多视角行人.除文献[22]外,大多数文献中行人视角的划分都是人为主观进行的,如何更好地自动划分行人子类,研究则相对较少.

针对上述问题,本文做了以下工作.首先,提出一种后验 HOG 特征,该特征通过统计全部行人样本的共性信息——梯度特征能量,对个体样本的 HOG 特征进行加权处理,得到表征行人边缘的后验 HOG 特征.该特征融合了行人共性先验信息与个体信息,能强化行人

轮廓边缘,有效减少背景影响.而且,这种利用共性信息加权个体样本特征的方法也可以推广到其它用向量表示形式的特征,如 LBP 特征,可显著提高特征的描述能力.其次,利用 S-Isomap 特征降维方法和 K-means 聚类方法,自动按行人姿态和视角划分子类,降低行人的类内差异,对每一子类分别训练分类器,并采用加权集成方式集成这些姿态-视角子类分类器的输出值,有效的提高行人检测的精度.在不同测试集上的测试结果表明,同现有方法相比,本文提出的方法具有更好的准确率.

## 2 基于共性信息的后验 HOG 特征

在自然环境中,背景的边缘是杂乱无序的,而与背景相比,行人则体现了明显的规律.具有很强的共性信息.因此,本文提出基于梯度特征能量的后验 HOG 特征.梯度特征能量是大量行人正样本梯度信息的统计体现,可以分析行人的共性边缘信息.

对于一个行人样本  $f_i$ ,  $i = 1, 2, \dots, N$ ,  $N$  为正样本个数,将其归一化 $[W, H]$ 大小.计算每一像素的水平和垂直方向的梯度  $G_h(x, y)$  和  $G_v(x, y)$ , 得到  $f_i$  的梯度幅值矩阵  $m_i$ , 其中  $(x, y)$  位置的元素值为:

$$m_i(x, y) = \sqrt{G_h(x, y)^2 + G_v(x, y)^2} \quad (1)$$

其中  $1 \leq x \leq H, 1 \leq y \leq W$ .

计算所有样本的梯度幅值矩阵和  $M$ , 其中  $(x, y)$  位置的元素值为:

$$M(x, y) = \sum_{i=1}^N m_i(x, y) \quad (2)$$

归一化  $M$ :

$$\tilde{M}(x, y) = \frac{M(x, y)}{\max_{1 \leq x \leq H, 1 \leq y \leq W} M(x, y)} * 255 \quad (3)$$

这里,将  $\tilde{M}(x, y)$  称为无方向梯度能量图.

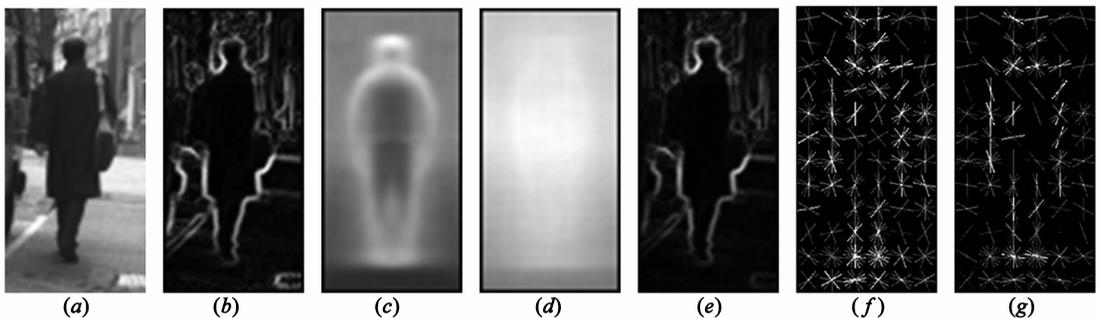


图1 梯度能量示意图.(a)背景复杂的行人样本;(b)(a)的梯度图;(c)正样本集的方向无方向梯度能量图;(d)负样本集的方向无方向梯度能量图;(e),(c)对(b)平滑后的梯度图;(f)经典HOG特征(SVM正权重作用后的HOG特征)示例图;(g)(f)与边缘共性信息作用后的后验HOG特征示例图

图1给出了戴姆勒训练数据集[7]的行人、背景的梯度能量示意图.由图1(c)可以看出,行人的无方向梯度能量图有一个明显的人体轮廓,在背景区域则表现

出较弱的梯度.而背景的方向无方向梯度能量图(图1(d))没有明显的轮廓特征.对于一些背景复杂的行人(如图1(a)),运用梯度能量图对其梯度图平滑,即梯度图与

梯度能量图对应位置的像素值相乘,然后全图像素值归一化到 $[0, 255]$ ,可以强化行人的轮廓边缘信息,减弱背景的影响,如图 1(e)所示.不过,在公式(3)中仅是对每个像素的梯度值加和,没有考虑梯度方向.由于背景边缘杂乱,且梯度方向无规律性,若同时考虑梯度方向,将能更有效突出行人边缘.为此,本文借鉴无方向梯度能量图思想,提出一种后验 HOG 特征,该特征考虑了所有样本的有方向梯度特征能量,能表达行人边缘特征的共性信息,强化每个样本的有用边缘特征.图 1(f)为经典 HOG 特征,从中能够看出明显的人体轮廓,但周边有杂乱的边缘噪声,图 1(g)是用本文提出的后验 HOG 特征图,该图能够看出明显的人体轮廓,而背景梯度特征被明显弱化.由此可以看出,本文提出的方法能够有效的强化人体轮廓边缘,弱化背景影响.后验 HOG 特征提取方法如下:

首先,计算每个行人训练样本的 HOG 特征向量  $x_i$  ( $i = 1, \dots, N$ ). 并把所有样本的 HOG 特征向量的对应元素加和,得到所有训练正样本的 HOG 梯度和值的特征向量,记为先验特征向量  $X$ :

$$X = \sum_{i=1}^N x_i \quad (4)$$

仿照 HOG 特征的归一化方式,以 Block 为单位对  $X$  做归一化. 本文实验比较了不同的归一化方式: L1-norm, L2-norm, 最终选择了 L1-norm. 归一化后的特征向量记为  $\bar{X}$ , 表达了行人边缘特征的共性信息.

其次,用归一化的先验特征向量  $\bar{X}$  对单个样本的 HOG 特征  $x_i$  做加权处理,即  $\bar{X}$  与单个样本特征  $x_i$  的对应元素相乘,得到单个样本的后验 HOG 特征向量  $\bar{x}_i$ , 其第  $j$  个元素值为:

$$\bar{x}_{i,j} = \bar{X}_j * x_{i,j} \quad (5)$$

在本文中,Block 和 Cell 的设置与经典 HOG 相同: Block 为  $16 * 16$  像素,Cell 的大小为  $8 * 8$  像素,方向 bin 个数为 9.

本文提出的后验 HOG 特征描述能力优于经典 HOG 以及其它典型特征的描述能力,更多信息见第本文 4 节实验结果.事实上,本文提出的利用样本共性信息提取个体样本的后验特征,不仅适用于 HOG 特征,同样也可推广到其它用向量形式表示的特征,如 LBP 特征. 这里,按照同样的步骤,首先计算所有样本的 LBP 特征,并把所有样本的 LBP 特征向量的对应元素加和,得到所有训练正样本的 LBP 特征和向量,记为先验特征向量  $X^{LBP}$ , 并以 Block 为单位对  $X^{LBP}$  做 L1-norm 归一化,归一化后的特征向量记为  $\bar{X}^{LBP}$ , 表达了行人纹理特征的共性信息.

其次,用归一化的先验 LBP 特征向量  $\bar{X}^{LBP}$  对单个样本的 LBP 特征做加权处理,即  $\bar{X}^{LBP}$  与单个样本的 LBP

特征向量  $x_i^{LBP}$  的对应元素相乘,得到单个样本的后验 LBP 特征向量  $\bar{x}_i^{LBP}$ , 其第  $j$  个元素值为:

$$\bar{x}_{i,j}^{LBP} = \bar{X}_j^{LBP} * x_i^{LBP} \quad (6)$$

为验证后验特征思想的适用性,本文同时给出了后验 LBP 特征的实验结果. 实验中采用的是均匀 LBP 特征<sup>[23]</sup>, 共有 59 个 LBP 模式, Block 大小为  $16 * 16$  像素.

### 3 多姿态-视角集成分类器

由于相机拍摄视角以及行人自身姿态变化,很难找到一种有效的特征能够同时表征不同姿态、视角的行人. 为此,一些文献按照姿态的规律,把行人分成多个子类,降低行人姿态的类内差异,取得了良好的效果. 文献<sup>[22]</sup>提出一种基于 LLE 和 K-means 聚类算法的子类自动划分方法,并采用纠错输出编码(ECOC)方法检测多姿态多视角行人. 考虑到 LLE 方法对噪声比较敏感<sup>[24]</sup>, 而且在高维空间中,样本具有很强的稀疏性,本文提出了利用 S-Isomap 特征降维方法和 K-means 聚类方法对不同姿态和视角的行人做子类划分,并集成各子类分类器的方法.

首先,利用 S-Isomap 方法<sup>[24]</sup>对特征降维. S-Isomap 是对 Isomap<sup>[25]</sup>方法的一种改进,能够把高维数据嵌入到低维空间,同时保持数据间的几何分布结构. Isomap 步骤如下:

(1) 构建输入数据的特征空间上的邻接图  $G$ . 距离度量为欧式距离  $d(x_i, x_j)$ , 处在  $\epsilon$  球内或  $k$  最近邻的点称为邻接关系点.

(2) 计算邻接图  $G$  上两点间的最短路径  $d_G(x_i, x_j)$ , 并估计流形上的测地线距离  $d_M(x_i, x_j)$ , 得到的矩阵  $d_G = \{d_G(x_i, x_j)\}$  为邻接图  $G$  上任意两点间的最短路径距离.

(3) 优化函数  $E = \|\tau(D_G) - \tau(D_Y)\|_{L^2}$  得到欧式空间  $Y$  的低维嵌入. 其中  $D_Y$  是空间  $Y$  中所用样本点的欧式距离矩阵,  $\|\cdot\|_{L^2}$  是矩阵的  $L2$  范数, 操作符  $\tau(\cdot)$  把欧式距离计算转化为内积计算  $\tau(D) = -HS\mathbf{H}/2$ , 其中  $D$  是  $N * N$  对称矩阵, 矩阵  $S$  的第  $(i, j)$  个元素  $S_{ij} = D_{ij}^2$ ,  $H$  是中心矩阵,  $H = I_N - \frac{1}{N}$ , 其中  $I_N$  是  $N * N$  单位矩阵. 本文中,空间  $Y$  的维数是 2.

S-Isomap 算法是一种有监督的降维算法,利用数据中的类别标签完成流形学习. 与 Isomap 不同,它将 Isomap 算法中的距离矩阵改进为

$$D(x_i, x_i) = \begin{cases} \sqrt{1 - \exp(-d^2(x_i, x_j)/\beta)}, & y_i = y_j \\ \sqrt{\exp(d^2(x_i, x_j)/\beta) - \alpha}, & y_i \neq y_j \end{cases} \quad (7)$$

其中  $y_i$  是样本  $x_i$  的类别标签,  $d(x_i, x_j)$  表示  $x_i$  和  $x_j$  之

间的欧式距离,参数  $\beta$  用于防止在  $d(x_i, x_j)$  较大时  $D(x_i, x_j)$  增长过快,参数  $\alpha$  用于控制不同类之间的距离不至于过远.从式(6)可以看出,当欧式距离  $d(x_i, x_j)$  相同时,类间距离比类内距离大.

然后,运用 K-means 方法对多姿态多视角行人自动聚成  $K$  类.图 2 给出了聚类数  $K=4$  时的戴姆勒训练数据集的聚类图,4 种不同颜色的点代表了聚类结果的分布.图中也给出了每个子类的代表样本,从代表样本可知,聚类算法大致将行人分为了前后视角/左右侧面视角腿并拢、左侧面视角腿中等分开、右侧面视角腿中等分开和腿大角度叉开的 4 个行人子类.

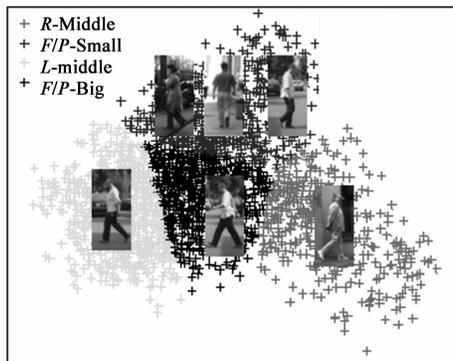


图2 行人聚类图

从图中聚类结果可以看出,本文提出的划分方法兼顾了行人的视角和姿态,同时,子类的划分也降低了行人的类内差异.事实上,K-means 方法聚类时必须事先指定类别个数  $K$ .而类别个数的多少将直接影响集成分类器的性能.在第 4 节中,本文分析了子类类别个数  $K$  与检测性能的关系.

在本文,对每一子类样本分别提取后验特征后,分别训练  $K$  个线性 SVM 子类分类器,最后再集成  $K$  个子分类器的输出,构成一个统一的多姿态行人检测器.对于每个子分类器的集成,采用将各子分类器的输出值直接加和方式集成,即:

$$O_T = \sum_{j=1}^K S_j + b \quad (8)$$

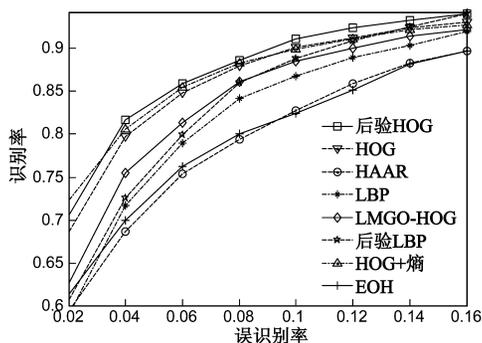
其中,  $S_j$  为第  $j$  个子分类器输出值.  $b$  为阈值.

这样集成的原因在于,视角和姿态是一个连续变化的量,聚类得到的有限姿态—视角划分是离散的,对于一个待识别行人,其姿态—视角很可能介于  $K$  个有限的姿态—视角之间,因此,可能同时被多个临近该行人姿态—视角的子分类器所识别.被越多的子分类器识别,就意味着待识别行人真正是行人的可能性越大,同时各子分类器之间的重要性相同.因此,本文采用加和原则集成子分类器,这样也相当于等权重求和集成.本文也与其它集成方法做了比较,见第 4 节实验结果.

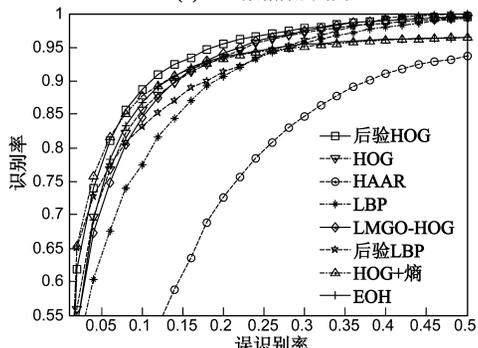
## 4 实验结果

为验证提出的后验 HOG 特征(后验 LBP 特征),以及划分行人子类进行多姿态-视角集成的方法的有效性,本文使用了两个公共数据集的数据进行实验:ETH 行人数据集<sup>[26]</sup>和 Daimler 行人数据集<sup>[6]</sup>.本文选择 ETH 数据集的 5 段视频作为行人分类器的训练集:正样本为 9022 个视角姿态各异的行人图像,负样本为从视频中与行人不相交区域随机取出的 10000 个尺寸各异的非行人图像.选用 ETH 数据集的剩余 1 段视频和 Daimler 数据集的测试视频作为测试集.其中,ETH 测试集包含 2684 名行人和随机产生的与行人区域不相交的 30000 个尺寸各异的非行人图像;Daimler 测试集包含 5054 名行人和随机产生的与行人区域不相交的 30000 个尺寸各异的非行人图像.

为了比较不同特征的描述能力,本文在上述训练集上分别采用不同特征训练线性 SVM 分类器,然后在测试集上测试.本文把样本图像归一化到  $96 \times 48$  大小,按照前文提出的 Block 和 Cell 的大小设置,得到 1980 维 HOG 特征,3245 维 LBP 特征.图 3(a), (b) 给出了本文提出的后验 HOG 特征(后验 LBP 特征)与其它六个现有特征(HOG<sup>[8]</sup>、EOH<sup>[9]</sup>、HAAR<sup>[14]</sup>、LMGO-HOG<sup>[15]</sup>、HOG + 熵<sup>[16]</sup>、LBP<sup>[23]</sup>)的 ROC 曲线比较结果.从比较结果看出,本文所提后验 HOG 特征明显优于其它特征,且后验 LBP 特征也好于原有的 LBP 特征.同时,本文也做了后



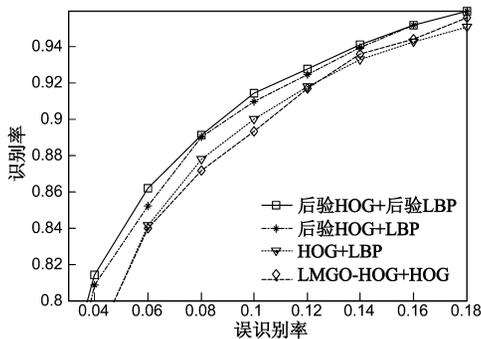
(a) ETH数据集的结果



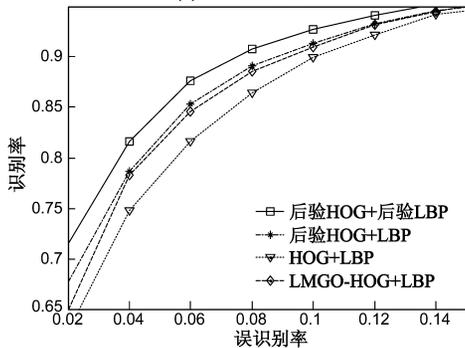
(b) Daimler数据集的结果

图3 在不同数据集上7种特征的比较结果

验 HOG + 后验 LBP、后验 HOG + LBP 与 HOG + LBP、LMGO-HOG + LBP 组合特征的比较结果,见图 4(a), (b), 结果表明本文的后验 HOG 特征(后验 LBP 特征)对于组合特征也有明显改善。



(a) ETH数据集的结果



(b) Daimler数据集的结果

图4 在不同数据集上组合特征的比较结果

为了分析不同子类个数  $K$  对检测性能的影响,本文分别设定了不同的子类个数:  $K = 2, 4, 6$ , 然后按 3 节中方法聚类,训练集成分类器,然后在测试集上比较不同  $K$  值对应的集成分类器性能,图 5 给出了在 ETH 测试集上的 ROC 曲线。

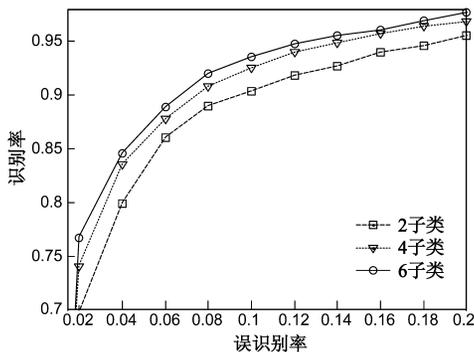


图5 不同子类个数的性能比较

从图 5 可以看出,子类个数 6 的集成检测器性能比子类个数 2 的性能有很大的提高,但是,与子类个数 4 的性能差不多.考虑到时间性能,本文确定  $K = 4$ 。

为验证提出的自动聚类方法训练集成分类器的有

效性,本文把行人划分为 4 个姿态子类,即  $K = 4$ ,并按 3 节所述方法训练集成分类器,并且与 LLE 方法划分行人子类(同样划分成 4 个子类)训练集成分类器、不划分行人子类直接训练一个分类器的方法进行了比较,结果显示在图 6. 这里,本文也同时给出了不同集成方法的比较结果。

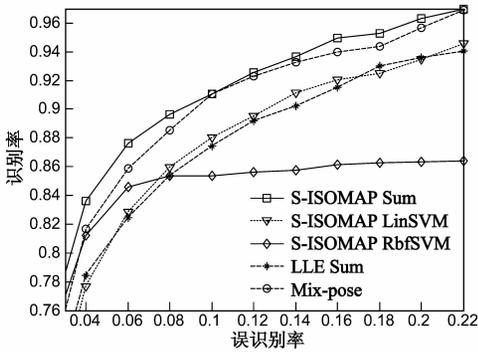
图 6 中,除 Mix-pose 方法外,其余方法都采用了 K-means 方法对降维后的样本进行聚类.其中,S-Isomap Sum 指用 S-Isomap 方法降维聚类,根据聚类结果分别采用线性 SVM 训练每个子类分类器,然后将各子类分类器输出值等权重加和集成;S-Isomap LinSVM 指用 S-Isomap 方法降维聚类,根据聚类结果分别采用线性 SVM 训练每个子类分类器,然后将全部子类分类器的输出值作为特征向量,用线性 SVM 训练,确定各子类分类器的权重,再根据该权重加和集成各子类分类器输出值<sup>[27]</sup>;S-Isomap RbfSVM 指用 S-Isomap 方法降维聚类,用高斯核函数 SVM 训练确定各分类器输出值的权重,再根据该权重加和集成各子类分类器输出值<sup>[20]</sup>;LLE Sum 指用 LLE 方法降维聚类,根据聚类结果分别采用线性 SVM 训练每个子类分类器,然后将各子类分类器输出值等权重加和集成;Mix-pose 指不划分行人子类直接训练一个线性 SVM 分类器.这里,为保证比较公平,所有方法都采用了本文提出的后验 HOG 特征。

从图 6 中的两个测试集结果可以看出,本文所用的 S-Isomap 方法聚类的结果明显好于 LLE 方法和不聚类(Mix-pose)方法.而且,当同时采用 S-Isomap 方法聚类时,本文所用的等权重加和集成的方法明显好于线性 SVM 权重加和、RbfSVM 权重加和两种集成方法.图 7 也给出了所有方法都基于后验 HOG + LBP 组合特征的比较结果。

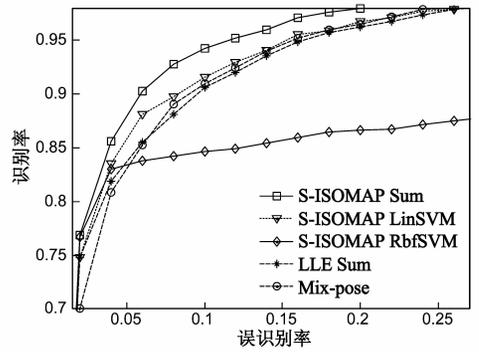
结果同样可以看出,本文所用的 S-Isomap 方法聚类的结果仍然明显好于 LLE 方法和不聚类(Mix-pose)方法.在不同集成方法比较方面,等权重加和集成方法依然表现出明显优势,除了在 ETH 数据集的一小段区间上略差于 RbfSVM 权重加和集成外.若同时考虑到时间性能开销,等权重加和集成方法则明显优于其它两种集成方法。

为进一步验证本文所提出检测方法的有效性,与文献[17, 27 ~ 29]四种方法进行了对比,结果如图 8、图 9 所示。

其中,图 8 中的本文方法<sup>[1]</sup>使用后验 HOG 特征结合提出的多姿态-视角集成分类器,图 9 中本文方法<sup>[2]</sup>使用后验 HOG + 后验 LBP 特征结合提出的多姿态-视角集成分类器.本文方法[3]使用后验 HOG 特征 + LBP 特征结合提出的多姿态-视角集成分类器,其它四种对比方法完全按照原文献描述方式实现,特征也采用各自文献所提到的特征.从比较结果可以看出,除了在图 8(a)所示的 ETH 数据集的一小段区间上本文方法[1]

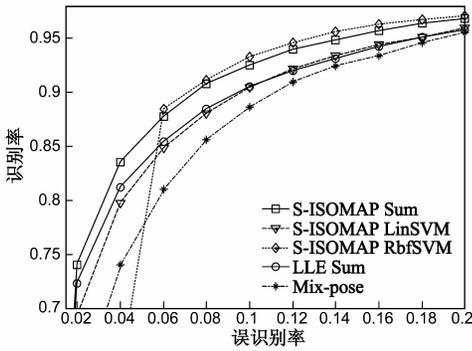


(a) ETH数据集的结果

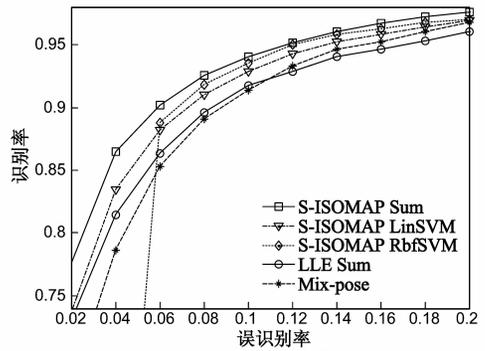


(b) Daimler数据集的结果

图6 基于后验HOG特征的比较结果

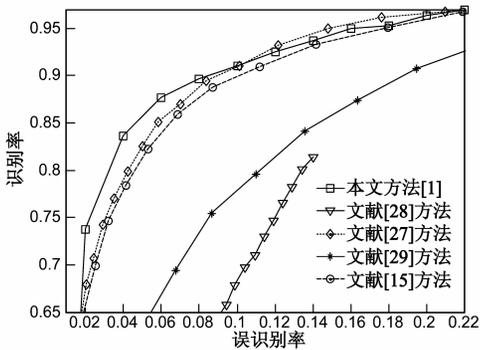


(a) ETH数据集的结果

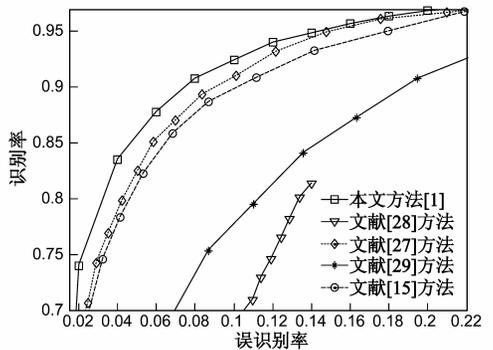


(b) Daimler数据集的结果

图7 基于后验HOG+LBP特征的比较结果

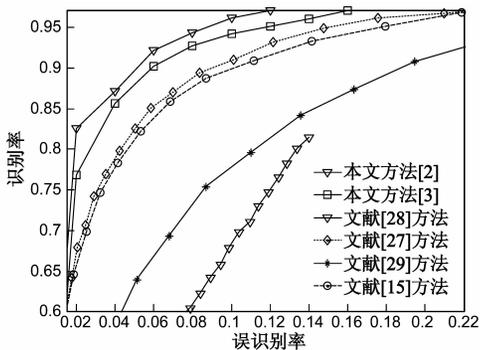


(a) ETH数据集的结果

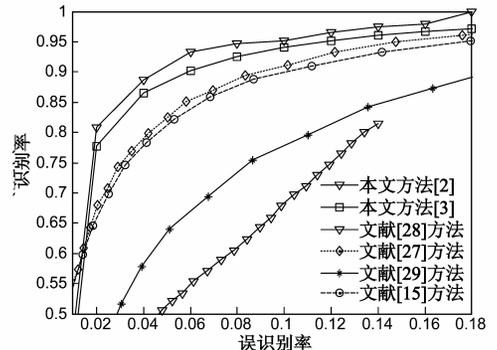


(b) Daimler数据集的结果

图8 本文方法与其它方法的比较结果



(a) ETH数据集的结果



(b) Daimler数据集的结果

图9 本文方法与其它方法的比较结果

略差于文献[27]方法外,本文方法([1]~[3])均明显好于其它四种方法,说明了所提出方法的有效性.特别是图 9 中的本文方法[3]明显好于本文方法[2],说明相比后验 HOG 特征 + LBP 特征,后验 HOG + 后验 LBP 特征明显提高了分类能力,这也说明了本文提出的基于共性信息获得后验特征的有效性.

## 5 结论

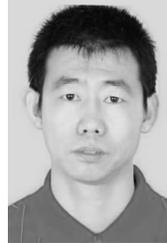
本文提出了基于后验 HOG 特征的多姿态行人检测方法.后验 HOG 特征继承了 HOG 特征的优点,同时有效利用了行人样本的共性信息和单个样本的个性信息,有效突出了行人轮廓特征,对背景噪声具有很好的抑制作用.而且,本文所提出这种利用共性信息加权个体样本特征的方法也可以推广到其它用向量形式表达的特征,显著提高特征的描述能力.同时,考虑到行人姿态多样性,本文提出基于 S-Isomap 和 K-means 的行人自动聚类方法,利用聚类结果训练不同姿态和视角的行人子类分类器,并利用等权重加和方式集成不同姿态和视角的子分类器,有效降低了行人类内差异的影响.与现有方法在不同数据集的实验对比结果表明了本文提出方法的有效性.

## 参考文献

- [1] Y W Xu, D Xu, S Lin, et al. Detection of sudden pedestrian crossings for driving assistance systems[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2012, 42(3): 729 – 739.
- [2] D Geronimo, A M Lopez, A D Sappa, et al. Survey of pedestrian detection for advanced driver assistance systems[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(7): 1239 – 1258.
- [3] 田广, 戚飞虎. 移动摄像机环境下基于特征变换和 SVM 的分级行人检测算法[J]. *电子学报*, 2008, 36(5): 1024 – 1028.  
G Tian, F H Qi. Feature transformation and SVM based hierarchical pedestrian detection with a monocular moving camera [J]. *Acta Electronica Sinica*, 2008, 36(5): 1024 – 1028. (in Chinese)
- [4] B Li, Q Yao, K Wang. A review on vision-based pedestrian detection in intelligent transportation systems[A]. 9th IEEE International Conference on Networking, Sensing and Control[C]. Beijing: IEEE, 2012. 393 – 398.
- [5] P Dollar, C Wojek, B Schiele, et al. Pedestrian detection: An evaluation of the state of the art[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(04): 743 – 761.
- [6] M Enzweiler, D M Gavrila. Monocular pedestrian detection: survey and experiments [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(12): 2179 – 2195.
- [7] 苏松志, 李绍滋, 陈淑媛, 等. 行人检测技术综述[J]. *电子学报*, 2012, 40(4): 814 – 820.  
S Z Su, S Z Li, S Y Chen, et al. A survey on pedestrian detection[J]. *Acta Electronica Sinica*, 2012, 40(4): 814 – 820. (in Chinese)
- [8] N Dalal, B Triggs. Histograms of oriented gradients for human detection[A]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* [C]. San Diego, CA, USA: IEEE, 2005. 886 – 893.
- [9] D Gerónimo, A López, D Ponsa, A Sappa. Haar wavelets and edge orientation histograms for onboard pedestrian detection [A]. *Pattern Recognition and Image Analysis* [C]. Germany: Springer Berlin Heidelberg, 2007. 418 – 425.
- [10] B Wu, R Nevatia. Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors[A]. *IEEE International Conference on Computer Vision* [C]. China: IEEE, 2005.
- [11] P Sabzmeydani, G Mori. Detecting pedestrians by learning shapelet features[A]. *IEEE Conference on Computer Vision and Pattern Recognition* [C]. Minneapolis, MN: IEEE, 2007. 1 – 8.
- [12] Y Mu, S Yan, Y Liu, et al. Discriminative local binary patterns for human detection in personal album[A]. *IEEE Conference on Computer Vision and Pattern Recognition* [C]. Anchorage, AK: IEEE, 2008. 1 – 8.
- [13] S Walk, N Majer, K Schindler, B Schiele. New features and insights for pedestrian detection [A]. *IEEE Conference on Computer Vision and Pattern Recognition* [C]. San Francisco, CA: IEEE, 2010. 1030 – 1037.
- [14] C Papageorgiou, T Poggio. A trainable system for object detection[J]. *International Journal of Computer Vision*, 2000, 38(1): 15 – 33.
- [15] X Su, W Lin, X Zheng, et al. A new local-main gradient orientation HOG and contour differences based algorithm for object classification[A]. *IEEE International Symposium on Circuits and Systems* [C]. Beijing: IEEE, 2013. 2892 – 2895.
- [16] 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, 37(12): 375 – 381.  
Y W Chong, H L Kuang, Q Q Li. Two-stage pedestrian detection based on multiple features and machine learning[J]. *Acta Automatica Sinica*, 2012, 37(12): 375 – 381. (in Chinese)
- [17] X Wang, T X Han, S Yan. An HOG-LBP human detector with partial occlusion handling [A]. *IEEE International Conference Computer Vision* [C]. Kyoto: IEEE, 2009. 32 – 39.
- [18] W R Schwartz, A Kembhavi, D Harwood, L S Davis. Human

- detection using partial least squares analysis[A]. IEEE International Conference on Computer Vision[C]. Kyoto: IEEE, 2009. 24 – 31
- [19] A Shashua, Y Gdalyahu, G Hayun. Pedestrian detection for driving assistance systems: Single-frame classification and system level performance[A]. IEEE Intelligent Vehicles Symposium[C]. Italy: IEEE, 2004. 1 – 6.
- [20] A Prioletti, A Møgelmoose, P Grisleri, M Trivedi, A Broggi, T Moeslund, Part-based pedestrian detection and feature-based tracking for driver assistance: real-time, robust algorithms, and evaluation[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(3): 1346 – 1359.
- [21] K Guto, K Kidono, Y Kimura, T Naito. Pedestrian Detection and Direction Estimation by Cascade Detector with Multi-classifiers Utilizing Feature Interaction Descriptor[A]. IEEE Intelligent Vehicles Symposium[C]. Germany: IEEE, 2011. 224 – 229.
- [22] Q Ye, J Liang, J Jiao. Pedestrian detection in video images via error correcting output code classification of manifold subclasses [A]. IEEE International Conference on Intelligent Transportation Systems [C]. Anchorage, AK, USA: IEEE, 2012. 193 – 202.
- [23] T Ojala, Multi resolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971 – 987.
- [24] X Geng, D C Zhan, Z H Zhou. Supervised nonlinear dimensionality reduction for visualization and classification[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2005, 35(6): 1098 – 1107.
- [25] J Tenenbaum, V D Silva, J Langford. A global geometric framework for nonlinear dimensionality reduction [J]. Science, 2000, 290(5500): 2319 – 2323.
- [26] A Ess, B Leibe, K Schindler, et al. A mobile vision system for robust multi-person tracking[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Anchorage, AK: IEEE, 2008. 1 – 8.
- [27] A Mogelmoose, A Prioletti, M Trivedi, et al. Two-stage part-based pedestrian detection[A]. IEEE International Conference on Intelligent Transportation Systems [C]. Anchorage, AK, USA: IEEE, 2012. 73 – 77.
- [28] I Alonso, D Llorca, M Sotelo, L Bergasa, P de Toro, J Nuevo, M Ocana, M Garrido, Combination of feature extraction methods for SVM pedestrian detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2007, 8(2): 292 – 307.
- [29] L Oliveira, U Nunes, P Peixoto, On exploration of classifier ensemble synergism in pedestrian detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 11(1): 16 – 27.

#### 作者简介



刘 威 男, 1975 年 6 月出生于沈阳. 博士, 东北大学副教授, 教授级高级工程师, 主要研究领域为汽车辅助驾驶、智能交通、智能安防, 研究方向为计算机视觉、图像处理、模式识别.

E-mail: lwei@neusoft.com



段成伟 男, 1987 年 4 月出生于河南. 硕士, 研究方向为计算机视觉、图像处理、机器学习.