

doi: 103969/j. issn. 0490-6756. 2016. 09. 011

一种改进的梯度直方图远红外行人检测方法

周红志¹, 刘 祺²

(1. 阜阳师范学院信息工程学院, 阜阳 236041; 2. 华南理工大学计算机科学与工程学院, 广州 510640)

摘 要: 在远红外行人检测领域, 实时性以及鲁棒性如何提高, 这一问题一直得不到有效的解决. 笔者以感兴趣区域 ROIs 为切入点, 在图像预处理阶段充分应用像素梯度相关知识, 总结出有效的垂直投影方法. 该方法原理为: 搜集图像梯度信息, 采用初定位方法对垂直方向上的条状区域展开处理并获取包括行人在内的图像, 采取局部双阈值图像分割算法获得相对较为可靠性 ROIs, 避免对整幅输入图像进行搜索. 检测过程中, 充分发挥全局信息作用, 采用以金字塔熵加权的梯度直方图方法, 针对远红外行人进行描述. 本方法引入了支持向量机算法的识别技术. 文章最后针对该方法进行了实验检验. 实验结果显示, 该方法大大缩短了 ROIs 提取过程的搜索耗时, 而且可以对一些以背景目标为主的候选区域进行抑制.

关键词: 行人检测; ROIs; 像素梯度; 双阈值; 梯度直方图; SVM 算法

中图分类号: TP393 **文献标识码:** A **文章编号:** 0490-6756(2016)05-1018-09

A far infrared detection recognition description method based on improved gradient histogram

ZHOU Hong-Zhi¹, LIU Qi²

(1. College of Information Engineering, Fuyang Teachers' College, Fuyang 236041, China;
2. School of Computer Science and Engineering, South China University of Technology, Guangzhou 51064, China)

Abstract: For the real-time and robustness of far-infrared pedestrian detection remains a very challenging problem, the paper presents a vertical gradient projection method based on the pixels from ROIs (Regions of Interest) extracted angle. This method utilizes the image gradient information to conduct early positioning for the vertical banded image area which may contain pedestrians, then forms relatively reliable ROIs by partial dual-threshold image segmentation algorithm inside obtained image. In the pedestrian detection process, it combines global information and gradient histogram method based on pyramid weighted entropy to describe the far-infrared pedestrian, at the same time assist in SVM algorithm to complete the identification process. Experiments show that the method can improve the search efficiency of ROIs extraction stage, and suppress portion of the candidate region included in the background target.

Key words: Pedestrian detection; ROIs; Pixel gradient; Dual Threshold; Gradient histogram; SVM algorithm

1 引 言

如今,越来越多的科学技术被应用到交通管理中,尤其是远红外行人检测技术,在这一领域中的应用前景十分广阔. 通过人们不断地研究,这种技术检测结果的准确性和效率有了很大的提升,不过同时也伴随着无法消除的迟钝以及高漏检率现象. 目前市场上常见的行人检测方法很多,不过这

些方法的流程都是相同的,需要首先进行感兴趣区域 ROIs(Regions of Interests)提取,然后才是行人识别,总共通过这两个环节达到检测目的.在前一环节内帧间差分法^[1]表现较好,该方法对紧挨着的不同帧的图片灰度参数大小展开差处理,完成这该步骤后对图片里面行人的移动区域展开分析.在移动式行人检测设备上比较常见,其主要缺陷在于识别准确性和行人移动速度有关.背景去减法^[2]以某一帧图像灰度为参照,对下一帧图像进行对比分析,确定移动区域,不过该方法在运动场景下无法使用,只可以安装在固定电子检测设备上使用.国内学者 Xu 完成图像拍摄后,紧接着运用直方图方法对其进行均衡化处理^[3],然后将灰度均值乘以某一系数,将得到的结果当做分割阈值. Nanda 及其研究小组通过人工方式,由图像内截取一千个行人样本^[4],以像素点的灰度均值为依据,从而获得分割阈值.这种方法的应用环境仅限于行人和背景间存在较大差异的情况,若拍摄到的图像中行人不够清晰的话,会给区域的分割带来极大困难. Curio 及其研究小组首先提取出行人的步行姿态特点,然后构建模型^[5],运用模型和行人的图像进行匹配,应用行人侧面信息完成识别,该方法的缺陷在于:当行人对着的摄像设备走过来时,识别准确率很低. Gavrilu 以及很多学者采用了构建模板的方法,主要依据为行人的不同形态特征^[6-8],然后将红外仪器拍摄的图像和模板进行对比,从而达到行人识别的目的.另外,还有一种模板是以行人轮廓为依据的^[9],原理上基本是一致的,不过这种模板匹配算法在应用中需要进行大量的计算,实用性因此受到限制.文献^[10]中的方法首先从头到脚搜集行人的信息,然后总结出其移动的特点,选择固定的周期,对这一行人在不同时刻的图像的相关性进行计算,这种方法避免了行人大量信息的搜集,不过需要某些信息达到很高的准确性水平,否则就会影响识别准确率.考虑到上述方法都有一定的局限性的,笔者在进行 ROIs 提取研究时,引入了垂直投影技术,从而大大提高了全局搜索的效率.同时,笔者在进行行人识别时,采用了以金字塔熵加权为基础的梯度直方图方法,并充分利用全局信息的作用,展开对远红外行人的描述,最后应用支持向量机 SVM(Support Vector Machines)算法进行有效的识别.对分类器进行离线训练时,将更多的关注点放在困难样本和迭代收集上,从而促进了分类器鲁棒性的提升.实验结果显示,这种方法在行人检

测方面表现出更好的实时性以及鲁棒性.

2 优化的 ROIs 提取方法

笔者提出的 ROIs 提取方法包括两个步骤,即预分割和 ROIs 产生.预分割是通过垂直投影方法所实现的,其原理在于首先总结行人在图像中灰度值的特征,然后运用这些特征对条形区域展开评测,最后剔除掉所有整体灰度比较一致的区域; ROIs 的形成,有赖于双阈值分割算法,通过这一方法对图像进行分割处理,从而获取 ROIs.

对曲线进行描述,从而得到所有列上像素梯度均处于某一值以上的像素数,通常来说,曲线中都存在一些有着下降沿或是上升沿(也就是转折点)的条形区域.在这些条形区域里面,转折点的存在表示像素数的亮度值发生了剧烈的变化.为了判断其是否位于局部边缘范围中,本文采用了梯度幅度分割阈值 T_s ,以该阈值 T_s 为标准,筛选出所有高于这一标准的点,然后对这些点进行登记处理.

利用梯度投影曲线的统计特征可以计算出 T_s ,具体计算公式如下.

$$T_s = \omega \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2} \quad (1)$$

$$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

其中, x_i 代表列数为 i 的序列上所有像素的数量; ω 代表 ROIs 的权重系数; ROI_s 代表输入图像列数.根据这一公式计算出来的 T_s 的值即为图 1(c) 中的水平虚线.

采用前文中提到方法步骤,等到搜索范围缩小至若干垂直方向上的条形区域后,就能够看到行人目标(如图 1(d)),图像中亮度比较高的地方是需要搜索的最后一块区域.通常来说,通过远红外设备获取到的背景信息,面积以及灰度同质度都处于较高水平.这一特点的存在,有效地让梯度信息更加明显,从而提高目标区域的确定效率.

和背景面积比较大的目标相比,绝大多数水平方向上的条形范围中都会有很多灰度不一致的背景,难以针对其归纳出某些灰度情况特点,这对于提取的准确性来说是很不利的,因此在进行预分割时,刻意忽视了以像素梯度为基础的水平方向上的投影.完成搜索后,下一步就需要使用到局部自适应双阈值分割算法^[11].此时针对以上成果的区域进行二值图像分割,从而获取 ROIs.

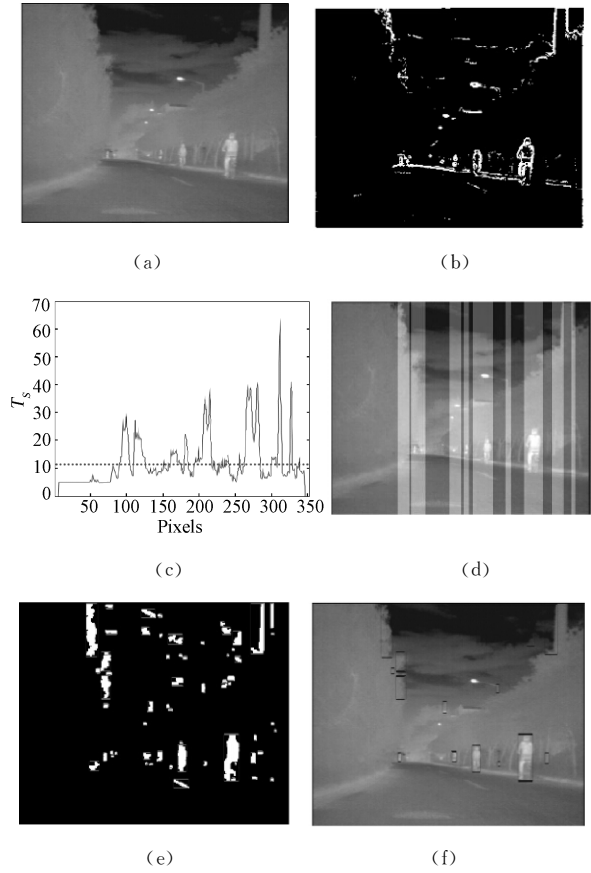


图 1 以某一区域为例,所进行的 ROIs 提取
Fig. 1 The extract of ROIs in a region as an example
(a) 原始远红外图像;(b) 二值梯度图像;(c) 梯度垂直投影曲线;(d) 通过本文设计的方法得到的需要进一步搜索的区域;(e) 采用二值分割对条形区域进行分割后得到的结果;(f) 最终得到的 ROIs

为了进一步消除所得二值图像中的噪声和填补部分弱连接区域,先后采用形态学腐蚀运算和膨胀运算处理二值图像中的连通区域,腐蚀运算采用结构元素为 1×3 的掩膜(mask),膨胀运算采用结构元素为 3×3 的掩膜.经过上述处理后的二值图像分割结果如图 1(e)所示,图中所有独立的 8 连通区域均采用粉色的最小外接矩形进行标识.

3 行人识别

在进行离线训练时,笔者引入了迭代训练机制,并应用 SVM 学习算法,从而有效地提高了行人分类器的准确性和效率.在进行检测时,为了能够对识别结果展开进一步的处理,降低误检发生的可能性,使用了多帧校验的方法,促进结果鲁棒性的显著提升.

3.1 金字塔熵加权方向梯度直方图

笔者对文献[12]中的熵加权特征提取法进行了改良,从而得到金字塔法(下称 HOG_PEW),该

方法能更有效地提取部分或全部与形状有关的数据信息,从而更好地对目标展开描述.即对输入模式进行细化处理,从而得到尺寸不断提高的不同层次的图像 o,一层比一层更加细化,最后得到数量为 2×2 的十分细致的图像片,过程图详见图 2(c)~(e).完成细化处理后,应用 HOG 方法总结并提取出其中的特点,从而对该层的局部结构进行表达.最后将不同层次的图相片的熵加权 HOG 特征结合在一起,即为 HOW_PEW 特征向量.

针对该描述模式而言,我们用 L 表示金字塔层数.那么在第 l 层的水平以及垂直两个方向上图相片数量分别为 2^{l-1} 以及 2^l ,那么该层总图相片数量为 $2^l \times 2^{l-1} = 2^{2l-1}$ 个.当然, l 的范围是位于 1 到 L 之间的正整数.运用 r 维特征分量对所有局部结构特征进行表达,此时能够得到区间数量 $bin = r$,在 1 图像块中图相片的数量为 2×2 个;采用第 l 层描述方式,水平方向上的图相片的数量只有 1 个,如图 2(c).因此在不考虑紧挨着的两个图像块存在的重叠区时,可以利用某 $r \times 2^{2l-1}$ 维特征分量对第 l 层进行表达,最终计算出维数为 $r \sum_{l=1}^L 2^{2l-1}$.

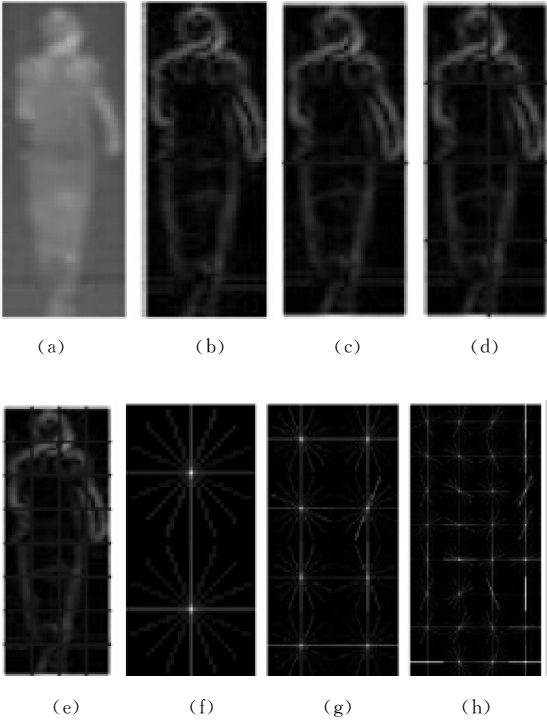


图 2 总结金字塔熵加权方向梯度直方图特征的过程
Fig. 2 Conclusion pyramid process of entropy weighted direction gradient histogram feature
(a) 待测行人;(b) 行人的梯度图像;(c)~(e) 应用图像空间金字塔描述方法得到的三层图像片划分;(f)~(h) 和(c)~(e)图中的图像片对应的方向梯度直方图特征情况

需要注意的是,和尺度空间中的表达存在不同,在对 HOG_PEW 特征进行求解时,不同层次彼此间不需要实施平滑处理,只需要根据原始图像来计算熵加权 HOG 特征,从而大大节省了计算过程的开销.以图 2 来看,若是 L 的取值为 3,那么拍摄像素就应该为 32×80 .

3.2 支持向量机算法分类器

在行人检测方面,SVM 算法表现出了良好的性能,其决策公式为

$$h(x) = \sum_{i=1}^v \alpha_i y_i K(x, x_i) + b \quad (3)$$

从文献[13-15]可知,直方图交叉核能够给出比较正确的两者统计特征的相似百分比,这是 SVM 算法的重要理论基础,也是该算法能够在这一领域普及开来的原因,表现出十分出色的效果.为了提高 ROI 判断结果的有效性,本文采用 HIKSVM 算法针对行人分类器展开训练,此时式(3)就能够转换为以下形式.

$$h(x) = \sum_{i=1}^v \alpha_i y_i \left(\sum_{j=1}^d \min(x_i(j), x(j)) \right) = \sum_{i=1}^v \left(\sum_{j=1}^d \alpha_i y_i \min(x_i(j), x(j)) \right) + b \quad (4)$$

本文引入图像空间金字塔的表示方法,同时提取输入模式的局部形状信息和全局结构信息,提出 HOG_PEW 特征描述远红外行人,这种描述识别方法虽然通过金字塔熵加权方向加以改进,但仍然属于直方图统计特征,因此引入用基于直方图交叉核(Histogram Intersection Kernel, HIK)的 SVM 分类器实现行人识别,直方图交叉核能有效度量两个直方图统计特征之间的相似度,因此基于直方图交叉核的 SVM 算法在目标检测与识别方面获得了较大的成功,根据文献[12]的描述,这种方法与业内常用的 FCSVM、s-SVM 相比具有明显的优势.

3.3 基于 Boot-strapping 算法的训练机制

根据上文分析可知,从本质上看,行人检测实际上就是对稀有事件的检测.所以,和不含行人的大面积区域相比,训练数据的搜集难度更低.通常情况下,从处于控制外的场景中采集能够对所有潜在代表性训练进行描述的数据是非常难的.Boot-strapping 算法的应用,能够通过迭代完成分类器在“复杂”训练数据位置上的集合,从而促进其准确性的大幅上升,从技术的角度使得背景目标数据的搜集和作用的实现成为现实.

要对过拟合这一实际情况进行稳妥处理,很多

专家提出采取提前终止策略^[16],其有效性早已在 BP 神经网络中就得到了证实.所以,本文也采取了提前终止策略以及 bootstrapping 方法,从而构建全新的迭代训练机制,运用这一新的机制对具有描述性作用的“复杂”训练数据进行表达,并在最大程度上利用其价值,从而提高行人分类器的准确性和高效性.另外,一般情况下,一项搜集工作是无法提供全部有价值的训练数据的,而本文所设计的训练机制却能做到,即代替人工完成大量训练方式的搜集.

4 实验分析

为了检验笔者所提出的新方法,本文从三个方面展开实验:(1) 测量 ROI 提取方法参数改变所带来的影响.对于以像素梯度为基础的垂直投影方法而言,当设施不同的参数时,分割结果也会随之发生变化,详见本文第 4.3 节;(2) 测量行人识别性能水平的高低.实验结果显示,以 HOG_PEW 特征为识别依据,比以往的 HOG_like 特征以及金字塔二进制模式特征的性能水平更高,详见本文第 4.4 节;(3) 了解行人检测方面的实际能力,检测对象为迭代训练机制如何对分类器的识别能力造成影响,对于实践应用来说能够达到期望标准,详见 4.5 节.

4.1 实验素材及实验平台

就目前的现状来看,能够应用的以远红外检测为基准的数据库只有 OCTBVS^[17],笔者在对行人分类器进行训练时,所采用的素材是本研究小组自行拍摄的一系列视频文件,这些视频文件中环境所包含的干扰性因素很多,而且这些因素在实际生活中都是存在的,比如天气、动态背景和交通工具等.视频中超过一半的图片均为利用笔者设计的 ROI 提取方法获得的,除此之外还应用了人工裁剪的方法.同时鉴于垂直投影技术是本文的创新点之一,为确保其能始终发挥作用,本文在设计远红外采集平台时,本文在轿车的车头安装了红外单目远红外传感器,即热像仪,其安装高度为 24 cm(以水平路面作为参考面).该摄像头的空间分辨率为 384 像素 $\times$ 288 像素、数据采集速率为 25 帧/s、焦距为 3 m、视场角为 $28^\circ \times 21^\circ$,在这种实验场景下,可以确保场景的路是垂直的.

另外,本文构建了四个不同的测试视频序列.然后将所有行人目标标识出来,为验证算法提供更多依据.为了完成对训练机制的验证,笔者准备了两个仅有背景模式的视频片段,从而提供“复杂”训练数据.而且,还需要准备两个视频序列,以完成扩展验证.表 1 列出了以上实验所需要的所有素材.

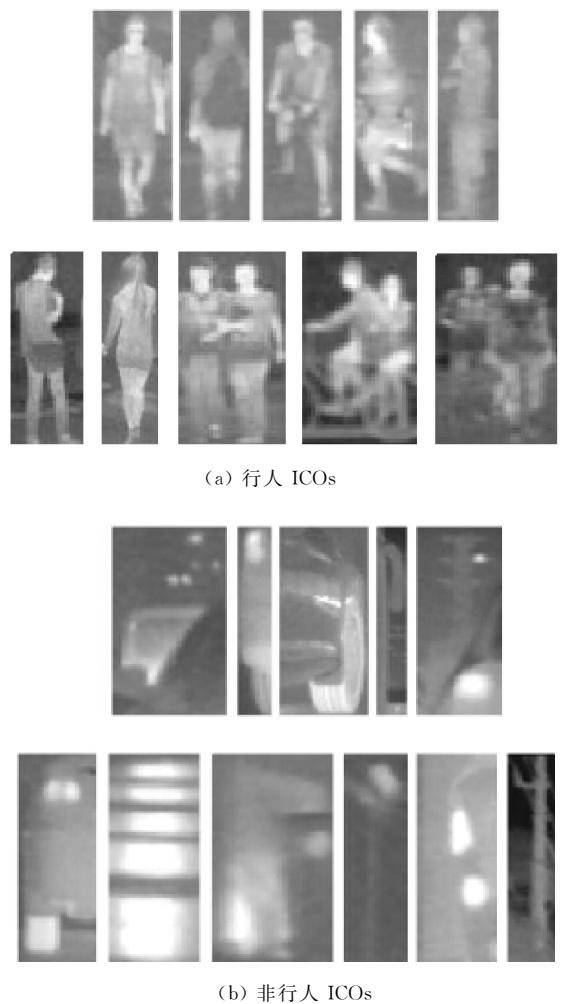


图 3 视频集中的部分样本示例

Fig. 3 Some samples of the concentrated sample video

表 1 仿真实验采用的视频序列数据

Tab. 1 Video sequence data used in simulation

视频序列名称	视频长度(帧)	行人数量	备注
SummerSeq01	1567	494	测试作用, 包括 Test_S08 及 Test_S09
SummerSeq02	1448	1143	测试作用, 包括 Test_U10 及 Test_U11
SummerSeq03	824	108	验证作用
Train_D1	—	3876	提供“多人粘连”样本数据
Train_D4	2239	—	提供“复杂”数据
WinterSeq01	2419	1139	测试作用, 包括 Test_S05、Test_U05 及 Test_U07
WinterSeq02	1553	204	测试作用, 包括 Test_U08 及 Test_U09
WinterSeq03	1000	966	验证作用
Train_D3	2208	—	提供“复杂”数据

后续的验证需要构建以下软硬件环境:CPU: Inter Pentium E5800(3. 220 GHz);RAM:2 GB;操作系统:Windows 7 Ultimate;软件:MATLAB.

4.2 性能评估标准

笔者在选择性能评估标准时,始终坚持实时性以及鲁棒性两大目的. 考虑到 ROIs 预分割是为了消除漏检现象的出现,同时尽量缩小搜索范围,因此确定了漏检率(Miss Rate,MR)以及垂直方向上条形区域的比例(Image Stripes Ratio,ISR)两项评估指标. MR 的计算公式为

$$MR = \frac{n_M}{n_A} \tag{5}$$

$$DR = \frac{n_D}{n_A} \tag{6}$$

其中, n_M 以及 n_A 代表区域以外以及被标注的行人的数量; n_D 代表可以被检测到的行人数量. MR 以及式(6)中的 DR 具有互补性,即 $MR = 1-DR$.

我们定义 ISR 为需要被搜索的垂直方向上的条形区域面积和图像总面积两者间的比值. MR 是检验本文所提出的方法的鲁棒性的最重要指标. ISR 则是检验本文提出方法的实时性的重要指标. 这是由于传统方法对系统运行时间进行检测,就不能忽视硬件平台的类型. 原因在于当检测方案或是硬件平台中的任何一项发生改变时,运行时间都会发生变化. 但对于实践而言,硬件平台是多种多样的,因此笔者创新性地引入 ISR 指标,以完成对实时性的检验.

根据文献[12]可知,我们假设识别某一行人必须以这一条件为前提的话:目标行人的轮廓面积和客观标准面积两者间的比值达到甚至高于 35%. 考虑到当平台不发生变化时,误检率以及检测率是非常重要的评价参数,不过在实际应用中,两项参数都实现最优化的可能几乎为零,因此笔者引入了另外一项指标,即加权准确率(Weighted Accuracy, WA),通过这一指标就可以了解训练机制对性能产生的影响,其计算公式为

$$WA = p \frac{n_D}{n_A} + (1 - p)(1 - \frac{n_{FP}}{n_w}) \tag{7}$$

其中, n_w 代表目标仅有背景的 ROIs 的数量; n_{FP} 代表误检数量; p 代表权重系数, $0 \leq p \leq 1$; WA 代表对误检率与检测率两者的权衡.

为了对本文所提出的方法的整体性能有更为直接的了解,笔者使用了目前应用非常广泛的用于评估行人检测性能的指标,即 True Positive Rate (TPR)、False Positive Rate(FPR),两者的计算公

式如式(8)、式(9)

$$TPR = \frac{\text{被正确判断的行人 } ICOs \text{ 数量}}{\text{行人 } ICOs \text{ 总数}} \quad (8)$$

$$FPR = \frac{\text{被错误判断的非行人 } ICOs \text{ 数量}}{\text{非行人 } ICOs \text{ 总数}} \quad (9)$$

4.3 预分割性能评估

我们为检验预分割准确程度,对两项重要指标 T_g 以及 T_s 进行完善,以表 1 为源,从中筛选了 971 帧图像,从而建立测试图像序列,其中有 759 位行人被标注出来。

根据式(1),能够计算出 T_s 的具体值,而 x_i 的值和 T_s 有关,因此我们假设 $\omega = 1$,选择不同 T_s 值,研究本文设计的方法对预分割准确程度的影响。依次选择 $\{12, 14, 16, 18, 20, 22, 24, 26, 28\}$ 作为 T_s 的值,预分割效果的变化情况见图 4(a)。

从实验结果可知,当 T_s 的值不太大时,MR 和其彼此间的关系比较疏远;当 T_s 不断上升,达到甚至超过 24 时,我们选择的方法会导致部分携带目标的垂直方向上的条形图像被剔除出去,从而产生漏检情况。原因在于,如果行人和摄像头之间的距离太远的话,其热辐射在传递的过程中,能量大幅降低,从而变成了模糊的目标。和近景相比,它们不但像素更少,而且和其周围背景之间的灰度差异很小。通常来说,它们的局部边缘只能够被表征成具有更少不明显灰度变化的像素。因此由于这些模糊目标的 T_s 处于较高水平,因此其识别难度也更高。从图 4(a)中可以看出, T_s 和 ISR 之间呈明显的负相关关系,再加上 T_s 对 MR 及 ISR 所产生的影响,不难得知:如果 T_s 比较保守的话,那么所有的行人目标都不会超过方法所得到的垂直方向上的条形区域,不过这会造成搜索目标区域范围过大;随着 T_s 的上升,这一区域范围会不断缩小,不过也会导致漏检率随之上升。根据前文分析的结果,我们假定 $T_s=20$,并针对这一值展开下一步的检验,此时方法的性能水平就比较理想,在彻底避免预分割出现漏检情况的前提下,将搜索目标区域面积控制在图像总面积的 47% 水平。

然后,我们对 T_s 发生改变时,方法对预分割产生的影响进行研究。从式(1)可知,当曲线的统计特征发生改变时, T_s 的值会保持不变,也就是说影响 T_s 值的量只有 ω 。考虑到 ω 的取值过大或过小都会造成统计特性发生变化,因此只能在比较大的区间范围内进行取值。笔者将 ω 的取值区间限定为 $[0.95, 1.05]$,部分结果详见图 4(b)。上述实验表

明,笔者所提出的方法的性能和 ω (或 T_s),特别是 MR 之间没有直接关联。不过,在 ω 持续上升的过程中,ISR 会随之降低。原因在于随着 ω 的提高, T_s 会有所上升,且二者变化幅度十分接近。从图 1(c)可以知道,二者的变化会给转折点的确定造成影响。也就是说本文实施的预分割方法具有一定的合理性,在后续的实验,假定 ω 的值为 1。

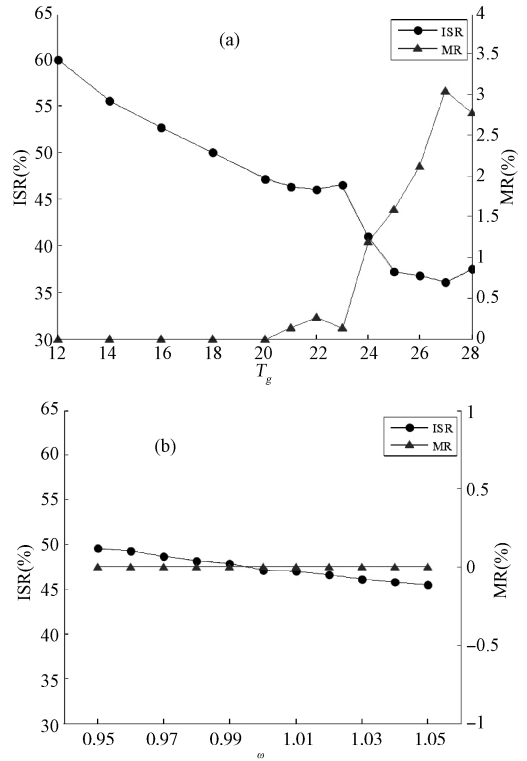


图 4 垂直投影方法的性能表现

Fig. 4 The performance of vertical projection methods
(a) 在 ω 取值为 1、 T_g 取值变化的情况下,性能和变化的走向;
(b) 在 T_g 取值为 20、 ω 取值变化的情况下,性能和变化的走向

4.4 行人识别性能比较

为了了解识别性能的真实水平,笔者从视频文件中选择行人和非行人 ICOs 训练样本集 779、1890 个以及测试样本集 1060、4067 个,从而得到两个不存在重叠的样本集,为分析不同行人识别方法的性能水平提供基础。

首先测试当层数 L 依次为 2、3、4 时, HOG_PEW 特征是对性能的影响。笔者引入了 10 重交叉验证方法。考虑到样本的选择都是随机的,为了降低随机选择对评估结果造成的影响,有必要重复展开十次的验证,最后计算出结果的标准差和均值,从而完成评估。当层数为 3 时,得到的验证结果为 $90.85\% \pm 7.06 \times 10^{-3}\%$, $93.67\% \pm 7.34 \times 10^{-3}\%$ 和 $95.48\% \pm 8.67 \times 10^{-3}\%$ 。研究结果表明,

HOG_PEW 特征的准确性和层数之间呈正相关关系,考虑到不同描述目标之间互补性的存在,因此可从不同被检测图像空间中提出得到. 不过,在这样的条件下,会出现拥有较高维度的 HOG_PEW 特征向量. 比如,在层数为 4 的情况下,可以得到维度在 5000 左右的 HOG_PEW 特征. 这会导致识别过程计算量上升,从而拉低方法的实时性. 为了达到计算量和准确性两者间的平衡,结合实验结果和应用经验,在实验中,假定层数 L 的值为 3.

接下来,我们针对 HOG_PEW 特征以及目前应用最为广泛的具有典型意义的 HOG_like 方法展开对比,具体比较对象有 HOG^[18]、PHOG、PBP^[19] 以及 HOG_EW^[12] 等不同的特征方法. 在对比过程中,保持分类器结构以及 SVM 算法不变,而且以视频文件中检测目标外接矩形的高度的最低值当做依据,对训练样本集进行细化处理,从而得到三个互相不影响训练子集:近景和近景目标的高度像素分别在 60 个以上以及 32 个以下,然后将其优化成 32×80 像素、 12×32 像素;剩余全部归为中景目标,并将其优化成 24×64 像素. 图 5 详细描绘了不同的 ROC 曲线.

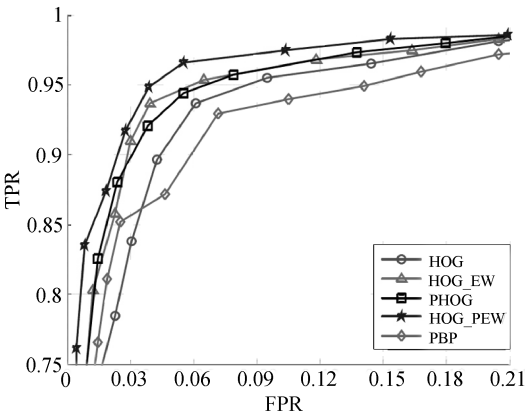


图 5 基于不同描述特征的行人识别性能比较
Fig. 5 The performance compare of pedestrian detection based on different describe characteristics

通常,考虑到 PBP 特征的用途是对幕布结构进行表达,所以应用这样的纹理特征对近景目标进行表达能够取得不错的效果. 但测试样本中有很多目标纹理较为欠缺,尤其是远景行人模式,在这一点上表现得更加突出. 从图 5 得知,PBP 拥有比 HOG_like 更高的准确性,但却比剩下的特征法更低. 和行业内应用最为普及 HOG 特征法相比,文献^[12]提到的 HOG_EW 提取方法以及本文所提出的 HOG_PEW 提取方法在识别性能方面更高.

即便是采用非行人模式的空间结构信息时,HOG_EW 以及 HOG_PEW 的性能也和 PHOG 基本持平. 本文引入了金字塔表示方法,从而有效地获取到不同图像空间下的 HOG_EW 特征全面表达的识别目标的空间结构数据,达到了更好的感知效果.

4.5 行人检测性能评估

在对行人检测方法的性能水平进行评估之前,必须首先对训练机制有更全面的了解. 从视频文件里面选择行人和非行人 ICOs1179 和 788 个,从而构建原始训练样本集,在该样本集范围内,应用 bootstrapping 法对具有代表性的非行人样本进行采集和处理. 从表 1 可知,表 1 中存在两个视频序列,即 Train_D4 和 Train_D3,这两个序列,其主要作用是提供前文所述的“复杂”样本,其余的序列可以用于衡量其工作的有效程度,图 6 为本次实验的实验结果. 在实施评估时,要使用前文中提到的位置约束信息,尽可能多地把非行人 ROIs 给剔除掉.

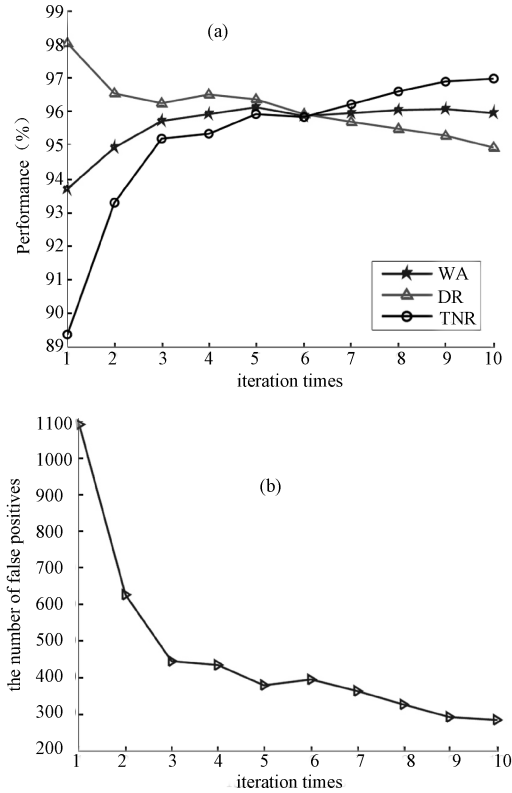


图 6 通过评估得到的训练机制的有效性水平
Fig. 6 The level of effectiveness on training mechanisms by assessing

(a) 迭代次数发生变化时,方法性能和变化的走向;
(b) 迭代次数发生变化时,误检数量和变化的走向

图 6(a)告诉我们,TNR(True Negative Rate)能够表示非行人 ROIs 召回率的实际水平. 考虑到

在进行离线训练时,行人样本集基本上不会发生变化,不过当 Bootstrapping 方法进行迭代时,都会产生一些“复杂”数据,从而对非行人样本集进行扩展.因此,特征空间经过迭代后,通过 SVM 算法所获得的最合理分类超平面和非行人样本存在空间之间有着一定差异,从而导致随着迭代次数增多,DR 会不断降低,详见图 6(a).不过,通常来说,DR 的降低幅度是能够接受的,而且这一过程有助于降低误检出现的可能性,详见图 6(b).通过对公式(7)的分析可知,WA 在此过程中会有所提高.

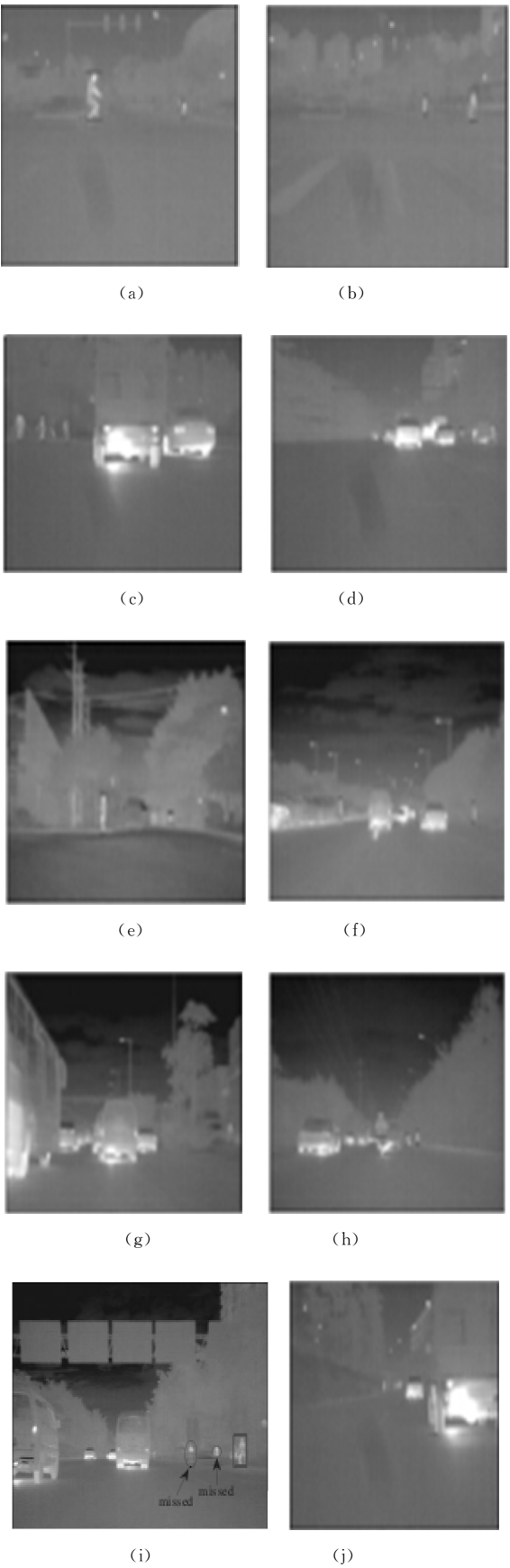
从图 6(a)可以得知,当 bootstrapping 迭代进入到第五轮时,WA 就表现出收敛的趋势,考虑到这一现象的存在,根据提前终止策略,训练到此结束.此时得到的描述非行人目标的样本集中 ICOs 的数量为 2014 个,运用这些 ICOs 就可以构建行人分类器,且准确性很有保障.结合夜间行人辅助驾驶系统的功能需求来看,对笔者所提出的方法的漏检数、误检数以及检测速率进行评价和测量,是很有实践意义的.表 2 中的数据就是改变测试序列时,笔者所提出的方法的性能指数.应用 MATLAB 进行仿真运算,执行速度基本上能够维持在 5 帧/s 左右.

表 2 系统在测试视频序列上的检测性能

Tab. 2 Detection performance on the test of video sequence

测试视频序列	视频长度(帧)	召回率*	漏检数量	误检数量	DR	FPPF
SummerSeq01	1567	487/494	7	40	0.986	0.026
SummerSeq02	1448	1112/1143	31	19	0.973	0.013
WinterSeq01	2419	1101/1139	38	45	0.967	0.019
WinterSeq02	1553	192/204	12	38	0.941	0.024
结果平均值					0.967	0.021

从图 7 可以得知,笔者所设计的方法能够有效地检测到各种各样的外部特征以及运动特征,尤其是当行人在穿过马路、或是在路边等待绿灯时,检测有效率更高.不过,该方法也没有彻底避免漏检和误检情况的出现.究其根源在于,红外摄像头分辨率水平较低,远景行人模式极有可能变得非常模糊,因此难以得到足量的信息依据进行判断,所以导致了出现了漏检的情况.图 7(i)即为漏检画面,以蓝色椭圆标出漏检对象,是由局部遮挡现象引起的,该行人的下半身受道路旁边的花台所遮挡.此外还存在误检现象,这主要是由于道路及两旁区域中常见的背景物,比如树木、车灯等的轮廓和行人轮廓较相近.图 7(j)~(l)列出了一些典型的误检画面.



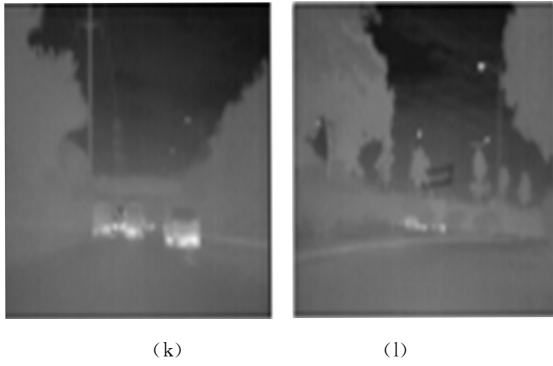


图 7 在测试视频序列上的部分远红外行人检测结果示例
Fig. 7 The part of sample results on testing Far-infrared pedestrian detection

5 结 论

本论文所设计的远红外检测行人方法是以像素梯度的垂直投影方法为基础,对 ROIs 方法进行改良后得到的方法。通过远红外设备获取到的环境图像中,天空、地面等背景在灰度方面表现得比较一致,这是本文方法能够实现的重要事实基础。该方法过程为:首先发挥图像梯度信息的作用,采用初定位的方法对垂直方向上的条状区域展开处理并获取包括行人在内的图像,运用局部双阈值图像分割算法获取具有一定可靠性的 ROIs。行人识别过程中,充分发挥全局信息的作用,采用以金字塔熵加权的梯度直方图方法,针对远红外行人展开描述。本方法引入了 SVM 算法进行行人识别。该方法在行人检测方面表现出良好的鲁棒性以及效率,在对行人目标进行搜索时,能够有效地抑制很多不包含行人目标的区域范围。

参考文献:

- [1] 叶谷平,王伟东.一种基于脉冲耦合神经网络的运动目标提取方法[J]. 四川大学学报:自然科学版, 2014, 51(1): 69.
- [2] 邱光能,全惠敏.基于对称差分 and 背景减的运动目标检测[J]. 计算机工程与应用, 2014, 50(13): 158.
- [3] Xu F L, Liu X, Fujimura K. Pedestrian detection and tracking with night vision [J]. IEEE Trans Intell Transp, 2005, 6(1): 63.
- [4] Nanda H, Davis L. Probabilistic template based pedestrian detection in infrared videos [C]//Proceedings of IEEE Intelligent Vehicle Symposium (IV). Versailles, France: IEEE, 2002.
- [5] Curio C, Edelbrunner J, Kalineke T, *et al.* Walking pedestrian recognition [J]. IEEE Trans Intell Transp, 2000, 1(3): 155.

- [6] Broggi A, Bertozzi M, Fascioli A, *et al.* Shape-based pedestrian detection [C]//Proceeding of IEEE Intelligent Vehicles Symposium (IV). Dearborn, MI: IEEE, 2000.
- [7] Gavrilă D M. Pedestrian detection from a moving vehicle [J]. LNCS, 2000, 1843: 37.
- [8] Gavrilă D M, Munder S. Multi-cue pedestrian detection and tracking from a moving vehicle [J]. Int J Comput Vision, 2007, 73(1): 41.
- [9] Gavrilă D M, Giebel J, Munder S. Vision-based pedestrian detection: the PROTECTOR system [C]// Proceeding of IEEE Intelligent Vehicles Symposium (IV). Parma, Italy: IEEE, 2004.
- [10] Cutler R, Davis L S. Robust real-time periodic motion detection, analysis and applications [J]. IEEE Tpmi, 2000, 22(8): 781.
- [11] Ge J F, Luo Y P, Gyomei T. Real-time pedestrian detection and tracking at nighttime for driver-assistance systems [J]. IEEE Trans Intell Transp, 2009, 10(2): 283.
- [12] Liu Q, Zhuang J J, Kong S F. Detection of pedestrians for far-infrared automotive night vision systems using learning-based method and head validation[J]. Meas Sci Technol, 2013, 24(7): 74022.
- [13] Su S Z, Chen S Y. Analysis of feature fusion based on HIK SVM and its application for pedestrian detection [J]. Abstr Appl Anal, 2013(3): 112.
- [14] 孙锐,侯能干,陈军.基于特征融合和交叉核 SVM 的快速行人检测方法[J]. 光电工程, 2014, 41(2): 53.
- [15] Wu J X. A fast dual method for HIK SVM learning [J]. LNCS, 2010, 6312: 552.
- [16] 吴银花,金龙旭,张宁.针对 H. 264 改进的快速整像素运动估计算法[J]. 光学精密工程, 2013, 21(4): 1017.
- [17] Davis J W, Keck M A. A two-stage template approach to person detection in thermal imagery[C]// Proceedings of IEEE Workshops on Application of Computer Vision (WACV). Breckenridge, CO, USA: IEEE, 2005.
- [18] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Diego, CA, USA: IEEE, 2005.
- [19] Sun H, Wang C, Wang B L, *et al.* Pyramid binary pattern features for real-time pedestrian detection from infrared videos [J]. Neurocomputing, 2011, 74(5): 797.