

车载热成像行人检测 RoI 提取方法

刘 琼, 罗 晴, 彭绍武
(华南理工大学 软件学院, 广东 广州 510000)

摘 要: 热成像适合低照度行人检测, 车载热成像背景灰度分布变化大, 行人目标易与背景干扰物混淆, 捕捉远距离行人困难, 以双阈值分割方法提取 RoI(region of interest)很难满足系统召回率和虚警率要求. 针对车载热成像行人检测, 提出新的 RoI 提取方法, 包括图像预处理、RoI 提取和后处理. 设计膨胀最大值滤波器进行图像增强; 利用 Haar-like 特征来改进自适应双阈值分割法进行 RoI 提取, 设计增量计算模型以提高计算效率; 设计目标时序特性和空间对称性过滤器来排除虚假 RoI. 与基准方法相比, 当虚警率不高于 40 时, 本文方法提高召回率 49%, 且 RoI 提取速率不低于 18 帧/s.

关 键 词: RoI 提取; 膨胀最大值滤波器; Haar-like 特征; 双阈值分割; 车载热成像行人检测
中图分类号: TP 311.5 **文献标志码:** A **文章编号:** 1005-3026(2020)08-1083-08

RoI Extraction for Vehicular Thermal Infrared Pedestrian Detection

LIU Qiong, LUO Qing, PENG Shao-wu

(School of Software Engineering, South China University of Technology, Guangzhou 510000, China.
Corresponding author: LIU Qiong, E-mail: liuqiong@scut.edu.cn)

Abstract: Thermal infrared images are suitable for pedestrian detection in low illumination such as at night. The grayscale distribution of background in an infrared image vehicle-mounted varies obviously and pedestrian is easily confused with background interference and it is difficult to catch a pedestrian in the distance. The system recall rate and false alarm rate requirements can't be achieved through extracting RoI with double threshold segmentation method. We construct a new RoI extraction method consisting of image preprocessing, RoI generation and RoI post processing etc. An expanding maximum filter is designed to enhance image contrast. Adaptive double threshold segmentation is improved by Haar-like feature. Computing efficiency is raised by designing incremental model. Besides, filters considering gray-scale temporal feature and spatial symmetry feature of a pedestrian are presented to remove false RoI. Comparing with the benchmark method, our method improves recall rate by 49% when the number of false RoIs is less than 40 per frame. And RoI extraction speed isn't lower than 18 frames per second.

Key words: RoI extraction; expansion maximum filter; Haar-like feature; double threshold segmentation; vehicular infrared pedestrian detection

车载热成像行人检测是高级驾驶辅助系统(advanced driver assistance systems, ADAS)的重要组成部分, 感兴趣区域(region of interest, RoI)提取是行人检测系统的重要环节, RoI 提取品质直接影响系统的检测率和虚警率. 由于行人目标极易与背景干扰物混淆, 远距离行人难识别, 以及背景灰度分布变化较大, 提高车载热成像行人检测的检测率, 降低虚警率, 具有极大挑战性.

滑窗法以多尺度穷尽方式提取 RoI^[1]可以获得高召回率; 但是非行人 RoI 过多, 分类阶段虚警率居高不下, 且其计算开销限制了实际应用. Ge 等^[2]线性扫描双阈值 RoI 提取方法直接面向车载热成像行人检测, 计算开销较低, 能较好解决简单场景的图像目标分割碎裂问题; 但是由于热成像图像品质通常不够稳定, 道路场景复杂, 其分割效果呈现多样性, 实用性有限. Liu 等^[3]对自适应双

阈值分割方法进行改进,采用像素梯度竖直投影预定位热成像中的潜在行人区域,再针对区域进行双阈值分割提取 RoI,以减小阈值分割的搜索范围,大幅降低计算开销.石祥滨等^[4]提出基于十字窗口灰度方差的阈值分割算法获取 RoI,尽量保留竖直方向的灰度信息,但是当行人目标的灰度分布不够均匀时,这种阈值计算方法极易产生分割碎裂. Kim 等^[5]提出合并分割子区域的 RoI 提取方法,利于连通热成像中分割碎裂的行人目标.

多种阈值分割方法致力于解决分割碎裂、背景干扰物引起的分割黏连或计算效率问题,但是往往顾此失彼,综合效果不尽如人意,很难满足实际需求;而且,对于多尺度行人目标,负样本类内差异较大. RoI 提取效率低下、检测率较低,以及虚警率较高等问题始终未能得到较好解决.

本文 RoI 提取框架包括图像预处理、RoI 提取和 RoI 后处理三个模块,创新特点如下:

- 1) 膨胀最大值滤波器. 根据热成像特点,应用膨胀滤波算法扩展行人 RoI,凸显灰度图像的对比度,提高行人目标的显著性.
- 2) Haar-like 增强自适应双阈值分割算法. 利用 Haar-like 特征调整双分割的低阈值,促成行人目标与背景干扰物较好分离;设计增量计算,实现多尺度目标分割,提高效率.
- 3) RoI 级联过滤器. 基于目标灰度时序特性和空间对称特性构建级联过滤器,滤除 RoI 中可能隐含的单干扰物虚警或组合虚警.
- 4) 集成测试. 构建 RoI 提取器,设计实验,在 SCUT 数据集上验证本文方法的效果和效率.

1 阈值分割方法

阈值分割方法即利用阈值划分灰度图像前景与背景的方法;对于灰度分布不均匀的热成像,一般采用自适应阈值,像素阈值取决于其邻域的灰度分布. 在热成像车载行人检测领域,经典自适应双阈值分割算法^[2]的作者通过实验研究发现,如果行人比背景亮度更高,即便行人亮度不一致,从水平扫描的角度行人亮度也将一致高于背景.

1.1 自适应双阈值分割算法的阈值计算

自适应双阈值分割算法设定图像的背景像素灰度最大阈值 $T_H(x,y)$ 和目标像素灰度最小阈值 $T_L(x,y)$,且 $T_L(x,y) < T_H(x,y)$. 从上至下逐行逐像素扫描当前图像:灰度高于 $T_H(x,y)$ 的像素置为前景,灰度低于 $T_L(x,y)$ 的像素置为背景;否则当

前像素灰度等同水平方向前一像素灰度. $T_L(x,y)$ 和 $T_H(x,y)$ 分别由式(1)和式(2)计算^[2]:

$$T_L(x,y) = \frac{\sum_{j=y-w}^{y+w} I(x,j)}{2w+1} + \alpha, \tag{1}$$

$$T_H(x,y) = T_L(x,y) + \beta. \tag{2}$$

式中:对于图像 $I, I(x,y)$ 代表位于 (x,y) 的像素灰度值; w 为与行人宽度相关的参数; α 和 β 均为偏移量,通过实验获取. 计算 $T_L(x,y)$ 的关键参数 w ,根据具体实例优化获取.

文献[2]为了在行人目标和背景灰度变化情况下获得较好的热成像分割效果,通过实验推导并根据 $T_L(x,y)$ 进行分段计算,求出最终 $T_H(x,y)$,如式(3)所示. 本文后续采用式(3)计算 $T_H(x,y)$.

$$\left. \begin{aligned} T_H(i,j) &= \max\{T_1(i,j), T_L(i,j)\}, \\ T_1(i,j) &= \min\{T_2(i,j), 230\}, \\ T_2(i,j) &= \min\{T_3(i,j), T_L(i,j) + 8\}, \\ T_3(i,j) &= \max\{1.06 \times (T_L(i,j) - \alpha), T_L(i,j) + 2\}. \end{aligned} \right\} \tag{3}$$

1.2 双阈值分割区域

对输入热成像进行全图分割,计算开销仍然较大,不利于嵌入式应用. 为此,提出基于图像梯度的垂直投影法^[1]:计算所有像素的梯度,在图像竖直方向上将具有较高梯度幅值的像素进行投影,获得梯度投影曲线;扫描该曲线的上升沿和下降沿,判断竖直方向上灰度值发生显著变化的转折点;通过匹配邻近上升沿和下降沿转折点的“点对”,从输入图像中提取竖直带状区域,作为初定位的潜在目标区域,以避免分割算法对输入图像进行全局搜索. 本文后续所述阈值分割界定于这种带状区域.

1.3 阈值分割算法的缺陷

为了避免热成像行人目标亮度不一致时被分割为多个碎片,自适应双阈值分割算法采用水平线性扫描法计算 $T_L(x,y)$. 但是,在复杂场景中,行人目标亮度分布通常不一致,水平线性扫描方式仅关注水平邻域的像素灰度分布,面对与行人目标处于同一竖直方向(横向位置相同)、背景亮度较高的干扰物体,可能产生行人目标与背景干扰物黏连的分割结果.

如图 1 所示,针对原始图像(第一行),采用自适应双阈值分割算法获得对应的二值图像(第二行). 由图可见,人眼完全能够区分背景与行人目标,而二值图的行人 RoI 多半与背景干扰物(如树干、路灯等)黏连,造成 RoI 尺度过大且偏

离真实行人目标的外接矩形,以致后续分类器执行失误.

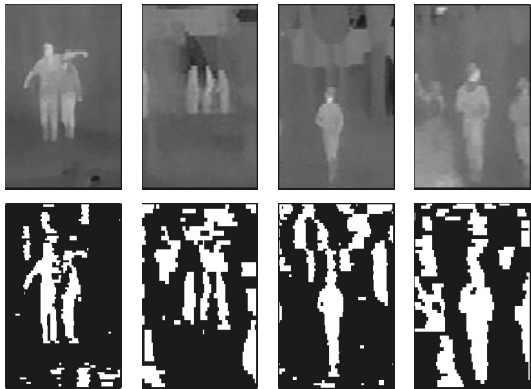


图 1 原始图与对应的二值化图
Fig. 1 Original image and binary image

2 RoI 提取方法

RoI 提取方法包括:①预处理模块——设计膨胀最大值滤波器,以增强图像对比度,提高行人目标的显著性;②RoI 提取模块——采用 Haar - like 特征辅助修订自适应双阈值分割的 $T_L(x,y)$ 计算,使提取的 RoI 与背景干扰物有效分离;设计增量方式实现多尺度目标分割计算,提高 RoI 提取效率;③后处理模块——基于目标灰度的时序特性和空间对称特性设计过滤器,排除虚假 RoI.

2.1 图像预处理

利用卷积神经网络通过扩张卷积核提高输出感受野^[6-8]的技术方式设计膨胀最大值滤波器:逐像素遍历输入图像,择出当前像素膨胀邻域中亮度最高的像素灰度值,并以此替代当前像素灰度值,增强当前像素显著性.对于图像的亮目标区域,膨胀操作前后的像素灰度差异普遍较明显,经膨胀滤波不仅目标亮度增强,而且目标区域也得到扩展;而对于图像的背景区域,膨胀操作前后的像素灰度差异较小,经膨胀滤波后背景区域灰度变化并不明显.以下给出膨胀最大值滤波器的定义及算法.

定义 1 设输入图像为 I ,位置处于 (x,y) 的当前像素灰度值记为 $I(x,y)$. 定义膨胀最大值滤波器如图 2 所示,位置处于 (x,y) 的当前像素以图 2 中黑色方块表示,其 8 邻域像素为图 2 中的阴影方格. 图像 I 经膨胀最大值滤波器处理得到图像 P ,则 $P(x,y) = \max \{ I(x,y), I(x-h,y-h), I(x-h,y), I(x-h,y+h), I(x,y+h), I(x+h,y), I(x+h,y-h), I(x,y-h) \}$, 其中, h 为膨胀因子,默认值设置为 3. 图像 I 四周

边缘像素的处理:依次在 I 的左、右、上或下边缘,逐像素向左、右、上或下填充 h 个灰度取值为 0 的像素.

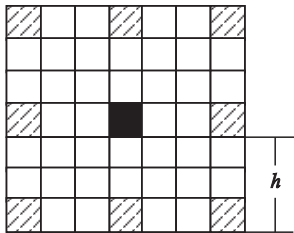


图 2 膨胀最大值滤波器
Fig. 2 Expanded maximum filter

车载热成像行人检测系统的远距离行人目标与传感器的距离定义为 85 m,远距离行人目标外接矩形尺度为 15×24 像素;近距离行人目标与传感器的距离定义为 20 m,行人目标外接矩形尺度为 90×140 像素. 计算结果表明,经膨胀最大值滤波器操作,远距离行人目标尺度扩展至 18×27 像素,是原尺度的 1.35 倍;近距离行人目标尺度扩展至 93×143 像素,是原尺度的 1.06 倍. 由此可见,膨胀最大值滤波器对远距离目标的扩展效果更显著,更符合实际需求. 膨胀最大值滤波器的处理效果如图 3 所示.

图 3 中自左至右共五组对比热成像行人目标的膨胀滤波操作效果,每组两列,左列为操作前,右列为操作后.

2.2 RoI 提取

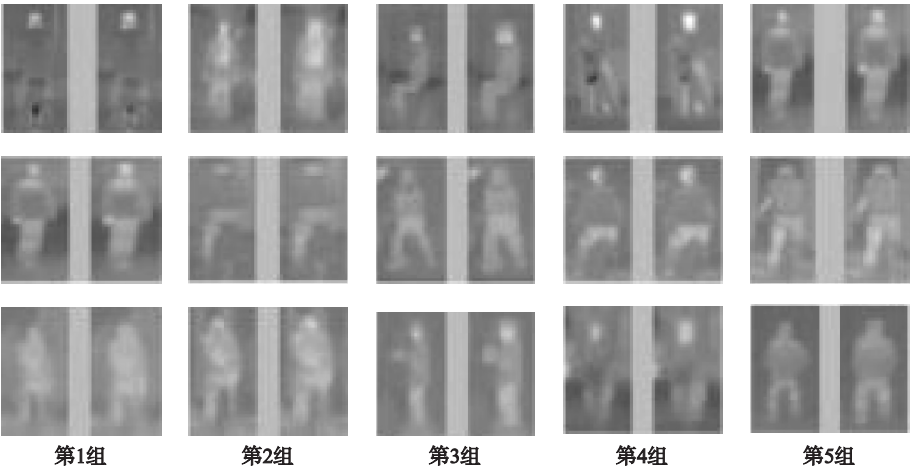
RoI 提取模块包括两个创新点:①Haar - like 增强自适应双阈值分割——借助 Haar - like 特征修订分割阈值计算;②高效率多尺度目标分割——采用增量计算进行多尺度目标分割,以单尺度目标 RoI 提取计算开销,实现多尺度目标 RoI 提取.

2.2.1 Haar - like 增强分割阈值计算

Liu 等^[3]的自适应双阈值分割 RoI 提取方法直接针对车载热成像行人检测,不仅有效解决因分割造成的目标碎裂问题,而且计算开销较低,适合资源有限的嵌入式应用. 但这种方法可能造成行人目标与背景干扰物体之间产生分割黏连.

实验分析表明,在这类分割黏连中,前景目标与背景干扰物的连接点大多出现在疑似行人目标的头部. 借助 Haar - like 特征调整自适应双阈值的分割阈值计算,可以同时兼顾水平和竖直方向像素邻域的灰度分布信息,排除行人目标与背景干扰物可能形成的黏连,具体算法描述如下.

经典自适应双阈值分割方法的 $T_L(x,y)$ 和



每组左右两列分别是处理前后的目标

图 3 膨胀滤波处理前后的行人目标

Fig. 3 pedestrian before and after expanded filter process

$T_H(x,y)$ 计算见式(1)和式(2). 本文提出的阈值计算方法描述如下,以原始输入图像提取的带状区域作为以下分割算法的输入图像 $I^{[4]}$,按照式(4)逐像素计算栅格区域 Haar-like 算子 $H(x,y)$,则 Haar-like 增强自适应双阈值分割算法的 $T_L(x,y)$ 计算见式(5).

热成像行人目标通常比背景干扰物更亮,围绕行人目标 RoI 可能的头部边缘区域像素点来计算 Haar-like 算子,特征值均大于 0,则低阈值 $T_L(x,y)$ 被提高,促成头部边缘区域的像素点归类为背景,从而使行人目标 RoI 与背景干扰物分离.

$$H(x,y) = \frac{\sum_{i=x+1}^{x+k} I(i,y) - \sum_{i=x-k}^{x-1} I(i,y)}{2k}, \quad (4)$$

$$T_L(x,y) = \frac{\sum_{j=y-w}^{y+w} I(x,j)}{2w+1} + H(x,y) + \alpha. \quad (5)$$

式中: $I(x,y)$ 为位置 (x,y) 的像素灰度值; k 取 $w/3$,与行人目标宽度相关的参数 w 默认取值为 12.

2.2.2 Haar-like 增强自适应双阈值分割法的 RoI 提取

在车载热成像行人检测应用场景中,行人目标通常分远、中和近三种尺度,如果针对不同尺度目标计算阈值,计算性能较差. 设计十字模板扫描图像,以增量方式实现多尺度目标灰度统计值的“或”运算和 Haar-like 算子的计算,以此降低计算开销. 十字模板如图 4 所示.

图 4 中下标 s,m 和 l 分别表示小尺度、中尺度和大尺度; w,b 分别表示目标水平邻域宽度的 $1/2$ 和目标竖直邻域高度的 $1/2$; Δ 表示增量.

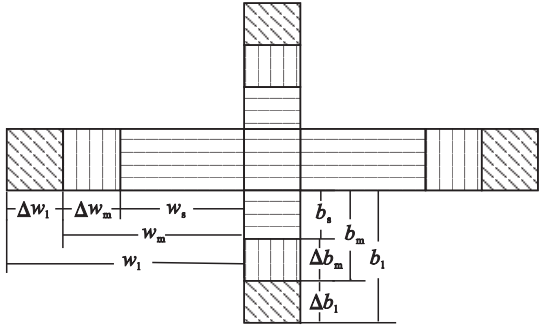


图 4 十字模板

Fig. 4 Cross template

为了进行阈值计算,采用模板逐行逐像素扫描输入图像,建立并实现关系为“或”的两类增量计算. 其一,当前像素属于初始像素(在图像的第 1 列或第 1 行),扫描统计执行尺度级别的增量计算:先计算小尺度目标的扫描统计值,则中尺度目标的扫描统计值等于小尺度目标的统计值加上中尺度增量;大尺度目标的扫描统计值等于中尺度目标的统计值加上大尺度增量. 其二,当前像素属于非初始像素,扫描统计执行像素级别的增量计算:当前像素的扫描统计值等于前一像素的统计值加上当前像素增量. 对于输入图像 I ,执行以下计算.

1) 阈值计算:

情形 1 像素 (x,y) 位于图像第 1 列($y=1$)或第 1 行($x=1$). 小尺度(远距离)目标邻域像素灰度统计值 $\text{sum}_s(x,y)$ 和竖直邻域像素灰度值之差 $h_s(x,y)$ 按式(6)计算.

$$\left. \begin{aligned} \text{sum}_s(x, y) &= \sum_{j=y-w_s}^{y+w_s} I(x, j), \\ h_s(x, y) &= \sum_{i=x+1}^{x+b_s} I(i, y) - \sum_{j=x-b_s}^{x-1} I(i, y). \end{aligned} \right\} \quad (6)$$

中尺度目标邻域像素灰度值之和 $\text{sum}_m(x, y)$, 竖直邻域像素灰度值之差 $h_m(x, y)$, 如式(7)所示.

$$\left. \begin{aligned} \text{sum}_m(x, y) &= \text{sum}_s(x, y) + \\ &\quad \sum_{j=y+w_s+1}^{y+w_m} I(x, j) + \sum_{j=y-w_s-1}^{y-w_m} I(x, j), \\ h_m(x, y) &= h_s(x, y) + \\ &\quad \sum_{i=x+b_s+1}^{x+b_m} I(i, y) - \sum_{i=x-b_s-1}^{x-b_m} I(i, y). \end{aligned} \right\} \quad (7)$$

式(7)所包含的两个式子分别表达在小尺度目标的基础上增加像素灰度值增量和 Haar-like 增量.

大尺度目标邻域像素灰度值之和 $\text{sum}_l(x, y)$, 竖直邻域像素灰度值之差 $h_l(x, y)$, 如式(8)所示.

$$\left. \begin{aligned} \text{sum}_l(x, y) &= \text{sum}_m(x, y) + \\ &\quad \sum_{j=y+w_m+1}^{y+w_l} I(x, j) + \sum_{j=y-w_m-1}^{y-w_l} I(x, j), \\ h_l(x, y) &= h_m(x, y) + \\ &\quad \sum_{i=x+b_m+1}^{x+b_l} I(i, y) - \sum_{i=x-b_m-1}^{x-b_l} I(i, y). \end{aligned} \right\} \quad (8)$$

式(8)所包含的两个式子分别表达在中尺度目标的基础上增加像素灰度值增量和 Haar-like 增量.

情形2 像素 (x, y) 不在图像左上角, 即 $y > 1, x > 1$, 则执行像素级别的增量计算. $d_k(x, y)$ 表示 k 尺度位置 (x, y) 像素的分割响应值; $\text{sum}_k(x, y)$ 表示 k 尺度目标邻域像素灰度统计值; $h_k(x, y)$ 表示 k 尺度目标邻域像素灰度值之差; $T_L^k(x, y)$ 表示 k 尺度分割低阈值, $T_H^k(x, y)$ 表示 k 尺度分割高阈值; $k = s, m, l$ 分别表示小尺度、中尺度和大尺度目标. 在十字模板扫描图像的过程中, 当前像素作为模板中心像素, 但凡逢图像边缘时, 伸出图像边缘的模板像素值均取“0”值, 如图4所示.

根据水平方向前一像素统计值 $\text{sum}_k(x-1, y)$ 和相应目标尺度像素灰度值增量计算当前像素灰度统计值 $\text{sum}_k(x, y)$, 基于竖直方向前一像素统计的邻域灰度值之差 $h_k(x, y-1)$ 和相应目标尺度像素灰度值增量计算当前像素的和, 以及当前像素的竖直邻域像素灰度值之差 $h_k(x, y)$:

$$\left. \begin{aligned} \text{sum}_k(x, y) &= \text{sum}_k(x, y-1) - \\ &\quad I(x, y-1-w_k) + I(x, y+w_k), \\ h_k(x, y) &= h_k(x-1, y) + I(x-1-b_k, y) - \\ &\quad I(x-1, y) - I(x, y) + I(x+b_k, y); \\ k &= s, m, l. \end{aligned} \right\} \quad (9)$$

2) 针对小尺度、中尺度和大尺度目标, 分别计算改进的双阈值分割算法的阈值 $T_L^k(x, y)$ 和 $T_H^k(x, y)$:

$$T_L^k(x, y) = \frac{\text{sum}_k(x, y)}{2w_k + 1} + \frac{h_k(x, y)}{b_k} + \alpha; k = s, m, l. \quad (10)$$

$$\left. \begin{aligned} T_1(x, y) &= \max \{ 1.06 \times (T_L^k(x, y) - \alpha), \\ &\quad T_L^k(x, y) + 2 \}, \\ T_2(x, y) &= \min \{ T_1(x, y), T_L^k(x, y) + 8 \}, \\ T_3(x, y) &= \min \{ T_2(x, y), 230 \}, \\ T_H^k(x, y) &= \max \{ T_3(x, y), T_L^k(x, y) \}; \\ k &= s, m, l. \end{aligned} \right\} \quad (11)$$

3) 针对小尺度、中尺度和大尺度目标逐像素求得 (x, y) 分割响应值 $d_k(x, y)$:

$$d_k(x, y) = \begin{cases} 1, & I(x, y) > T_H^k(x, y); \\ 0, & I(x, y) < T_L^k(x, y); \\ I(x, y-1), & \text{其他}. \end{cases} \quad (12)$$

$k = s, m, l.$

2.3 RoI 后处理

RoI 后处理包含两个级联过滤器: 基于目标时序灰度变化特性的过滤器, 用于滤除单干扰物虚警; 基于目标空间灰度对称特性的过滤器, 用于滤除组合虚警.

2.3.1 时序灰度不变性过滤器

在短时间内单干扰物虚警的灰度分布基本不变, 而行人目标的局部灰度分布变化较大. 设计“RoI 下肢跨度”模型, 并基于该模型构造时序灰度不变性过滤器.

1) RoI 下肢跨度模型:

行人目标与单干扰物虚警的局部灰度分布变化差异性较大, 若将整个 RoI 用于灰度分布计算, 非运动区域不仅耗费计算资源, 还降低真伪行人目标之间的可区分度. 选取反映 RoI 动态特性的下肢, 建立“下肢跨度”模型表示 RoI 的下肢高亮区域两侧边缘的距离, 记作 d . 算法流程见图5. 图5a表示单个 RoI, 纵向三等分 RoI, 自底截取 RoI

的 $1/3$ 区域作为下肢,其高度为 H ,宽度为 W ,见图 5b. 按式 (13) 计算 RoI 下肢的竖直灰度投影 $V(x)$,见图 5c. 按式 (14) 计算竖直灰度投影的一阶差值,见图 5d. 一阶差值曲线的最高波峰取作 RoI 下肢左边缘位置,一阶差值曲线的最低波谷取作 RoI 下肢右边缘位置,左、右边缘位置横坐标之差的绝对值即为 RoI 的下肢跨度 d ,见图 5e.

$$V(x)=\sum_{y=0}^{H-1}f(x,y),\tag{13}$$

$$V'(x)=\frac{dV(x)}{dx}=\frac{V(x+1)-V(x)}{(x+1)-x}=\frac{V(x+1)-V(x)}{1}.\tag{14}$$

2) 时序灰度不变性过滤器:

依据 RoI 动态特性,建立基于“下肢跨度”模

型的 RoI 过滤器,滤除与行人目标轮廓极为相似的单干扰物. 过滤器算法描述如下.

步骤 1 下肢跨度指标设为 ch^* . 在 SCUT Dataset 中选 S0 ~ S10 作为基准数据集,统计数据集中单行人目标下肢跨度总变化量,进而观察不同单行人目标下肢跨度总变化量的分布区间,截取 95% 置信度区间下界截点作为下肢跨度判别指标 ch^* .

步骤 2 标记 RoI. 对当前图像的 RoI 逐个标记下肢跨度 d 、下肢跨度变化量 Δd 和下肢跨度总变化量 ch . Δd 表示相邻两帧 RoI 下肢跨度的变化量, ch 表示 RoI 经历首帧至当前帧的 Δd 累计总和.

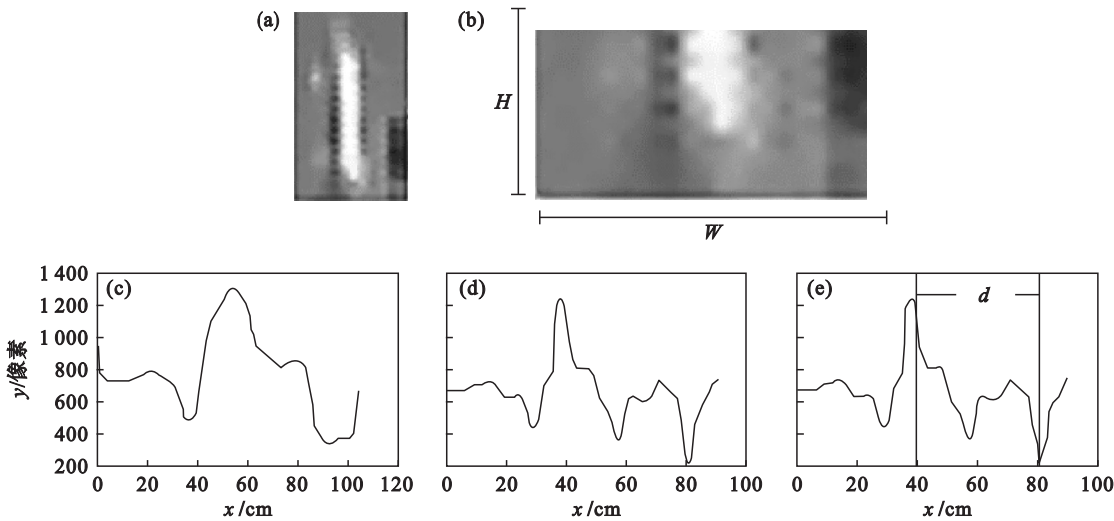


图 5 下肢跨度 d 计算流程

Fig. 5 Calculating the span d of a pedestrian limbs

(a)—RoI; (b)— $1/3$ RoI (下肢); (c)—RoI 下肢竖直灰度投影; (d)—竖直投影一阶差值曲线; (e)—下肢跨度 d .

3) 计算下肢跨度总变化量 ch :

①按 1) 所述计算 d ;

②计算 Δd :逐一匹配当前帧 RoI 与前一帧 RoI (逐个 RoI 匹配),若前一帧 RoI 的高、宽不小于当前帧 RoI 的高、宽的 $1/2$,且前一帧 RoI 的横、纵坐标与当前帧 RoI 的横、纵坐标差值的绝对值分别小于当前帧 RoI 高、宽的 $1/2$,则当前帧的当前 RoI 与前一帧对应 RoI 的 d 值之差的绝对值即为当前 RoI 的 Δd ;

③当前帧当前 RoI 的 Δd 与前一帧相应 RoI 的 ch 之和为当前 RoI 的 ch .

对于未能匹配前一帧 RoI 的当前帧当前 RoI (首次出现的 RoI),则 $\Delta d = ch = 0$,首次匹配前一帧 RoI 的当前帧 RoI, $ch = \Delta d$,否则 $ch \geq \Delta d$;逐个 RoI 计算 $ch - ch^*$,当 $ch - ch^* < 0$ (不具备动态性),滤除该 RoI.

2.3.2 空间灰度对称性过滤器

道路行人目标通常呈步行或骑行状态,普遍具有对称的性质,而组合干扰物的灰度分布复杂,基本不具备对称性. 本文通过设计最大池化特征来构造空间灰度对称性过滤器.

1) 最大池化特征:

池化特征的概念来自人们可以用图像一个区域上的某个特征的平均值(或最大值)代表这个区域的特性. 道路行人姿势变化多样,为了鲁棒度量 RoI 的对称性,设计 RoI 的最大池化特征(包括符号及其含义),如图 6 所示. RoI 最大池化特征 f_R 算法如下.

步骤 1 纵向五等分 RoI,从上至下取 $1/5$, $2/5$ 和 $2/5$ 区域分别定义为当前 RoI 的头部和上、下躯干.

步骤 2 按部位求取 RoI 的部件池化特征:

- ①基于外接矩形水平对称轴划分当前部件为大小相等的矩形 A 和 A' ;
- ②分别在 A 和 A' 内部各取 3 个随机子块,子块的长、宽大于所属矩形长、宽的 $1/2$ 但小于长、宽,3 个子块两两部分重叠但不完全重叠;
- ③计算各子块灰度均值;
- ④在 A 和 A' 内分别选取最大均值.

步骤 3 按部件计算 A 和 A' 区域间最大灰度均值之差并取绝对值 $f(A,A')$ 作为 RoI 部位池化特征:

$$f(A,A') = \left| \max_{i=1,2,3} \left(\frac{S_{Ai}}{N_{Ai}} \right) - \max_{i=1,2,3} \left(\frac{S_{A'i}}{N_{A'i}} \right) \right|. \quad (15)$$

式中: S_{Ai} 和 $S_{A'i}$ 分别表示图 6 所示子块的特征,如直方图特征和局部均值特征; N_{Ai} 和 $N_{A'i}$ 代表子块包含的像素数.

步骤 4 从三部件池化特征中选取最大值作为 RoI 最大池化特征 f_R .

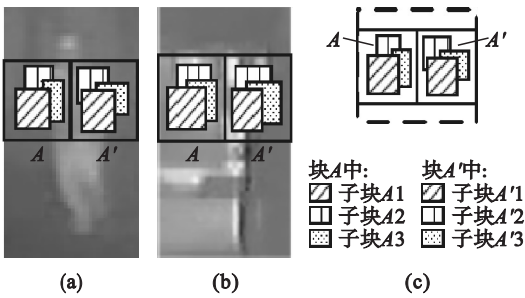


图 6 RoI 部件池化特征
Fig. 6 Component pooling features of RoI

(a)—行人对称块; (b)—虚警对称块; (c)—对称块说明.

2) 空间灰度对称性过滤器:

根据 RoI 空间对称特性,建立基于最大池化特征的 RoI 过滤器,滤除与行人目标轮廓极其相似的组合干扰物.过滤器算法描述如下:

①计算对称性判别指标 mf^* . 在 SCUT Dataset 中选 S0 ~ S10 作为基准数据集,统计数据集中单行人目标最大池化特征,进而观察不同单

行人目标最大池化特征的分布区间,并取 95% 置信度区间的上界截点作为判别指标 mf^* .

②计算 RoI 的最大池化特征 mf .

③逐个 RoI 计算 $mf - mf^*$.

当 $(mf - mf^*) > 0$ 时,RoI 不具备对称性,滤除 RoI.

3 实 验

为了评价本文的方法,针对图像增强、RoI 提取、RoI 过滤等单个功能模块和新的 RoI 提取框架,分别设计专门的实验验证方案.实验在 Pentium IV 3.0 GHz 台式计算机上执行.

3.1 实验方案

基准方案的选取:所有工作针对实际应用展开,选择双阈值分割 RoI 提取方法^[3]作为基准方案并进行复现.设计 4 个验证方案:方案 1,验证膨胀最大值滤波器,RoI 提取方案为“图像预处理模块 + 基准方案”;方案 2,验证 RoI 后处理模块,RoI 提取方案为“基准方案 + RoI 后处理模块”,RoI 后处理模块即时序灰度不变性过滤器和空间灰度对称性过滤器构成的级联过滤器;方案 3,验证新 RoI 提取模块,即 Haar-like 增强双阈值分割方案;方案 4,验证 RoI 提取集成方法,方案为“图像预处理模块 + Haar-like 增强双阈值分割 + RoI 后处理模块”.4 个验证方案均与基准方案进行比较.

3.2 评价标准

相比其他车载热成像行人检测公开数据集(如知名的 KAIST Dataset)和基准方案的测试数据(自行采集的车载热成像数据),SCUT 数据集^[8]具有图像帧数和标注数量较大、标注类型和道路场景类型多样化等优势,实验结果更具有说服力.以下实验选 SCUT 测试集 S11 ~ S20 用于测试,测试集相关信息如表 1 所示.行人标签的具体释义参考 Xu 等对 SCUT 数据集的详细阐述^[9].

表 1 测试集信息
Table 1 Information of test dataset

行人标签	带有标注的图像帧数	闭合框数量	遮挡闭合框数量	行人数量	平均每个人出现的帧数
步行	43 421	90 526	26 185	1 523	59.44
骑行	43 153	86 201	8 689	1 017	84.76

评价指标:统计测试视频标注的行人数量、虚警数量和总帧数,计算系统召回率 (Recall) 和每帧虚警数 (false positive per frame, FPPF),用曲线直观地表现评价结果.在 25 帧/s 的视频中出现

1 s 及以上的行人目标认作有效行人,因此,每隔 25 帧计算一次评价指标;而且仅当检测框与标注框至少重叠 0.5 时,才认为检出行人 RoI.

3.3 结果分析

基准方案和 4 个验证方案的实验测试结果如图 7 所示,可以看出:FPPF 相同的条件下,方案 1 与方案 3 召回率相当,但都显著高于基准方案的召回率,验证了预处理和 Haar – like 改进双阈值分割(基准方法)的有效性;在相同召回率条件下,本文所提的 4 个方案的 FPPF 明显低于基准方案,验证了虚警过滤功能的有效性.

重要实验结果见表 2,当 FPPF 为 50 时,基准方案召回率为 74%,方案 1 召回率为 89%,方案 2 为 77%,方案 3 为 88%,方案 4 为 91%. 当 FPPF 为 40 时,基准方案召回率为 59%,方案 1 为 69%,方案 2 为 71%,方案 3 为 68%,方案 4 为 88%. 当 FPPF 为 50 时,方案 4 (RoI 提取集成方案) 召回率达 91%,较基准方案提高 23%;当 FPPF 为 40 时,方案 4 召回率达 88%,较基准方案提高 49%,方案 4 召回率显著高于基准方案.

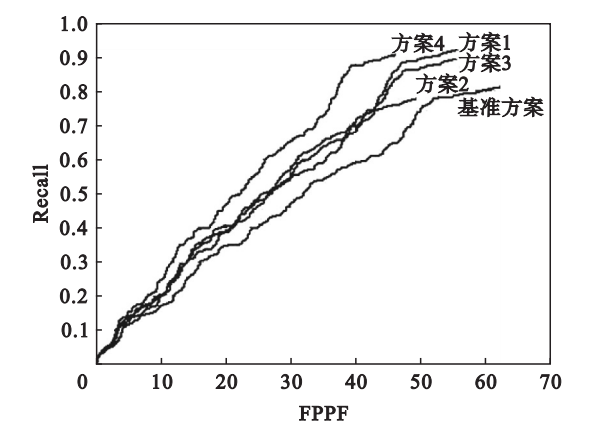


图 7 实验测试结果
Fig. 7 Test results

表 2 实验结果(召回率)					
条件	基准方案	方案 1	方案 2	方案 3	方案 4
FPPF = 50	74	89	77	88	91
FPPF = 40	59	69	71	68	88

为了证明本文方法的实用性,进一步比较方案 4 和基准方案的帧计算开销,如表 3 所示. 取 576 × 720 的热成像灰度图像,提取远、中和近三种尺度行人的 RoI,基准方案总耗时 90 ms,处理帧率为 11 帧/s;而方案 4,膨胀最大值滤波预处理模块耗时 8 ms,Haar – like 增强双阈值分割耗时 35 ms,级联过滤模块耗时 8 ms,总耗时 51 ms,

处理帧率约 18 帧/s. 方案 4 计算效率显著高于基准方案.

表 3 计算开销
Table 3 Computing overhead

方案	基准方案	方案 4
帧计算开销/ms	90	51

4 结 语

本文针对车载热成像行人检测的 RoI 提取方法,提出了新的 RoI 提取框架. 通过实验证明了新方法的有效性,并对计算开销进行比较,证明了新方法的实用性. 但是基于时序信息的 RoI 过滤方法易受车速影响,空间灰度对称性过滤器也可能存在片面性. 车载行人检测系统出现虚警极易对驾驶员造成干扰,过多虚警将影响系统的可用性;而实际道路场景复杂多变,虚警模式很难预测,如何提出更有效的方法防范虚警值得进一步探讨.

参考文献:

[1] Cheng M M, Zhang Z, Lin W Y, et al. BING: binarized normed gradients for objectness estimation at 300fps [J]. *Computational Visual Media*, 2019, 5(1): 4 – 21.

[2] Ge J, Luo Y, Tei G. Real-time pedestrian detection and tracking at nighttime for driver-assistance systems [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2009, 10(2): 283 – 298.

[3] Liu Q, Zhuang J J, Ma J. Robust and fast pedestrian detection method for far-infrared automotive driving assistance systems [J]. *Infrared Physics & Technology*, 2013, 60: 288 – 299.

[4] 石祥滨, 刘晓娜, 刘芳. 一种双阈值红外行人分割方法 [J]. *计算机工程*, 2012, 38(12): 5 – 8.
(Shi Xiang-bin, Liu Xiao-na, Liu Fang. Dual-threshold segmentation method for infrared pedestrian [J]. *Computer Engineering*, 2012, 38(12): 5 – 8.)

[5] Kim D S, Lee K H. Segment-based region of interest generation for pedestrian detection in far-infrared images [J]. *Infrared Physics & Technology*, 2013, 61: 120 – 128.

[6] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation [J/OL]. [2019 – 06 – 25]. <https://arxiv.org/pdf/1706.05587.pdf>.

[7] Yu F, Koltun V, Funkhouser T. Dilated residual networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 472 – 480.

[8] Vision Lab of SCUT. SCUT FIR Pedestrian Dataset [EB/OL]. [2019 – 07 – 05]. http://www2.scut.edu.cn/cv/scut_fir_pedestrian_dataset/main.psp.

[9] Xu Z W, Zhuang J J, Liu Q, et al. Benchmarking a large-scale FIR dataset for on-road pedestrian detection [J]. *Infrared Physical & Technology*, 2019, 96: 199 – 208.