



改进的多光谱聚合通道行人检测

彭志蓉, 赵美蓉*, 杨伟明, 郑叶龙

天津大学精密仪器与光电子工程学院, 精密测试技术及仪器国家重点实验室, 天津 300072



摘要: 针对外部环境的多变性和复杂性导致的单一波段下行人检测准确率较低的问题, 提出了一种改进的基于可见和红外双波段聚合通道特征的行人检测算法。分别提取可见图像与红外图像的聚合通道特征; 通过改变像素对比规则, 采用自适应的阈值进行比较, 将得到的改进的中心对称的局部二值模式特征添加到特征通道中; 针对多光谱聚合通道特征设计了不同滤波器组进行滤波; 训练分类器, 实现多光谱下行人检测。实验表明, 改进的局部二值模式特征能更好地描述红外图像中行人的对称性, 中间滤波层丰富了候选特征池, 算法在多种场景均能有效检测出行人, 提高了行人检测精度, 与利用多光谱聚合积分通道的检测工作相比, 平均漏检率有所降低。

关键词: 行人检测; 多光谱; 局部二值模式; 聚合通道; 滤波器组

中图分类号: TP751

文献标志码: A

Improved multispectral aggregate channel for pedestrian detection

Zhirong Peng, Meirong Zhao*, Weiming Yang and Yelong Zheng

State Key Laboratory of Precision Measuring Technology and Instruments, School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China

Abstract: To solve the problem of low pedestrian detection accuracy in a single band due to variability and complexity of external environment, an improved pedestrian detection algorithm based on multispectral aggregate channel feature is proposed. The aggregate channel features of visible images and infrared images are extracted, respectively. The pixel contrast rule is changed and the results are compared with the adaptive threshold. The improved central symmetric local binary pattern feature is added to the feature channels. Different filter banks are designed to filter the multispectral aggregate channel features. The classifier is trained to realize the multispectral pedestrian detection. Experiments show that the improved local binary pattern feature can describe the symmetry of pedestrians of infrared images better and the intermediate filter layer enriches the candidate feature pool. The algorithm can effectively detect pedestrians in various scenes and improve the pedestrian detection accuracy. Compared with the previous multispectral aggregate channel detection work, the algorithm reduces the log-average miss rate.

Keywords: pedestrian detection; multispectral; local binary pattern; aggregate channel; filter banks

DOI: 10.3969/j.issn.1003-501X.2017.09.004

Citation: *Opto-Elec Eng*, 2017, **44**(9): 882-887

1 引言

行人检测是计算机视觉领域的一个极具挑战却又

非常重要的问题, 在视频监控、智能交通、机器人及人机交互等领域应用广泛, 一直是研究的热点^[1]。过去对行人检测的研究多集中于单一可见图像或者单一红外图像。但是可见波段对于光照变化敏感, 在夜间、昏暗或雾霾等环境表现欠佳, 红外波段又缺乏细节信息, 对于热辐射差异较小或其他热源干扰情况下检测

收稿日期: 2017-05-22; 收到修改稿日期: 2017-07-02

*E-mail: meirongzhao@tju.edu.cn

效果较差^[2]。因此,可以通过融合可见与红外的多光谱图像对信息实现行人检测以提高检测精度。一种较好的多光谱图像对采集方法是在图像采集系统中加入分束器,使得入射光的红外波段透射进入红外相机,可见波段反射被可见相机接收;然后通过相机标定,调整相机位置,从采集系统硬件配置上实现图像对的配准。通常可见图像的空间分辨率高于红外图像,为了得到大小相同的图像对,一般采用牺牲可见图像的边缘部分来配准。配准的多光谱图像对视野大小相同,既有可见图像的细节信息,又有红外热辐射信息,即使在较为恶劣的条件下也能较好地刻画行人的特征。

对于静态场景下的多光谱行人检测, Davis 等^[3]利用背景减法分别对可见和红外图像提取行人轮廓,融合得到行人的位置; Castillo 等^[4]采用 INT³-Horus 框架对双波段图像自适应分割,将分别得到的感兴趣区域(region of interest, ROI)进行融合。但是这些方法不适合辅助系统或自动驾驶等运动场景,应用受限。对于动态场景下的行人检测, Yuan 等^[5]在可见和近红外多光谱图像对上分别提取梯度直方图(histograms of gradients, HOG)特征,利用潜变量支持向量机(latent variable support vector machine, L-SVM)对各波段图像分别进行训练和检测,最后将各自的检测结果加权来判断是否为行人。这种方法忽略了图像的低级特征,对于多光谱信息的融合位于分类阶段,图像对的信息利用明显不足。Hwang 等^[6]建立了可见与远红外配准的多光谱图像对数据库 KAIST,利用多光谱的聚合通道特征^[7](aggregate channel feature, ACF)进行训练和检测,证实了配准的多光谱图像对在解决复杂多变条件下的行人检测问题上具有重大意义。但是该方法并没有区分对待可见图像和红外图像,没有探讨红外图像较于可见的优势和区别。

通过对聚合通道特征进行分析,本文提出了针对恶劣环境改进的基于多光谱聚合通道特征(modified multispectral aggregate channel feature, MMACF)的行人检测算法,通过融合多光谱信息来补偿单一波段下行人特征的缺失,克服了原有的仅使用可见波段的聚合通道特征,在昏暗、模糊、雾霾等复杂环境下的行人检测所遇到的问题。算法着重分析红外图像,提出了具有自适应阈值的改进的中心对称局部二值模式(modified center symmetric local binary pattern, MCS-LBP)特征,并将其添加到多光谱的聚合通道特征中;在融合双波段的通道特征后,为了学习局部特征并丰富候选特征池,提出在聚合多光谱通道特征基础上加入滤波器组。在 KAIST 数据库上进行实验,实验结果表明,该算法即使在较为复杂恶劣的条件下也能够较好地实现行人检测,具有较高的检测精度。

2 算法实现

首先,MMACF 分别提取多光谱图像对中不同的通道特征信息;其次,利用中间滤波层学习聚合通道特征中的局部细节信息;最后采用 Real Adaboost 进行分类器的训练。算法流程如图 1 所示,分为三个部分:特征提取、滤波与分类器训练。具体描述如下。

1) 特征提取:分别从可见波段提取 LUV 色彩通道(3个)、梯度幅值(1个)、HOG(6个)通道特征,将这 10 个通道(LUV+M+HOG)整体记为可见的 ACF,从红外波段提取 LUV 空间亮度分量、HOG(9个)和 MCS-LBP 通道,分别记作 T、THOG 和 TMCS-LBP。

2) 滤波:参考卷积神经网络中卷积核的设计,通过引入合适的滤波器组对上述提取得到的通道特征进行滤波,可以获取图像更为充分的特征信息,使得分类器的学习更为准确。

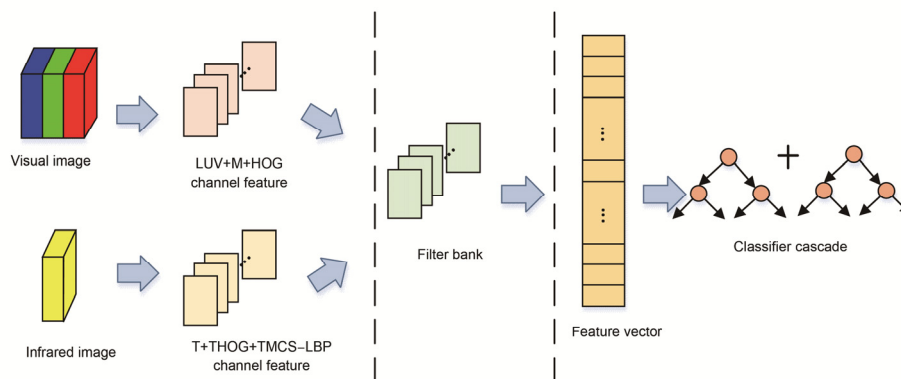


图 1 改进的多光谱聚合通道特征的行人检测算法流程。

Fig. 1 Flow chart of improved multispectral aggregate channel feature for pedestrian detection.

3) 分类器训练 :Real Adaboost 将弱分类器输出结果推广到连续置信度输出,因此拥有更小的分类误差和更快的收敛速度。通过 Real Adaboost 算法进行训练得到弱分类器,然后将多个弱分类器级联整合为强分类器。

2.1 改进的中心对称 LBP 特征

局部二值模式(local binary pattern, LBP)是一种计算简单、灰度和旋转不变的有效纹理描述子。LBP 算子没有关注图像中心像素邻域的梯度变化,使得特征对噪声较为敏感。Heikkila 等^[8]改变像素灰度值对比规则,对处于中心对称的像素进行比较,由此得到了中心对称的局部二值模式(center symmetric local binary pattern, CS-LBP)。CS-LBP 不仅将 LBP 的 256 种模式减少到 16 种,并且可以更好地捕捉梯度信息。两种形式 LBP 计算公式为

$$L_{LBP_{R,N}}(x,y) = \sum_{i=0}^{N-1} s(n_i - n_c)2^i, \quad (1)$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & \text{otherwise} \end{cases};$$

$$L_{CS-LBP_{R,N,T}}(x,y) = \sum_{i=0}^{(N/2)-1} s(n_i - n_{i+(N/2)})2^i, \quad (2)$$

$$s(x) = \begin{cases} 1, & x \geq T \\ 0, & \text{otherwise} \end{cases}.$$

式中: R 为计算区域的半径, N 为圆内采样点的个数, n_c 为中心像素灰度值, n_i 和 $n_{i+(N/2)}$ 为采样点的灰度值, T 为阈值,用来限制中心对称的像素对之间的差别。

行人的轮廓特征可以作为人体识别的依据,行人的头肩轮廓相对稳定^[9],在红外图像中虽然细节缺失,但人体轮廓的对称特征明显,并且纹理信息与色彩信息可以互补^[10],因此提出通过比较垂直对称的像素对的灰度差异来获得局部二值特征,即 MCS-LBP 特征。MCS-LBP 特征相对于 LBP 特征来说不仅速度加快,计算也更为简单。另外, MCS-LBP 比 CS-LBP 能更好地描述红外图像中行人明显的对称特征。上述各种 LBP

在 8 邻域情况下的编码方式如图 2 所示。

为了增强局部区域的对比度,在提取 MCS-LBP 特征前,需要先进行图像的归一化预处理。计算公式为

$$n'_i = (n_i - n_{\min}) / (n_{\max} - n_{\min}), \quad (3)$$

式中: n_i 和 n'_i 分别表示处理前后的像素点的灰度值, n_{\max} 和 n_{\min} 分别为图像的最大灰度值和最小灰度值。为了更好地获得图像局部纹理,编码计算中采用具有较好稳健性的自适应阈值,将邻域内像素点平均值与常数 λ 的乘积作为阈值。取值方法为

$$T = \lambda \times \frac{1}{N} \sum_{i=0}^{N-1} n_i, \quad (4)$$

式中各变量含义与(1)式和(2)式相同。

2.2 滤波器组设计

为了能在局部更好地描述各通道的特征,提出在 ACF 框架下引入滤波器组。引入滤波层,将通道特征与滤波器组乘积得到的特征矢量构成候选特征池,最后利用 Real Adaboost 训练弱分类器级联得到最终用于检测的强分类器。滤波器组的设计类似于卷积神经网络中的卷积核。在卷积神经网络中,卷积核与图像卷积可以学习得到图像的各种特征,其中卷积核的数值由多层神经网络训练得到。在 ACF 框架下,可以采用手动设计的滤波器组学习图像的特征。本文对双波段图像均利用滤波器组来丰富和学习特征。

局部去相关通道特征(local decorrelation channel features, LDCF)^[11]是利用每个特征通道进行主成分分析(principal component analysis, PCA)得到本征值构成滤波器,同时去除通道特征中数据的局部相关。随机滤波器(random filters)是在给定滤波器数目和最大滤波器大小的情况下,对滤波器大小按照均匀分布采样,并对滤波器中元素在[-1 1]之间按照均匀分布取值。棋盘格滤波器(checkerboard filters)^[12]是包含了均匀的正方形、水平垂直梯度检测模式和可能棋盘格模式的所有集合,不依靠先验知识得到的滤波器。

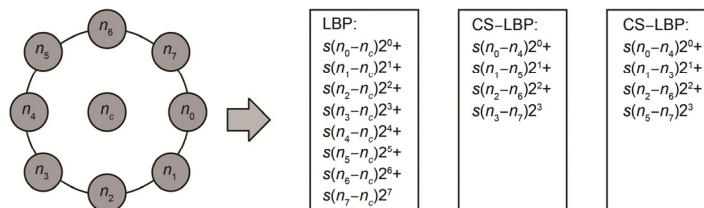


图 2 8 邻域下的 LBP 和 CS-LBP 的编码方式。

Fig. 2 LBP and CS-LBP features for a neighborhood of eight pixels.

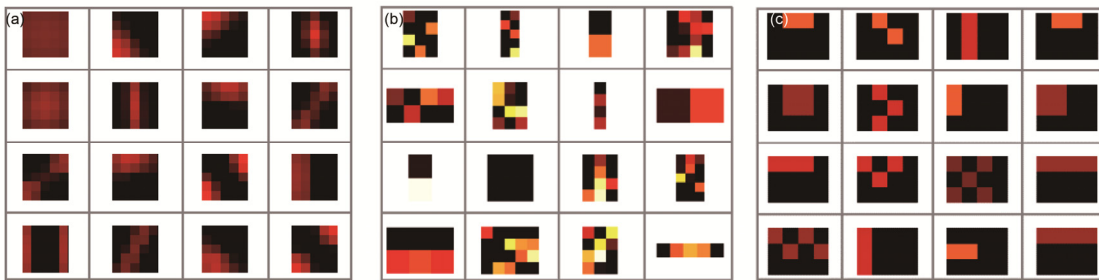


图3 不同滤波器组形式. (a) LDCF 滤波器组. (b) 随机滤波器组. (c) 棋盘格滤波器组.

Fig. 3 Illustration of different filter banks. (a) LDCF filter bank. (b) Random filter bank. (c) Checkerboard filter bank.

图3利用hot标准颜色映象,给出了各种手动设计的滤波器模式,hot颜色映象中颜色按照从黑到红到黄到白的顺序(黑色为0,白色为1)。利用上面的滤波器组对双波段图像进行运算,学习局部特征,构成候选特征,最后送入分类器训练。

2.3 分类器训练

分类器的训练是在ACF框架下,采用Real Adaboost算法训练并级联构成强分类器。ACF框架通过对输入图像的各种线性和非线性的变换,利用积分图快速计算出图像的特征,然后对这些特征进行筛选和训练得到分类器。这种方法参数设置少,并且融合了各种特征的信息,尤其与级联分类器结合具有很快的检测效果^[13]。提升算法中最常用的是Discrete Adaboost,它的每一个弱分类器输出结果只有1或者-1(二分类),而Real Adaboost是前者的改进,将其从离散二值判定规则推广到具有连续置信度输出,即输出扩展到了实数域。Real Adaboost具有更小的分类误差以及更快的收敛速度,所以本文选择该方法进行训练。

Real Adaboost算法会从聚合的多光谱特征矢量中选出典型特征,将其转换为对应的弱分类器,并将这些弱分类器通过一定的权重组合得到强分类器。多级强分类器级联得到最终的分类器。因此训练中分类器的设计需要考虑分类器级数、每级弱分类器的个数、决策树的深度等参数。分类器级数表示训练迭代次数,即强分类器的个数。强分类器个数越多,训练次数越多,难例样本的学习会更为充分。以上三个参数越大,行人检测的误检率越低,但是检测速率也会有所下降。

3 实验结果与分析

实验基于i7 CPU, 16 G内存的主机进行。选择KAIST数据库进行实验。KAIST数据库是由Hwang

等^[6]建立的配准多光谱图像对,该数据库中包含了白天和夜间图像昏暗、行人距离较远、图像对比度差等多种场景图像,共有95328对配准图像。实验使用简化后的数据库,由原数据库每隔20帧提取得到,其中2500对多光谱图像用于训练,2252对多光谱图像用于测试。为了评估算法性能,采用Dollar等^[14]提出的评价准则,利用平均漏检率(log-average miss rate, MR)与每幅图的误检数目(false positive per image, FPPI)曲线进行评价,以0.1 FPPI的MR值作为评价标准。图4给出了上述优化参数下各算法的检测结果。

图4中ACF(Vis)表示仅利用可见波段进行训练得到的分类器进行检测的结果,ACF(Inf)为仅用红外波段图像进行训练检测的结果。ACF+T+THOG表示利用可见的ACF通道、红外波段的亮度通道和HOG通道得到的结果。在ACF+T+THOG基础上对比增加红外波段中原始的LBP、CS-LBP和MCS-LBP对检测精度的影响,得到的结果分别为ACF+T+THOG+TLBP、ACF+T+THOG+TCS-LBP和ACF+T+THOG+TMCS-LBP。LDCF、random filters和checkerboard filters分别是在ACF+T+THOG+TMCS-LBP基础上引入不同滤波器的结果。

从图4可以看出,利用多光谱的信息可以极大地提高检测精度。事实上,KAIST数据库中双波段图像本身质量欠佳,此处利用单一波段仅检测行人的情况下,红外波段表现优于可见波段。在双波段实验中,通过调整参数,利用ACF+T+THOG,得到了53.47%的MR,较之Hwang等^[6]得到的64.76%已降低了近10%。图4中还比较了添加红外图像的不同形式LBP特征检测结果。加入红外LBP特征其MR降低到52.80%。由于红外图像行人主要表现的是垂直的对称特征,采用MCS-LBP优于原始的LBP和CS-LBP。

图4也对比了不同滤波器组作为中间滤波核的检

测结果。可以看到, LDCF 的检测效果不如随机滤波器, 前者的 MR 为 49.35%, 后者为 48.78%。这一方面是因为 LDCF 在应用 PCA 时隐含了数据符合高斯分布、具有较高信噪比的假设, 对噪声敏感。本文数据库图像质量本身具有挑战性, 尤其对红外图像的非均匀噪声反应敏感。另一方面, 随机滤波器与棋盘格滤波器组的多种滤波器形式可以捕获图像中的局部模式特征。同时可以看出, 棋盘格滤波器组检测效果稍微优于随机滤波器, 其 MR 为 47.38%。这主要是因为前

者的设计过程中包含了对图像边缘模式的考虑, 而后者只是随机生成的。由上面的分析可知, 通过双波段的信息补偿, 并在此基础上加入 MCS-LBP 特征和滤波层, 可以更有效地实现复杂环境下的行人检测。

因为行人检测中大多数仅利用可见波段信息, 为了突出红外波段对可见波段的补偿作用, 这里仅将利用单一可见波段进行检测的结果与 MMACF 中双波段检测结果进行了对比。图 5 为 4 种不同场景下的检测结果对比。其中第一列, 即图 5(a), 为仅利用可见图

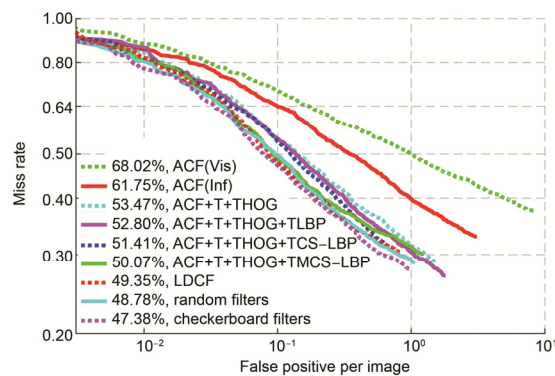


图 4 算法检测结果.

Fig. 4 Results of pedestrian detection.

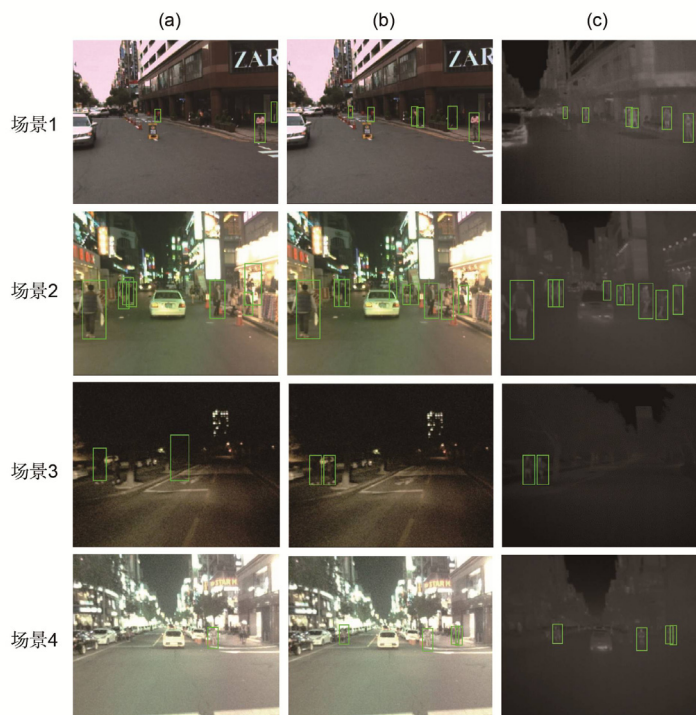


图 5 不同场景的检测结果. (a) 仅利用可见波段训练并在可见图像检测的结果. (b) 利用 MMACF 算法在可见图像检测的结果. (c) 利用 MMACF 算法在红外图像检测的结果.

Fig. 5 Detection results of different scenes. (a) Detection results of a visual images trained by visual band. (b) Detection results of visual images trained by MMACF. (c) Detection results of infrared images trained by MMACF.

像训练得到的分类器对 4 种不同场景可见图像进行检测的结果。场景 1 远处目标较小且与背景相似；场景 2 中行人较多，背景复杂且局部位置光照过亮；场景 3 夜间光照不明；场景 4 中目标较为严重地淹没在背景中，模糊不清。第二列和第三列，即图 5(b)和图 5(c)，分别是利用 MMACF 算法在 4 种场景的可见和红外图像上检测的结果。从图 5 可以看出，对于昏暗、成像不清、光照变化较大的情况，利用本文算法可以较为准确地检测行人。

4 结 论

提出了利用多光谱信息进行行人检测的 MMACF 算法。算法基于配准的多光谱图像对，对可见图像提取常用的 10 个通道特征，对红外图像提取 HOG 特征、亮度通道以及 MCS-LBP 特征；引入滤波层，将多光谱的聚合通道特征的滤波结果送入提升算法进行训练得到分类器。结果表明，MMACF 算法即使在昏暗、模糊或单一波段成像质量不佳情况下，仍然有较好的检测效果。

基金项目

国家自然科学基金资助项目(60874106)。

参考文献

- Zhang S S, Benenson R, Omran M, et al. How far are we from solving pedestrian detection?[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1259–1267.
- Huang Fuyu, Shen Xueju, Liu Xumin, et al. Detection of super wide-field infrared target based on spatial-temporal fusion processing[J]. *Optics and Precision Engineering*, 2015, **23**(8): 2328–2338.
黄富瑜, 沈学举, 刘旭敏, 等. 基于空时域融合处理检测超大视场红外目标[J]. *光学精密工程*, 2015, **23**(8): 2328–2338.
- Davis J, Sharma V. Background-subtraction using contour-based fusion of thermal and visible imagery[J]. *Computer Vision and Image Understanding*, 2007, **106**(2): 162–182.
- Castillo J C, Serrano-Cuerda J, Fernández-Caballero A, et al. Hierarchical Architecture for Robust People Detection by Fusion of Infrared and Visible Video[M]// Novais P, Camacho D, Analide C, et al. *Intelligent Distributed Computing IX. Studies in Computational Intelligence*. Cham: Springer, 2016: 343–351.
- Yuan Y, Lu X Q, Chen X. Multi-spectral pedestrian detection[J]. *Signal Processing*, 2015, **110**(5): 94–100.
- Hwang S, Park J, Kim N, et al. Multispectral pedestrian detection: Benchmark dataset and baseline[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1037–1045.
- Dollar P, Appel R, Belongie S, et al. Fast feature pyramids for object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(8): 1532–1545.
- Heikkila M, Pietikainen M, Schmid C. Description of interest regions with center-symmetric local binary patterns[C]// Kalra P K, Peleg S. *Computer Vision. Graphics and Image Processing*. Berlin, Heidelberg: Springer 2006: 58–69.
- Liu Lin, Geng Junmei, Gu Guohua, et al. Human detection based on contour features and neural networks[J]. *Opto-Electronic Engineering*, 2014, **41**(7): 50–56.
刘琳, 耿俊梅, 顾国华, 等. 轮廓特征与神经网络相结合的行人检测[J]. *光电工程*, 2014, **41**(7): 50–56.
- Gu Lingang, Zhou Mingzheng. Fast pedestrian detection based on multi-feature fusion[J]. *Opto-Electronic Engineering*, 2015, **42**(9): 14–20.
谷灵康, 周鸣争. 多特征融合下的快速行人检测[J]. *光电工程*, 2015, **42**(9): 14–20.
- Nam W, Dollar P, Han J. Local decorrelation for improved pedestrian detection[C]// *Proceedings of the 27th International Conference on Neural Information Processing Systems*, 2014: 424–432.
- Zhang S S, Benenson R, Schiele B. Filtered channel features for pedestrian detection[C]// *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1751–1760.
- Jin Zhigang, Zhao Mingxin, Zhang Rui, et al. Optimized aggregated channel features pedestrian detection algorithm based on binocular vision[J]. *Journal of Tianjin University: Science and Technology*, 2016, **49**(12): 1225–1230.
金志刚, 赵明昕, 张瑞, 等. 基于双目视觉的聚合积分通道行人检测优化算法[J]. *天津大学学报: 自然科学与工程技术版*, 2016, **49**(12): 1225–1230.
- Dollar P, Wojek C, Schiele B, et al. Pedestrian detection: An evaluation of the state of the art[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(4): 743–761.