



# 结合改进聚合通道特征和灰度共生矩阵的俯视行人检测算法

李琳\*, 张涛

(信息工程大学, 郑州 450001)

(\* 通信作者电子邮箱 llbnerfg@163.com)

**摘要:** 针对传统俯视行人检测方法提取的头部特征单一、检测错误率高的问题, 提出了结合改进聚合通道特征 (ACF) 和灰度共生矩阵 (GLCM) 的俯视行人检测算法。首先, 将提取到的 HSV 颜色特征、梯度幅值大小以及改进后的梯度方向直方图 (HOG) 特征组合成 ACF 描述子; 然后, 采用窗口法计算改进的 GLCM 参数描述子, 提取纹理特征, 串联每个窗口的特征向量得到共生矩阵特征描述子; 最后, 将聚合通道和共生矩阵特征分别输入 Adaboost 训练得到分类器, 并进行检测得到最终结果。实验结果表明, 所提算法能在干扰背景存在的情况下有效检测目标, 提高了检测的准确率和召回率。

**关键词:** 俯视行人检测; 聚合通道特征; 灰度共生矩阵; 梯度方向直方图; Adaboost

**中图分类号:** TP391.4 **文献标志码:** A

## Zenithal pedestrian detection algorithm based on improved aggregate channel features and gray-level co-occurrence matrix

LI Lin\*, ZHANG Tao

(Information Engineering University, Zhengzhou Henan 450001, China)

**Abstract:** Aiming at the uniqueness of head feature and high detection error rate extracted by traditional zenithal pedestrian detection method, a multi-feature fusion zenithal pedestrian detection algorithm based on improved Aggregate Channel Feature (ACF) and Gray-Level Co-occurrence Matrix (GLCM) was proposed. Firstly, the extracted Hue, Saturation, Value (HSV) color features, gradient magnitude and improved Histogram of Oriented Gradients (HOG) feature were combined into ACF descriptor. Then, the improved GLCM parameter descriptor was calculated by the window method, and the texture features were extracted. The co-occurrence matrix feature descriptor was obtained by concatenating the feature vectors of each window. Finally, the aggregate channel and co-occurrence matrix features were input into Adaboost for training to get the classifier, and the final results were obtained by detection. The experimental results show that, the proposed algorithm can effectively detect targets in the presence of interference background, and improve the detection precision and recall.

**Key words:** zenithal pedestrian detection; Aggregate Channel Feature (ACF); Gray-Level Co-occurrence Matrix (GLCM); Histogram of Oriented Gradients (HOG); Adaboost

## 0 引言

目前, 智能监控场景下的俯视行人检测技术被广泛应用于人流量统计<sup>[1]</sup>以及行人智能分析<sup>[2]</sup>中, 以实现对人流量密集场所的有效监管。基于俯视的行人检测近些年取得了突飞猛进的发展; 然而行人的发型、发色以及复杂的周围背景加大了检测的难度, 导致检测效果仍不理想, 需进一步改进。因此, 俯视行人检测技术具有极大的研究空间和重要的现实意义。本文重点研究在复杂背景下, 基于俯视视角的行人检测。

俯视行人检测方法有以下几类。第一类是基于形状分析的方法, 主要思想是通过描述行人头部的外观形状, 逐个筛选图像样本, 得到最终检测结果。Chen 等<sup>[3]</sup>通过计算区域面积来识别人头; Yu 等<sup>[4]</sup>采用改进后的 Hough 变换, 通过检测类圆来达到检测行人头部的目的; Li 等<sup>[5]</sup>通过 Canny 边缘提取算法和 Hough 变换结合来提取行人头部特征。基于形状分析的方法计算效率高, 但是实际检测过程中, 容易将外观类

似头部的背景误检为目标。第二类是基于机器学习的方法, 随着机器学习的蓬勃发展, 基于机器学习的特征提取法逐步成为俯视行人检测的常用方法。此类方法首先借助合适的特征描述子, 对样本图片进行特征提取, 之后根据大量训练样本构建俯视行人头部检测分类器, 在此基础上采用分类器进行分类。相比于形状分析法, 机器学习法能适应环境的变化, 在很大程度上提高了检测系统的鲁棒性。牛胜石等<sup>[6]</sup>在梯度特征中加入自生长梯度特征, 并采用支持向量机 (Support Vector Machine, SVM) 和 Adaboost 组合分类器进行训练; 汤石晨等<sup>[7]</sup>通过线性回归模型训练行人的面积特征来检测运动斑块中的头部目标; Li 等<sup>[8]</sup>采用局部二值模式 (Local Binary Pattern, LBP) 提取行人头部特征, 输入 Adaboost 进行分类训练; Tang 等<sup>[9]</sup>采用改进后的梯度方向直方图 (Histogram of Oriented Gradients, HOG) 描述子提取行人头部特征, 输入 SVM 分类器训练得到分类模型。

上述方法通过提取具有一定鲁棒性的特征来训练分类

收稿日期: 2018-05-24; 修回日期: 2018-07-21; 录用日期: 2018-07-31。 基金项目: 国家自然科学基金资助项目 (61572518)。

作者简介: 李琳 (1994—), 女, 北京人, 硕士研究生, 主要研究方向: 图像处理、模式识别; 张涛 (1977—), 男, 湖北天门人, 教授, 博士, 主要研究方向: 图像处理、模式识别、信息隐藏。



器,达到了很好的分类效果,但是只利用单一的特征进行训练容易造成较高误检。在实际的监控视频中,存在诸多具有干扰性的背景,比如:行人穿着的接近发色的深色衣服、行人随身携带的背包,以及在其他方面与行人头部相似的物体。这些干扰的存在加大了检测的难度。单一的特征分类器在这种场景下的检测率较低,这是因为在背景存在诸多干扰时,单一的特征往往不能够全面地描述行人头部,易使分类器混淆行人头部和其他干扰背景,造成误检与漏检,所以研究者尝试使用多种特征来刻画行人目标。Liu等<sup>[10]</sup>结合行人头部颜色和形状描述子来提取行人头部特征,提高了算法的检测准确率;Zeng等<sup>[11]</sup>采用俯视行人的头肩模型,通过HOG-LBP特征构建行人特征描述子,该算法很好地表征目标的轮廓和形状,并且具有良好的灰度不变性。多特征融合的方法在很大程度上改善了分类器的分类效果。然而,几个单一特征的结合不仅提高算法性能有限,并且增大了计算复杂度。所以如何选择合适的特征使分类器在提高检测率的同时保持较快的检测速度是当前的研究重点。

目前,Dollár等<sup>[12]</sup>提出的聚合通道特征(Aggregate Channel Features, ACF)在直立行人领域获得了较好的检测效果,而灰度共生矩阵(Gray-Level Co-occurrence Matrix, GLCM)<sup>[13]</sup>用来提取目标的纹理特征,能够弥补ACF特征在纹理提取上存在的不足,成功分类出目标与与目标区域、灰度相似的干扰背景。

在上述研究的基础上,本文提出了结合改进聚合通道特征(ACF)和灰度共生矩阵(GLCM)的俯视行人检测算法。首先,改进ACF的颜色通道以及梯度直方图通道;然后,尝试采用多特征融合的方法,将改进后的ACF特征与GLCM结合起来,分别提取目标的颜色、梯度、梯度方向直方图以及纹理特征,输入Adaboost算法进行训练检测得到结果。实验结果表明,所提多特征融合的方法,能够过滤掉与人头相似的背景,有效降低分类器的漏检与误检率,实现了较高的检测性能。

## 1 特征分析

### 1.1 改进的聚合通道特征

#### 1.1.1 特征描述

ACF特征起初被用来解决人脸识别问题,之后逐渐被应用于行人检测中。ACF特征的基本结构是通道,它分别采用颜色、梯度以及梯度方向直方图三个不同的通道描述图像。这种混合通道的采用大大提高了描述子对不同种类信息的分类表示能力,有助于提高复杂背景下目标的检测效果。

ACF特征的计算流程示意图如图1所示,在积分通道特征的基础上,ACF描述子采用特征金字塔算法提高了特征提取的速度。下面就该描述子的计算方法进行简要介绍。

1)通道类型。ACF特征定义了10个通道,分别是3个LUV颜色通道、1个梯度通道和6个梯度方向直方图通道。计算梯度幅值的过程采用一维微分模板 $[-1, 0, 1]$ 和 $[1, 0, -1]$ 分别得到水平方向的梯度分量 $G_x(x, y)$ 和竖直方向的梯度分量 $G_y(x, y)$ 。梯度方向直方图通道遵循HOG的思想,构造方式是将图像均匀分割成若干块,计算每个块中像素的梯度方向,将其按照划分的6个梯度方向区域进行统计,其中每一个方向区域都分别是1个特征通道。图像中像素点 $(x, y)$ 的梯度幅值和梯度方向计算式如式(1)、(2)所示:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (1)$$

$$\alpha(x, y) = \arctan(G_y(x, y)/G_x(x, y)) \quad (2)$$

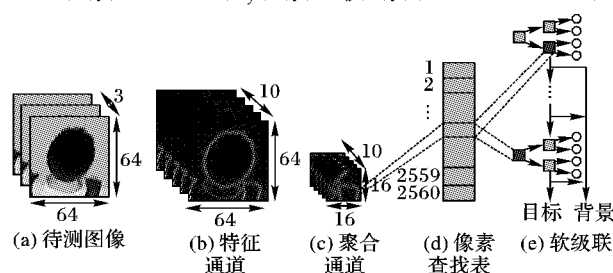


图1 聚合通道特征计算示意图

Fig. 1 Schematic diagram of aggregate channel features

2)特征计算。聚合通道的特征计算十分简单,首先给定一个彩色图像作为输入,计算预定义的通道。然后,根据预设因子对通道特征进行平均池化下采样,将下采样得到的矩阵矢量化为像素查找表,该矢量即为最终的特征描述子。

3)多尺度计算。为了提升计算效率,ACF特征采用一种特征金字塔算法。在行人检测中,目标的高矮、远近导致它到摄像头的距离不同,如果采用相同的尺度检测目标会导致漏检和误检,采用多尺度计算能够有效解决这一问题。传统的做法是先将图像进行重采样缩放到若干尺度,再采用滑动窗口法计算每个尺度下图像的特征。这种方法存在冗余计算,且耗时大。而ACF算法则先计算几个特定尺度层 $s \in \{1, 1/2, 1/4, \dots\}$ 下的特征,然后直接使用一种快速估计算法计算其余中间尺度下的特征值。采用快速特征金字塔作为估计算法,利用近似公式(3),在允许的误差范围内根据当前尺度的特征 $C$ 估计相邻尺度上对应的通道特征:

$$C_s \approx R(C, s) \cdot s^{-\lambda_n} \quad (3)$$

式中, $\lambda_n$ 表示每个通道特征对应的系数。

#### 1.1.2 改进颜色通道

ACF特征通道中融合了LUV颜色描述子。对于颜色通道的选择,本文作出改进,选用更适合的HSV(Hue, Saturation, Value)颜色通道代替原先使用的LUV颜色通道。

HSV颜色空间<sup>[14]</sup>属于非线性色彩表示系统,它的颜色参数分别是:色调(H),饱和度(S)和亮度(V)。HSV颜色空间对色彩的感知和人十分接近,它能够从人类的视角非常直观地表达色彩信息,达到更好的颜色对比效果。HSV颜色通道能够更细致地刻画目标和背景之间的灰度差别,有效提高检测率。利用RGB(Red, Green, Blue)计算HSV颜色空间的计算式如下:

$$S = \begin{cases} \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}, & \max(R, G, B) \neq 0 \\ 0, & \max(R, G, B) = 0 \end{cases} \quad (4)$$

$$H = \begin{cases} \frac{(-B + G) * p/3}{\max(R, G, B) - \min(R, G, B)}, & \text{当 } R = \max(R, G, B) \\ \frac{(-R + B) * p/3}{\max(R, G, B) - \min(R, G, B)}, & \text{当 } G = \max(R, G, B) \\ \frac{(-G + R) * p/3}{\max(R, G, B) - \min(R, G, B)}, & \text{当 } B = \max(R, G, B) \\ \text{未定义}, & R = G = B \end{cases} \quad (5)$$

$$V = \max(R, G, B) \quad (6)$$



### 1.1.3 改进 HOG 通道

传统 HOG 算法<sup>[15]</sup>的基本思想是通过提取图片的梯度大小和方向来描述行人特征。它将图像分成小的连通区域,名为细胞单元(cell),然后统计每个胞元内部像素的梯度和方向直方图,随后对梯度强度作归一化,将胞元有一定重叠地组合成块(block),以梯度作为权重统计块内像素的方向直方图,串联起来得到最终的 HOG 特征。

HOG 描述子对几何和光学形变有良好鲁棒性,但同时,它也有着计算量大的缺陷。为了提高算法的实时性,本文结合行人头部的特点,对 HOG 描述子作出如下改进,在不影响检测性能的基础上缩短提取特征的时间。

首先,由于俯视行人的头部特征形状近似圆形,较为统一,细节特征较少,所以考虑不再将块进一步细分为更小的细胞单元,并且采用无重叠的子块,减少提取到的冗余信息。其次,本文考虑将  $0^\circ \sim 180^\circ$  的角度范围扩展至  $-180^\circ \sim 180^\circ$ ,这样能够更完整地体现目标的变化特征。

对改进前后的 HOG 特征维数进行对比。当检测图像大小为  $64 \times 64$  时,传统 HOG 算法将方向直方图划分为 6 个相等大小的区域,采用  $8 \times 8$  大小的像素组成一个胞元,每  $2 \times 2$  大小的胞元组成一个块,且移动步长为 8,则传统 HOG 描述子的维数为  $6 \times 4 \times 7 \times 7 = 1176$ 。而本文改进算法将图像等分成  $4 \times 4$  共 16 个方形块,块与块之间互不重叠,则改进后的 HOG 描述子只有  $16 \times 6 = 96$  维。相比于传统方法,本文提出的改进算法突出了俯视行人的纹理和颜色特征,弱化了不必要的细节特征,大大降低了传统 HOG 描述子的维数,在保证分类准确率的前提下,减少了冗余的计算量。

### 1.2 改进的纹理特征

本文采用灰度共生矩阵来描述图像的纹理特征。灰度共生矩阵通过统计图像中一定距离  $d$  和一定方向  $\theta$  上两像素点灰度的空间相关性来描述纹理。

灰度共生矩阵定义彼此距离为  $d$ 、方向夹角为  $\theta$  的一对像素,其灰度值分别为  $i$  和  $j$  时的概率,记为  $P(i, j, d, \theta)$ 。通常方向夹角选取  $0^\circ, 45^\circ, 90^\circ, 135^\circ$  四个方向。下面给出像素灰度值分别为  $f(x_1, y_1) = a$  与  $f(x_2, y_2) = b$  的像素距离为  $d$  时的灰度共生矩阵计算公式:

$$P(i, j, d, 0^\circ) = \# \{ ((x_1, y_1), (x_2, y_2)) \in (L_r, L_c) \times (L_r, L_c) \mid x_2 - x_1 = 0, |y_2 - y_1| = d, f(x_1, y_1) = i, f(x_2, y_2) = j \} \quad (7)$$

$$P(i, j, d, 45^\circ) = \# \{ ((x_1, y_1), (x_2, y_2)) \in (L_r, L_c) \times (L_r, L_c) \mid (x_2 - x_1 = d, y_2 - y_1 = d) \text{ or } (x_2 - x_1 = -d, y_2 - y_1 = -d), f(x_1, y_1) = i, f(x_2, y_2) = j \} \quad (8)$$

$$P(i, j, d, 90^\circ) = \# \{ ((x_1, y_1), (x_2, y_2)) \in (L_r, L_c) \times (L_r, L_c) \mid |x_2 - x_1| = d, y_2 - y_1 = 0, f(x_1, y_1) = i, f(x_2, y_2) = j \} \quad (9)$$

$$P(i, j, d, 135^\circ) = \# \{ ((x_1, y_1), (x_2, y_2)) \in (L_r, L_c) \times (L_r, L_c) \mid (x_2 - x_1 = d, y_2 - y_1 = -d) \text{ or } (x_2 - x_1 = -d, y_2 - y_1 = d), f(x_1, y_1) = i, f(x_2, y_2) = j \} \quad (10)$$

其中:  $L$  为图像的灰度级;  $L_r$  和  $L_c$  分别表示图像的行数和列数;  $\# \{ \cdot \}$  表示该集合中的元素个数。

由于直接将灰度共生矩阵送入分类器会产生较大计算量,所以为了减小计算复杂度,本文采用文献[13]提出的 4 种特征参数来间接描述图像的纹理特征。

#### 1) 能量。

能量特征参数又被称为二阶矩,它是矩阵值的平方和。能量的大小表示的是图像灰度分布的匀称度以及纹理细腻度。若一幅图像的灰度分布均匀且纹理细腻,则能量值小;反之,如果图像中的灰度分布不均且纹理粗糙,则能量值大。能量的计算式如下:

$$ASM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P^2(i, j, d, \theta) \quad (11)$$

#### 2) 熵。

灰度共生矩阵的熵能够反映出图像中携带的信息量多少,属于随机性度量值。当图像的灰度共生矩阵中所有值均相等时,图像的随机性较大,熵值大;相反,当图像灰度共生矩阵中的值分布不均时,图像的随机性较小,熵值小。熵值的计算式如下:

$$ENT = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i, j, d, \theta) \log P(i, j, d, \theta) \quad (12)$$

#### 3) 对比度。

对比度指的是灰度共生矩阵主对角线附近的惯性矩,主要用来表示图像的清晰度以及沟纹深浅程度。如果图像清晰且纹理沟纹深,则对比度大;而当图像模糊且沟纹较浅时,对比度小。对比度采用式(13)进行计算:

$$CON = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} n^2 \left\{ \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i, j, d, \theta) \right\} \quad (13)$$

#### 4) 相关度。

相关度是用来度量灰度共生矩阵的元素在水平或垂直方向上的相似度。当矩阵中的值变化较小时或者图像纹理具备一些规律性时,相关值大;当矩阵中各元素的值相差很大时,相关值小。相关度的计算式如下:

$$CORR = \left( \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} ijP(i, j, d, \theta) - \mu_i \mu_j \right) / (\sigma_i^2 \sigma_j^2) \quad (14)$$

其中,  $\mu_i, \mu_j, \sigma_i^2$  和  $\sigma_j^2$  的定义如下:

$$\mu_i = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} i \cdot P(i, j, d, \theta) \quad (15)$$

$$\mu_j = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} j \cdot P(i, j, d, \theta) \quad (16)$$

$$\sigma_i^2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - \mu_i)^2 P(i, j, d, \theta) \quad (17)$$

$$\sigma_j^2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (j - \mu_j)^2 P(i, j, d, \theta) \quad (18)$$

灰度共生矩阵的特征参数能够对纹理进行细致的刻画。为了验证灰度共生矩阵作为分类特征的有效性,随机选取 100 张行人头部图像,以及 100 张干扰背景图像(包括深色背包、衣服等),分别计算所有行人头部和所有干扰背景在 4 个方向  $0^\circ, 45^\circ, 90^\circ, 135^\circ$  上的特征参数,取平均值进行对比,结果如图 2 所示。实验中取灰度级为  $L = 16$ ,距离  $d = 1$ 。由图 2 可以看出,头部目标和干扰背景的特征值之间存在明显差异,因此将灰度共生矩阵作为附加的特征输入分类器能够有效降低因干扰背景导致的误检率。

由于灰度共生矩阵的特征维数只有 4 维,过低的维数会导致较差的分类效果。所以本文参考文献[13]中的方法,在训练过程中采用大小为  $8 \times 8$ 、步长为 6 的窗口,按照从左至右、从上至下的顺序依次在图像上滑动,计算每个窗口内的特征参数。将每个窗口的 4 维特征参数依次串联起来,得到最终的灰度共生矩阵描述子,维数为  $10 \times 10 \times 4 = 400$ 。这种基





于小区域计算共生矩阵的方法不仅大幅度提高了训练效果,并且由于是在局部方格上进行操作,所以提取到的特征对图像的几何和光学形变具有良好的不变性。

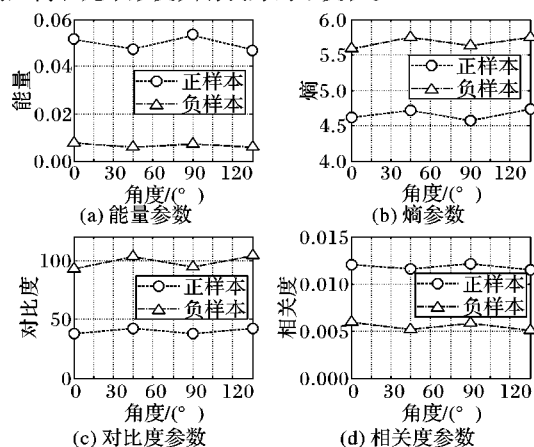


图2 四种特征参数对比

Fig. 2 Parameter comparison of four kinds of features

### 1.3 软级联 Adaboost 分类器

传统 Adaboost 算法的基本思想是将  $K$  个弱分类器 (Weak classifier) 级联组成更强的强分类器  $H(x)$ 。

为了提升分类效果,本文采用深度为 2 的决策树作为弱分类器,并且使用软级联 Adaboost<sup>[16]</sup> 取代传统分类器以实现更优的分类性能。

软级联 Adaboost 分类器相对于传统 Adaboost 分类器在算法上有了改进。不同于传统方法需要反复的调参才能获得近似最优的级联结构,软级联方法在每个弱分类器之后都设置了一个阈值,样本每通过一个弱分类器就作出一次决策,从而使分类器的复杂度和分类能力得到逐级提升,实现了更好的分类性能。

## 2 结合改进 ACF 和 GLCM 的俯视行人检测算法

### 2.1 算法流程

本文从多特征融合的角度出发,分别提取图像的聚合通道特征 (ACF) 和灰度共生矩阵 (GLCM) 特征,利用 Adaboost 分类器进行训练,得到分类器 1 和分类器 2,以级联的方式将聚合通道特征和灰度共生矩阵特征参数训练得到的两个分类器结合起来,采取这样的方法能够有效过滤掉干扰背景,准确检测出目标。本文算法流程如图 3 所示,具体步骤如下。

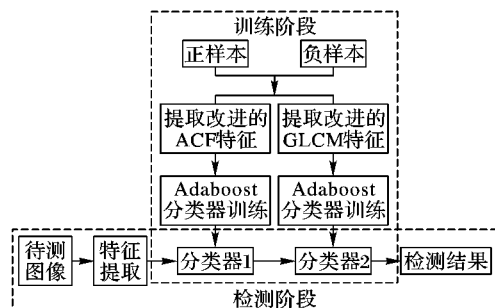


图3 本文算法流程

Fig. 3 Flow chart of the proposed algorithm

### 2.2 训练阶段

训练阶段的步骤如下:

1) 将训练的正样本和负样本归一化至  $64 \times 64$  大小,灰度

化之后作为输入。对样本进行下采样得到大小为  $16 \times 16$  的图像,分别计算训练样本 10 个通道的 ACF 特征,其中包括 3 个 HSV 通道、6 个改进的 HOG 通道,以及 1 个梯度幅值通道。计算得到特征向量。

2) 采用区域法计算归一化后正负样本的灰度共生矩阵,选取块大小为 8、步长为 6、灰度级为 16、距离为 1、方向为  $0^\circ$ 。利用得到的共生矩阵,分别计算能量、熵、对比度和相关度,得到 4 维的特征参数。最终每个样本得到 400 维特征向量。

3) 将得到的聚合通道特征向量和灰度共生矩阵特征向量分别送入软级联 Adaboost 分类器进行训练,其中:软级联阈值设置为 -1,弱分类器个数为 128。得到分类器 1 和分类器 2,用于后续测试。

### 2.3 测试阶段

测试阶段的步骤如下:

1) 首先按帧读入图像,进行图像灰度化。采用滑动窗口法计算原图的 ACF 特征,窗口大小为  $64 \times 64$ 。利用特征金字塔的思想直接对图像特征进行缩放,得到不同尺度图像下的特征向量。

2) 将计算得到的特征向量送入 Adaboost 分类器 1 进行分类得到检测结果的坐标 (bounding box) 和目标窗口。将检测到的行人头部窗口以及其坐标作为候选目标,进行下一步的筛选。计算该目标的共生矩阵参数向量,输入分类器 2,排除干扰背景,得到最终的输出结果。

## 3 实验结果及分析

为了验证本文所提算法的有效性,在配置为 Intel Core i5 CPU 2.40 GHz、内存 6.00 GB 的计算机上采用 Matlab 2014 平台编写算法验证实验。本文研究的对象是俯视镜头下拍摄到的行人目标。实验用到的样本图像取自某学校食堂抓拍的彩色监控图像序列,分辨率为  $320 \times 240$ 。训练集中的正样本通过人工裁剪图像序列中的头部目标区域得到,样本大小为  $64 \times 64$ ,数量为 850 张。在截取的头部样本中,约 60% 的头部样本选取较为规范的类圆形头像,约 40% 的头部样本则选取一些发型发色不规律的人头像,采用上述样本集训练得到的分类器泛化能力较强。训练集中的负样本从 1 000 多张不含行人的图片中随机截取得到,大小为  $64 \times 64$ ,数量为 4 000 张。从监控序列中选取完整连续的 3 段视频图像序列作为测试视频,每段视频有 200 帧,其中每帧至少有 1 个行人目标,行进速度和方向均为随机。用于测试的图像序列中具有较多干扰背景如深色的背包、服装等,且序列中头部目标包含多种发型及发色,增大了检测难度。测试集样本与训练集样本没有重叠。

对比实验采用准确率 (precision) 和召回率 (recall) 作为最终评价指标,计算式如下:

$$\text{precision} = \frac{tp}{(tp + fp)} \quad (19)$$

$$\text{recall} = \frac{tp}{(tp + fn)} \quad (20)$$

其中:  $tp$  是正确检测的行人数目;  $fn$  是漏检的人数;  $fp$  是算法误检出的行人数目。所以,  $tp + fp$  是算法总共检出的行人总数,  $tp + fn$  是所有正样本的数目。

实验采用本文算法和文献[12]中所提的 ACF 算法进行测试对比,结果如表 1 所示。

从表 1 中可以看出,本文提出的方法明显缩减了漏检和误检人数,在准确率和召回率上都比文献[12]方法略高。一



方面,本文采用的 HSV 颜色通道相比 LUV 通道能够达到更好的颜色对比效果,能够提取出鲁棒性更强的颜色特征;另一方面,本文结合了颜色、梯度、梯度方向直方图以及纹理 4 种特征,把改进后的聚合通道特征和灰度共生矩阵融合起来,将聚合特征描述子检测到的目标窗口输入灰度共生矩阵纹理描述子中进行筛选,极大程度上减少了误检与漏检目标,提高了准确率和召回率。两种检测方法的检测结果如图 4 所示。

表 1 不同方法实验结果对比

Tab. 1 Comparison of experimental results of different methods

实验组别	实验方法	目标人数	检测人数	漏检人数	误检人数	准确率/%	召回率/%
1	本文算法	145	142	12	9	93.83	91.72
	ACF 算法	145	144	15	14	90.27	89.65
2	本文算法	203	208	10	15	92.78	95.07
	ACF 算法	203	215	10	22	89.76	95.07
3	本文算法	231	236	18	23	90.25	92.20
	ACF 算法	231	244	22	35	85.65	90.47

在 Matlab 2014 环境下,一幅  $320 \times 240$  的 RGB 图像,用传统 ACF 算法结合灰度共生矩阵检测需要 2 s,而本文算法需要约 0.5 s,提升了运算速度。

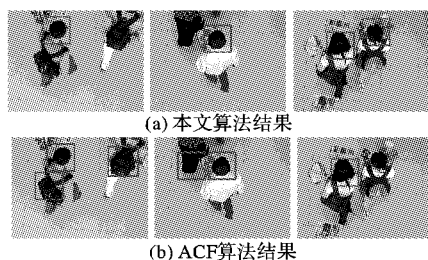


图 4 俯视行人不同算法检测结果对比

Fig. 4 Zenithal pedestrian detection result comparison of different algorithms

## 4 结语

本文提出了一种针对俯视视角的行人检测算法,从多特征融合的角度出发,将改进后的聚合通道特征和灰度共生矩阵训练得到的分类器级联起来,通过逐级筛选的方式进行分类,得到检测结果。该算法融合了颜色、梯度、梯度方向直方图和纹理特征,通过从不同角度描述行人来有效检测行人头部,减小了干扰背景对分类的影响,大幅提高了检测性能,为后续的行人跟踪与智能分析奠定了良好基础。

## 参考文献 (References)

- [1] QIAN X, YU X, FA C. The passenger flow counting research of subway video based on image processing [C]// Proceedings of the 29th Chinese Control and Decision Conference. Piscataway, NJ: IEEE, 2017: 5195–5198.
- [2] AGGARWAL J K, RYOO M S. Human activity analysis: a review [J]. ACM Computing Surveys, 2011, 43(3): Article No. 16.
- [3] CHEN T H, CHEN T Y, CHEN Z X. An intelligent people-flow counting method for passing through a gate [C]// Proceedings of the 2006 IEEE Conference on Robotics, Automation and Mechatronics. Piscataway, NJ: IEEE, 2006: 1–6.
- [4] YU H B, HE Z W, LIU J L. A vision-based method to estimate passenger flow in bus [C]// Proceedings of the 2007 International Symposium on Intelligent Signal Processing and Communication Systems. Piscataway, NJ: IEEE, 2007: 654–657.
- [5] LI J, HE Q, YANG L, et al. Pedestrian detection and counting in crowded scenes [C]// Proceedings of the 2016 International Conference on Green Intelligent Transportation System and Safety, LNEE 419. Berlin: Springer, 2016: 495–511.
- [6] 牛胜石, 毛晓晖, 侯建华, 等. 基于 Adaboost 和 SVM 的人头实时检测[J]. 微型机与应用, 2010(13): 33–36. (NIU S S, MAO X H, HOU J H, et al. Realtime head detection based on Adaboost and SVM [J]. Microcomputer & Its Applications, 2010(13): 33–36.)
- [7] 汤石晨, 陈锻生. 基于光流法和行人面积特征的人数统计方法研究[J]. 图学学报, 2013, 34(1): 139–144. (TANG S C, CHEN D S. A method for people counting based on optical flow and pedestrian size [J]. Journal of Graphics, 2013, 34(1): 139–144.)
- [8] LI B, ZHANG J, ZHANG Z, et al. A people counting method based on head detection and tracking [C]// Proceedings of the 2014 International Conference on Smart Computing. Piscataway, NJ: IEEE, 2014: 136–141.
- [9] TANG C H, CHEN Q J. Zenithal people counting using histogram of oriented gradients [C]// Proceedings of the 5th International Congress on Image and Signal Processing. Piscataway, NJ: IEEE, 2012: 946–951.
- [10] 刘恋, 贺赛先, 熊杰. 一种基于多特征融合的头部检测算法[J]. 电光与控制, 2016, 23(9): 106–110. (LIU L, HE S X, XIONG J. A head detection algorithm based on multi-feature fusion [J]. Electronics Optics & Control, 2016, 23(9): 106–110.)
- [11] ZENG C B, MA H D. Robust Head-shoulder detection by PCA-based multilevel HOG-LBP detector for people counting [C]// Proceedings of the 20th International Conference on Pattern Recognition. Piscataway, NJ: IEEE, 2010: 2069–2072.
- [12] 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.
- [13] HARALICK R M, SHANMUGAM K, DINSTEL I. Texture features for image classification [J]. IEEE Transactions on Systems, Man & Cybernetics, 1973, SMC-3(6): 610–621.
- [14] SMITH A R. Color gamut transform pairs [C]// Proceedings of the 5th Conference on Computer Graphics and Interactive Techniques. New York: ACM, 1978: 12–19.
- [15] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]// Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2005: 886–893.
- [16] BOURDEV L, BRANDT J. Robust object detection via soft cascade [C]// Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2005: 236–243.
- [17] YANG B, YAN J J, LEI Z, et al. Aggregate channel features for multi-view face detection [C]// Proceedings of the 2014 IEEE International Joint Conference on Biometrics. Piscataway, NJ: IEEE, 2014: 1–8.
- [18] OJALA T, PIETIKAINEN M, HARWOOD D. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions [C]// Proceedings of the 12th International Conference on Pattern Recognition. Piscataway, NJ: IEEE, 1994: 582–585.
- [19] VIOLA P, JONES M J. Robust real-time object detection [J]. International Journal of Computer Vision, 2004, 57(2): 137–154.
- [20] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object detection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627–1645.



- 2005: 507–514.
- [8] YANG Y, YANG Y, SHEN H T, et al. Discriminative nonnegative spectral clustering with out-of-sample extension [J]. *IEEE Transactions on Knowledge & Data Engineering*, 2013, 25(8): 1760–1771.
  - [9] XIA T, TAO D H, MEI T, et al. Multiview spectral embedding [J]. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 2010, 40(6): 1438–1446.
  - [10] BELKIN M, NIYOGI P. Laplacian eigenmaps and spectral techniques for embedding and clustering [C]// *Proceedings of the 2001 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*. Cambridge, MA: MIT Press, 2001: 585–591.
  - [11] CAI D, HE X F, HAN J W. Semi-supervised discriminant analysis [C]// *Proceedings of the 2007 IEEE 11th International Conference on Computer Vision*. Piscataway, NJ: IEEE, 2007: 1–7.
  - [12] XU Z L, LING I, LYU M R-T, et al. Discriminative semi-supervised feature selection via manifold regularization [J]. *IEEE Transactions on Neural Networks*, 2010, 21(7): 1033–1047.
  - [13] COELHO F, BRAGA A P, VERLEYSEN M. Multi-objective semi-supervised feature selection and model selection based on pearson's correlation coefficient [C]// *Proceedings of the 2010 Iberoamerican Congress Conference on Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Berlin: Springer, 2010: 509–516.
  - [14] DU L, SHEN Y D. Unsupervised feature selection with adaptive structure learning [C]// *Proceedings of the 2015 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2015: 209–218.
  - [15] CHEN X, LIU W, SU F L, et al. Semi-supervised multi-view feature selection with label learning for VHR remote sensing images [C]// *Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium*. Piscataway, NJ: IEEE, 2016: 2372–2375.
  - [16] SUN S L, JIN F, TU W T. View construction for multi-view semi-supervised learning [C]// *Proceedings of the 2011 8th International Conference on Advances in Neural Networks*. Berlin: Springer, 2011: 595–601.
  - [17] SHI C J, AN G Y, ZHAO R Z, et al. Multiview Hessian semisupervised sparse feature selection for multimedia analysis [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(9): 1947–1961.
  - [18] ZHU S H, SUN X, JIN D L. Multi-view semi-supervised learning for image classification [J]. *Neurocomputing*, 2016, 208: 136–142.
  - [19] FENG Y F, XIAO J, ZHUANG Y T, et al. Adaptive unsupervised multi-view feature selection for visual concept recognition [C]// *Proceedings of the 2012 11th Asian Conference on Computer Vision*. Berlin: Springer, 2012: 343–357.
  - [20] WEI X K, CAO B K, YU P S. Multi-view unsupervised feature selection by cross-diffused matrix alignment [C]// *Proceedings of the 2017 International Joint Conference on Neural Networks*. Piscataway, NJ: IEEE, 2017: 494–501.
  - [21] SONG Y Q, NIE F P, ZHANG C S, et al. A unified framework for semi-supervised dimensionality reduction [J]. *Pattern Recognition*, 2008, 41(9): 2789–2799.
  - [22] OLSHAUSEN B A, FIELD D J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images [J]. *Nature*, 1996, 381: 607–609.
  - [23] HOU C P, ZHANG C S, WU Y, et al. Multiple view semi-supervised dimensionality reduction [J]. *Pattern Recognition*, 2010, 43(3): 720–730.
  - [24] 陶红. 多视角数据分析算法研究[D]. 长沙: 国防科学技术大学, 2014: 15–28. (TAO H. Study on algorithms for analyzing multi-view data [D]. Changsha: National University of Defense Technology, 2014: 15–28.)
  - [25] 汪荆琪. 基于多视角的半监督特征选择算法研究[D]. 合肥: 中国科学技术大学, 2014: 5–8. (WANG J Q. Semi-supervised feature selection for multi-view data [D]. Hefei: University of Science and Technology of China, 2014: 5–8.)
  - [26] 郝伟, 刘忠宝. 基于 Fisher 准则的半监督特征提取方法[J]. *计算机工程与设计*, 2017, 38(1): 238–241. (HAO W, LIU Z B. Semi-supervised feature extraction method based on Fisher criterion [J]. *Computer Engineering and Design*, 2017, 38(1): 238–241.)

This work is partially supported by the Fundamental Research Funds for the Central Universities (2018IB016).

**SUN Shengzi**, born in 1996. Her research interests include machine learning, high dimensional image feature dimension reduction.

**WAN Yuan**, born in 1976, Ph. D., professor. Her research interests include machine learning, image processing, high dimensional image feature dimension reduction.

**ZENG Cheng**, born in 1997. His research interests include machine learning, high dimensional image feature dimension reduction.

(上接第3371页)

- [21] GAN G L, CHENG J. Pedestrian detection based on HOG-LBP feature [C]// *Proceedings of the 7th International Conference on Computational Intelligence and Security*. Piscataway, NJ: IEEE, 2011: 1184–1187.
- [22] 唐春晖. 一种基于多特征融合的单目俯视行人检测[J]. *系统仿真学报*, 2016, 28(9): 2146–2153. (TANG C H. Zenithal pedestrian detection using multiple feature fusion in monocular vision [J]. *Journal of System Simulation*, 2016, 28(9): 2146–2153.)
- [23] 王强, 冯燕. 基于颜色和形状信息的快速人数统计方法[J]. *计算机测量与控制*, 2010, 18(9): 2157–2159. (WANG Q, FENG Y. A fast people counting algorithm based on fusion of color and shape information [J]. *Computer Measurement & Control*, 2010, 18(9): 2157–2159.)
- [24] DOLLÁR P, TU Z W, PERONA P, et al. Integral channel features [C]// *Proceedings of the 2009 British Machine Vision Conference*. Durham, UK: BMVA Press, 2009: 91.1–91.11.
- [25] 栾书鹏. 基于图块和二阶统计特征的行人检测[J]. *电子设计工程*, 2016, 24(20): 190–193. (LUAN S P. Pedestrian detection based on blocks-graphs and second-order statistics [J]. *Electronic Design Engineering*, 2016, 24(20): 190–193.)

This work is partially supported by the National Natural Science Foundation of China (61572518).

**LI Lin**, born in 1994, M. S. candidate. Her research interests include image processing, pattern recognition.

**ZHANG Tao**, born in 1977, Ph. D., professor. His research interests include image processing, pattern recognition, information hiding.