

# 基于增强聚合通道特征的实时行人重识别

黄新宇, 许娇龙, 郭 纲, 郑二功

中国人民解放军空军航空大学飞行器控制系, 吉林 长春 130022

**摘要** 由于目标姿态、摄像头角度、光线条件等因素的影响,行人重识别仍然是一个具有挑战性的问题。目前大多数方法主要注重提高重识别精度,对实时性考虑较少。因此,本文提出了一种基于增强聚合通道特征(ACF)的实时行人重识别算法。利用 ACF 对行人进行检测,并在此基础上,结合直方图特征和纹理特征构成增强 ACF,作为行人重识别的特征描述子。利用测度学习方法对重识别模型进行训练。在 4 个数据集上的实验结果表明,与传统的重识别特征相比,提出的特征描述子逼近最好的重识别准确率,并且具有更快的计算速度。整个行人检测与重识别系统的运行速度达到  $10 \text{ frame}\cdot\text{s}^{-1}$  以上,基本可以满足实时行人重识别的需求。

**关键词** 图像处理; 实时行人重识别; 增强聚合通道特征; 测度学习; 深度学习

**中图分类号** TP391 **文献标识码** A

**doi:** 10.3788/LOP54.091001

## Real-Time Pedestrian Reidentification Based on Enhanced Aggregated Channel Features

Huang Xinyu, Xu Jiaolong, Guo Gang, Zheng Ergong

Academy of Aircraft Control, Aviation University Air Force, Changchun, Jilin 130022, China

**Abstract** The pedestrian reidentification is still a challenging problem due to various pedestrian poses, camera viewpoints and illumination conditions etc. Most of the reported works focus on improving the reidentification accuracy without considering the real-time capability. We propose a real-time pedestrian reidentification algorithm based on aggregated channel features (ACF). The ACF is applied to detect the pedestrian candidates, and the extracted ACF features are enhanced with histogram features and texture features and used as a pedestrian reidentification feature descriptor. Finally, based on the enhanced ACF features, we apply the metric learning to train the pedestrian identification model. The experimental results on four datasets show that the proposed feature descriptor obtains the higher recognition accuracy and much faster computation speed compared with the traditional reidentification features. The proposed pedestrian detection and reidentification framework has a running speed of above  $10 \text{ frame}\cdot\text{s}^{-1}$ , and it can basically meet the needs of real-time pedestrian reidentification tasks.

**Key words** image processing; real-time pedestrian reidentification; enhanced aggregated channel features; metric learning; deep learning

**OCIS codes** 100.4995; 040.5160; 100.3008

## 1 引言

随着社会公共安全越来越受到重视,许多重要场所(政府机关、商场、机场、车站、学校等)都安装了大量的监控摄像头,使得重识别技术可以很好地解决跨摄像头目标识别的问题。

当前行人重识别算法主要分为两类:基于特征直接判别的方法<sup>[1-3]</sup>和基于距离度量学习的方法<sup>[4-10]</sup>。文献[1]提出时空分割算法,利用颜色、边缘的显著特征来描述目标,获得一个响应特征图。文献[2]利用等级

**收稿日期:** 2017-02-17; **收到修改稿日期:** 2017-03-22

**基金项目:** 国家自然科学基金(6160011396)、吉林省教育厅“十三五”科学技术研究项目(吉教科合字[2016]第 515 号)

**作者简介:** 黄新宇(1992—),男,硕士研究生,主要从事计算机视觉与机器学习方面的研究。E-mail: 469051525@qq.com

**导师简介:** 郭 纲(1971—),男,博士,副教授,主要从事检测技术与自动化装置等方面的研究。

E-mail: 617611086@qq.com(通信联系人)

高斯分布来描述局部图像的均值和协方差信息,并形成相应的块特征,该方法使目标的纹理信息更为突出。文献[3]分别利用头部、躯干、腿的颜色直方图对行人特征进行了分块描述,使行人特征判别性更强。但基于特征直接判别方法得到的结果稳健性较差,很难找到一种对行人图像的各种变化具有不变性的稳健特征。

相比于基于特征直接判别的方法,基于距离度量学习的方法近年来更受关注,并且取得了良好的效果。文献[4]基于马氏距离提出保持简单有效原则下的距离测度学习(KISSME)算法,利用马氏距离得到度量矩阵,从而使正确匹配的似然率高于错误匹配的似然率,同时在预处理数据时利用主成分分析进行降维,减少维度之间的相关性。之后的大间隔最近邻(LMNN)<sup>[5]</sup>算法、信息理论度量学习(ITML)算法<sup>[6]</sup>、逻辑判别距离测度学习(LDML)算法<sup>[7]</sup>等都是基于马氏距离进行度量学习来实现重识别的。

文献[8]提出了改进的局部化的 Fisher 判别分析(LFDA)算法,主要是基于度量学习的降维处理。针对局部 Fisher 判别分析方法对于数据量较少时的信息弱化问题,文献[9]采用核空间的方法对其进行了改进,提出了核局部 Fisher 判别分析(kLFDA)算法,其计算速度比 LFDA 算法更快并且不会减少有用的判别信息。但由于 kLFDA 和 LFDA 使用多维直方图特征和矩特征来作为特征描述子,特征计算时间较长,不能满足实时行人重识别任务的要求。

相对传统的人工设计特征,近年来深度学习方法被广泛应用于行人重识别领域。因为深度学习本身就具有特征学习能力,能够从原始图像数据获得比人工设计特征具有更强表达能力的特征,为图像分类、目标检测、图像分割等领域带来了突破性的进展。文献[10-11]提出将深度学习用于重识别领域。文献[12]利用深度学习对重识别的联合特征进行排名,通过深度学习方法,使重识别的准确率得到了很大的提升。

关于行人目标检测的课题,近年来也有许多学者进行了研究<sup>[13-15]</sup>。但基于准确率、计算速度以及特征复用性的综合考虑,本文选用聚合通道特征(ACF)行人检测器<sup>[16]</sup>来实现目标检测。为了实现实时行人重识别,对行人重识别的描述特征进行了研究,利用增强 ACF 将行人目标检测与重识别结合到一起,构建了完整的实时行人重识别系统框架。

## 2 算法介绍

### 2.1 基于增强 ACF 的行人重识别基本框架

建立了实时行人重识别系统,如图 1 所示,主要分为三个模块:目标检测、特征提取以及基于距离测度学习的重识别。通过目标检测模块检测出所有的行人目标,以构建重识别的目标搜索库;通过特征提取模块可以将目标检测模块中计算的特征用来作为重识别的描述特征之一,并融合色度、饱和度、纯度(HSV)直方图特征和尺度不变的局部三值模式(SILTP)算法纹理特征;通过基于测度学习的重识别模块将待识别目标与搜索库中的目标进行匹配,从中找出待识别目标。

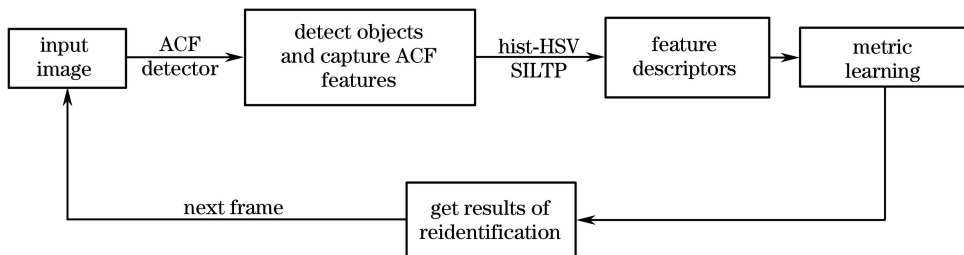


图 1 系统框架结构

Fig. 1 Framework of the system

### 2.2 增强 ACF

由于传统的 ACF 主要用于目标检测,即区分行人和背景。如果将其直接作为行人重识别任务的特征描述子,特征表示能力不足,尤其是对于类内目标的区分。而由于 HSV 直方图特征在局部形变上具有不变性,SILTP 算法特征在纹理信息描述上更详细,可以弥补 ACF 的不足,增强其特征稳健性。因此,将 ACF 特征、HSV 直方图特征、SILTP 纹理特征进行融合,形成一个对行人重识别特征描述性能更优的增强 ACF,记为 ACF+。

### 2.2.1 ACF 特征

ACF 在积分通道特征(ICF)<sup>[17]</sup>的基础上发展而来。二者采用了相同的特征通道:LUV 颜色通道、梯度幅值、方向梯度直方图(HOG),共计 10 个通道。这些通道特征能够从不同角度对图像信息进行描述,如图 2 所示。

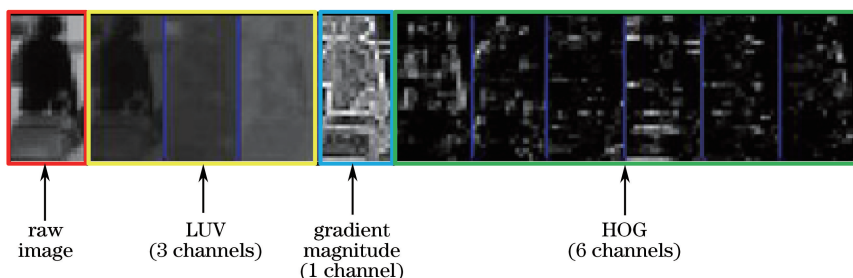


图 2 通道特征

Fig. 2 Channel features

ACF 通过对各个通道进行平滑和下采样,提取图像中各个位置的特征,在“聚合”的通道中进行单像素查找,因此,具有计算速度快的特点。采用 ACF 检测器进行行人检测,使各个行人样本的 ACF 特征可以在检测阶段直接获取,不需要额外计算,节省了重识别表现特征的计算时间。经过实验证明,ACF 在速度和结构上优于可变形部件模型(DPM)<sup>[18]</sup>所采用的 HOG 特征,对于行人和背景具有很好的区分能力。但是对于相同行人以及不同行人之间的区分,特征表示能力有限,因此需要结合直方图特征和 SILTP 纹理特征进行增强。

### 2.2.2 HSV 直方图特征

直方图特征主要反映各像素计算出的通道值的分布情况,并且能直观地反映出不同区域在颜色空间的差异性。HSV 直方图特征被证实对于行人重识别(区分类内样本)有较强的特征表示能力<sup>[19]</sup>。为了能使 HSV 直方图特征与 ACF 特征结合(图 3),采用  $8 \times 8$  的窗口大小,以步长为 4 进行滑动窗口扫描提取特征,这样就保证了 HSV 直方图特征与 ACF 特征有相同大小( $32 \times 12$ )的特征图谱。至于通道数,对 HSV 的 3 个通道分别计算其像素值的直方图分布(直方图个数设置为 8),形成 24 个特征通道。

### 2.2.3 SILTP 纹理特征

仅用外观特征对目标进行描述不够全面,所以加入了 SILTP<sup>[20]</sup>纹理特征,它是对局部三值模式(LTP)<sup>[21]</sup>的改进,对局部图像噪声和光线的急剧变化都具有很好的尺度不变特性。SILTP 将每个像素点与周围像素进行比较(邻域半径不同,尺度也不同),从而得到一个编码值,将其代替像素值,从而实现单像素映射。按照计算 HSV 直方图的方式同样进行滑动窗口扫描提取特征。由于选用了两个尺度的 SILTP 特征(忍耐值为 0.3,半径为 3 和 5,邻域点数目为 4)进行计算,所以通道数为 2。

由图 3 可知,通过滑动窗口的计算方式使得 HSV 直方图特征以及 SILTP 特征形成单独的通道特征能够与 ACF 进行组合,形成 ACF+,把所有的通道都融合到了一起,共计 36 个通道。

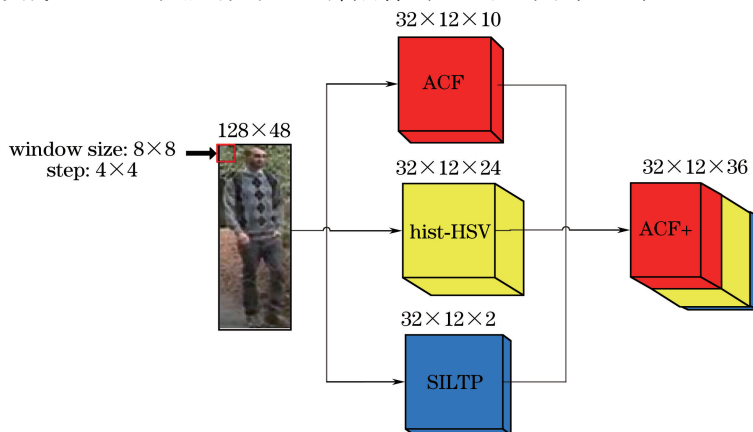


图 3 ACF+特征组成

Fig. 3 Components of ACF+

### 2.3 测度学习

在前面特征提取的基础上,采用测度学习来获得衡量行人特征相似度的距离函数,其目的是使得相同行人特征之间的距离尽可能小而不同行人之间的距离尽可能大。由于 ACF+特征维度比较高、计算量大,难以满足实时行人重识别计算要求,需要进行降维处理,并且在降维过程中尽量保留不同行人目标之间的一些重要判别信息。

采用文献[9]提出的 kLFDA 算法作为行人重识别的实验基线,对 ACF+特征下的行人重识别性能进行了实验。假设特征向量矩阵  $\mathbf{X} = (x_1, x_2, x_3, \dots, x_n)$ , 一共包含  $n$  个样本, 分别属于  $m$  类  $x_i \in \{1, 2, 3, \dots, m\}$ , 并且  $x_i \in \mathbf{R}^d$ 。  $z_i \in \{1, 2, 3, \dots, m\}$  为投影后的特征向量,  $z_i \in \mathbf{R}^D$ , 则有  $z_i = \mathbf{T}'x_i$ 。  $\mathbf{T}$  为投影矩阵,  $\mathbf{T}'$  是  $\mathbf{T}$  的转置矩阵, 其维度为  $D(1 \leq D \leq d)$ 。假设  $\phi(X)$  表示样本特征,  $\tilde{\mathbf{S}}^w$  和  $\tilde{\mathbf{S}}^b$  分别为类内和类间散度矩阵, 则

$$\mathbf{S}^w = \frac{1}{2} \phi(X) \tilde{\mathbf{S}}^w \phi(X)^T, \quad (1)$$

$$\mathbf{S}^b = \frac{1}{2} \phi(X) \tilde{\mathbf{S}}^b \phi(X)^T, \quad (2)$$

投影矩阵  $\mathbf{T}$  表示为

$$\mathbf{T} = \underset{\mathbf{T}}{\operatorname{argmax}} r [(\mathbf{T}'\mathbf{S}^w\mathbf{T})^{-1}\mathbf{T}'\mathbf{S}^b\mathbf{T}], \quad (3)$$

具体计算公式参考文献[9], 本文采用欧氏距离对降维后的特征相似度进行测量和计算。

### 2.4 算法步骤

通过对上述行人检测和重识别算法的具体研究, 针对整个实时行人重识别系统给出了具体的算法实现过程, 如算法 1 所示。

算法 1: 实时行人重识别

输入: 目标特征  $F$ 、图像  $I_i$

输出: 重识别目标边界框  $B$

1) for  $i=1:N$

2) 在  $I_i$  中检测出所有行人的  $B_{ij}$ , 并计算 ACF+特征  $F'_{ij}$

3) 利用测度学习训练得到投影矩阵  $\mathbf{T}$ , 对  $F$  及  $F'_{ij}$  降维得到  $F_d$  和  $F'_{dij}$ , 计算样本距离  $D_{ij} = \|F_d - F'_{dij}\|_2$

4)  $B = \min(B_{ij}), D_{ij} \leq \beta, \beta$  为设定阈值

5) end

输入图像序列, 在第一张图像中指定需要重识别的行人目标, 提取出特征  $F$ 。读取第  $i$  张图像, 利用 ACF 行人检测器进行检测, 保存检测出的所有行人图像  $j(1 \leq j \leq N)$ ,  $N$  为检测出的总行人数。同时提取每张图像的边界框  $B_{ij}$ , 并计算相应的 HSV 直方图和 SILTP 特征, 与检测过程中已经计算完的 ACF 结合到一起得到 ACF+特征, 用  $F'_{ij}$  表示。通过 kLFDA 算法得到  $F_d$  和  $F'_{dij}$ , 最后计算出特征距离  $D_{ij}$ , 设置阈值  $\beta$ , 当  $D_{ij} \leq \beta$  时判别为待识别目标, 选择距离最小的样本并返回其边界框  $B_{ij}$ 。

## 3 实验结果与分析

### 3.1 数据集及评估标准

将提出的 ACF+重识别性能分别在 VIPeR<sup>[22]</sup>、iLIDS<sup>[23]</sup>、3DPeS<sup>[24]</sup>、CAVIAR<sup>[25]</sup> 数据集上进行了测试, 并将数据集随机分成 5 个子训练集和测试集进行实验, 实验结果取其平均值, 实验中  $\beta$  设置为 8.0。数据集具体信息如表 1 所示。

表 1 数据集信息

Table 1 Information of datasets

Dataset	Camera	Size of dataset (pedestrian pairs)	Training set	Testing set
iLIDS	2	476(119)	416	60
VIPeR	2	1264(632)	948	316
3DPeS	8	1011(192)	916	95
CAVIAR	2	1220(72)	1184	36

采用累积匹配特性(CMC)曲线来评价重识别性能。该曲线表示在图像数据库中对待识别行人目标进行搜索,在前 $r$ 个结果中找到该目标的比率,通常以 $r=1$ 作为最佳参考值。同时,计算了不确定分数比例为

$$P_{\text{PUR}} = \frac{\lg N + \sum_{r=1}^N \mathbf{M}(r) \lg [\mathbf{M}(r)]}{\lg N}, \quad (4)$$

式中 $\mathbf{M}(r)$ 为累积匹配特性矩阵。

### 3.2 实验结果与分析讨论

#### 3.2.1 重识别特征性能比较

将ACF+与其他不同的重识别特征进行实验和对比分析,主要包括以下5种特征:

1) ACF特征,作为比较基线。

2) Hist-HSV特征,采用8个直方图,以搜索窗口为 $8 \times 8$ 、步长为4来提取目标特征。

3) 文献[9]提出的HistMoment-6Patch特征,采用颜色空间与局部二值模式(LBP)直方图特征以及颜色空间相应的矩特征。

4) imagenet-vgg-f-pool5特征,本文使用Matconvnet中的imagenet-vgg-f模型对图像特征进行计算。提取了conv1层、conv2层、pool5层的特征,由表2可知,对于重识别任务深度网络的底层特征表示能力比较有限(如conv1和conv2层),而较高层的特征对重识别性能有显著提升(如pool5层),所以选择用pool5层的特征作为描述特征。

5) 文献[26]提出的LOMO特征,采用3个尺度图像,分别提取HSV直方图特征和2个尺度的SILTP特征,目前该特征在行人重识别数据集上取得了较好的测试结果。

表2 不同网络层特征在VIPeR数据集上的识别率

Table 2 Recognition accuracy of different features of network layer on VIPeR dataset

Layer	Rank1	Rank5	Rank10	Rank20
Conv1	0.06	0.19	0.30	0.42
Conv2	0.13	0.36	0.51	0.68
Pool5	0.18	0.44	0.59	0.75

图4为上述5种特征基于kLFDA算法在4个数据集上的CMC性能曲线。

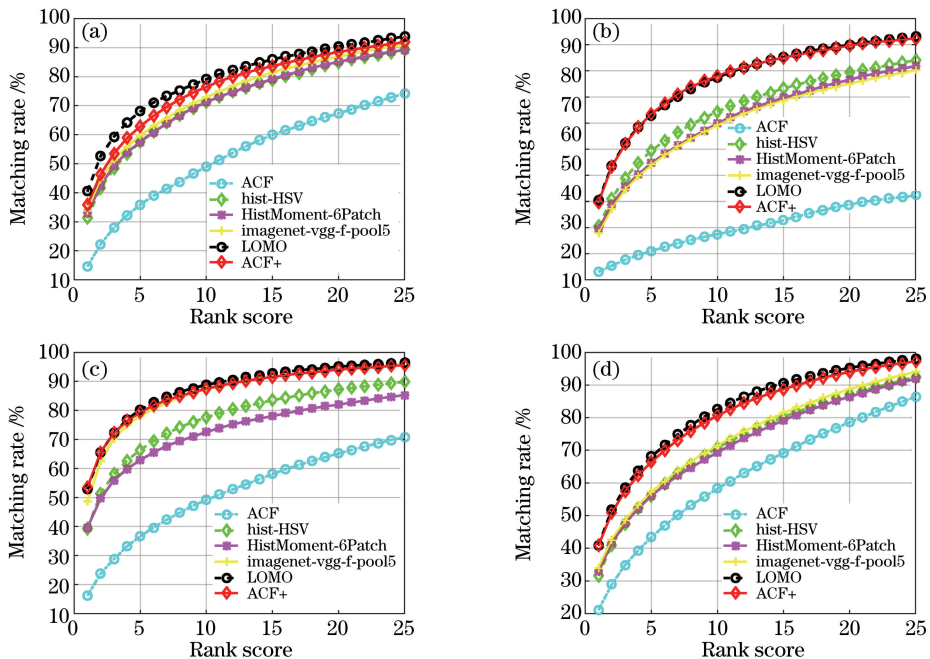


图4 5种特征在4个数据集上的CMC曲线。(a) iLIDS数据集;(b) VIPeR数据集;(c) 3DPeS数据集;(d) CAVIAR数据集

Fig. 4 CMC curves on four datasets with five features. (a) iLIDS dataset; (b) VIPeR dataset;

(c) 3DPeS dataset; (d) CAVIAR dataset



从图 4 可以看出,单独使用 ACF 特征重识别的性能不太理想,因此需要结合其他特征来增强识别效果。与 ACF 和 HistMoment-6Patch 特征相比,ACF+在性能上有了很大提高,甚至比 imagenet-vgg-f-pool5 特征还要好。而与 LOMO 特征相比,ACF+在 iLIDS 和 CAVIAR 数据集上性能稍微差一些,在 VIPeR 和 3DPeS 数据集上性能几乎持平。具体的实验数据结果见表 3~6( $P$  代表测试样本数),不确定分数比例的计算见(4)式。

表 3 不同特征在 iLIDS 数据集上的识别率( $P=60$ )Table 3 Recognition accuracy of different features on iLIDS dataset ( $P=60$ )

Rank	1	5	10	20	$P_{PUR}$
ACF	0.15	0.36	0.50	0.68	11.09
Hist-HSV	0.32	0.58	0.70	0.84	27.62
HistMoment-6patch	0.33	0.59	0.72	0.86	28.85
Imagenet-vgg-f-pool5	0.34	0.60	0.73	0.87	30.65
LOMO	0.40	0.68	0.79	0.90	37.97
ACF+	0.36	0.63	0.76	0.89	33.14

表 4 不同特征在 VIPeR 数据集上的识别率( $P=316$ )Table 4 Recognition accuracy of different features on VIPeR dataset ( $P=316$ )

Rank	1	5	10	20	$P_{PUR}$
ACF	0.03	0.11	0.17	0.28	8.42
Hist-HSV	0.20	0.50	0.64	0.79	41.19
HistMoment-6patch	0.19	0.45	0.60	0.77	37.76
Imagenet-vgg-f-pool5	0.18	0.44	0.59	0.75	37.79
LOMO	0.30	0.63	0.78	0.90	51.26
ACF+	0.30	0.63	0.78	0.90	50.77

表 5 不同特征在 3DPeS 数据集上的识别率( $P=95$ )Table 5 Recognition accuracy of different features on 3DPeS dataset ( $P=95$ )

Rank	1	5	10	20	$P_{PUR}$
ACF	0.16	0.37	0.50	0.65	16.48
Hist-HSV	0.39	0.66	0.78	0.87	40.57
HistMoment-6patch	0.40	0.63	0.73	0.82	37.28
Imagenet-vgg-f-pool5	0.49	0.78	0.86	0.94	52.57
LOMO	0.53	0.80	0.89	0.95	55.45
ACF+	0.54	0.79	0.87	0.94	54.83

表 6 不同特征在 CAVIAR 数据集上的识别率( $P=36$ )Table 6 Recognition accuracy of different features on CAVIAR dataset ( $P=36$ )

Rank	1	5	10	20	$P_{PUR}$
ACF	0.21	0.42	0.57	0.77	10.09
Hist-HSV	0.32	0.56	0.71	0.87	20.82
HistMoment-6patch	0.33	0.56	0.69	0.86	20.96
Imagenet-vgg-f-pool5	0.34	0.57	0.71	0.89	20.70
LOMO	0.41	0.68	0.83	0.95	32.91
ACF+	0.41	0.66	0.81	0.94	31.54

通过对表 3~6 中的实验数据的分析,可以得到以下结论:

- 1) ACF+ 优于 HistMoment-6Patch 特征。在 iLIDS 和 VIPeR 数据集上, Rank1、Rank5、Rank10、Rank20 分别提高了 3%、4%、4%、3%和 11%、18%、18%、13%。
- 2) Hist-HSV 特征对于重识别目标特征的表现作用很大,证明了将其作为 ACF+ 的一部分的重要性。
- 3) ACF+ 在性能上超过了 imagenet-vgg-f-pool5 特征,从一定程度上说明了深度学习特征在某些条件下并不是最优的特征描述子。

上述实验对比了不同特征在各个数据集上的重识别准确率,在此基础上,进一步对特征计算速度进行了对比。实验平台的软硬件配置为 AMD-A6 处理器、2GB 内存,本文算法采用 Matlab 软件实现。表 7 列出了不同特征在 VIPeR 数据集上的计算速度。

表 7 不同特征在 VIPeR 数据集上的计算速度

Table 7 Computation speed of different features on VIPeR dataset

Feature	Computation speed / (frame·s <sup>-1</sup> )
ACF	131
Hist-HSV	25
HistMoment-6patch	<1
Imagenet-vgg-f-pool5	6
LOMO	17
ACF+	20

从表 7 可以得出以下结论:

1) HistMoment-6patch 特征的计算速度非常慢,不适合用来作为实时行人重识别特征。

2) ACF+虽然在平均准确率上不如 LOMO 特征,但是在计算速度上稍快于 LOMO 特征,并且数据集越大,这种优势越明显。

通过对 ACF+和其他特征在重识别准确率与计算速度方面的综合比较,可以看出 ACF+具有更好的行人重识别综合性能。

### 3.2.2 重识别系统实时性分析

表 8 实时行人重识别系统速度

Table 8 Speed of real-time pedestrian reidentification

Detector	Feature descriptor	Metric learning	Detection speed / (s·frame <sup>-1</sup> )	Reidentification speed / (s·frame <sup>-1</sup> )
ACF	ACF+	kLFDA	0.036	0.0625

基于 ACF 行人检测器、ACF+特征描述子和 kLFDA 测度学习方法,构成了一个完整的实时行人重识别系统。在 3DPeS 数据集上的运行速度如表 8 所示,整个系统的运行速度为 0.0985 s·frame<sup>-1</sup>,即约 10.2 frame·s<sup>-1</sup>。为保证画面的连续性并尽量节省带宽,远程网络监控视频帧速率一般设置为 10 frame·s<sup>-1</sup>以上,因此,本文重识别系统基本可以满足实时性需求。

### 3.2.3 重识别系统定性分析

为了更直观地说明本文重识别系统的性能,给出一个在 3DPeS 视频图像序列上的测试实例。如图 5(a)所示,首先在摄像头 A 的图像序列中指定待识别行人目标,在图中用红色矩形框标记,在之后的图像中用 ACF 行人检测器检测出所有的行人,非待识别行人目标用绿色矩形框标记。实时重识别结果如图 5(b)所示,左侧图像中行人的特征距离 dis 大于  $\beta$ ,判断为非待识别目标,用绿色矩形框标记;中间和右侧图像中行

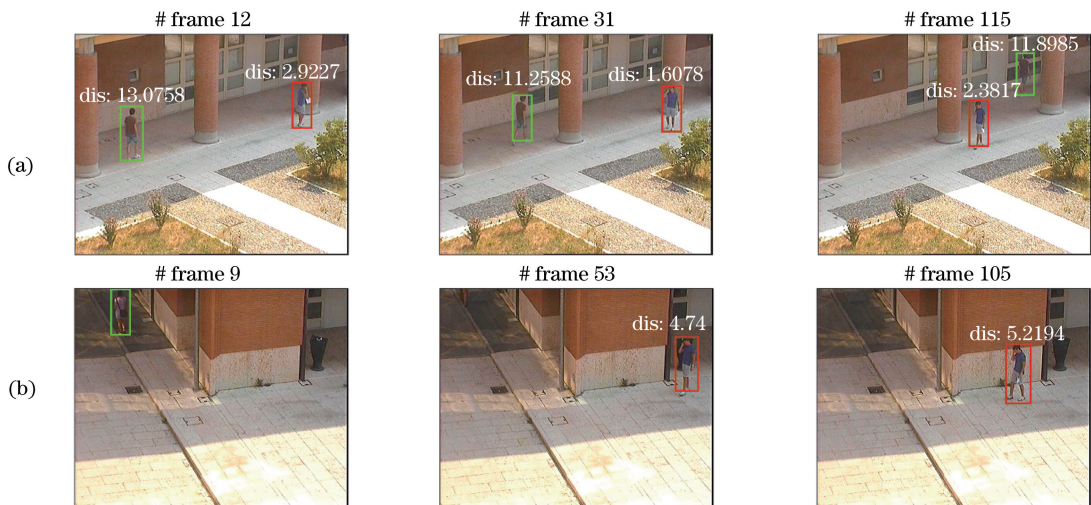


图 5 kLFDA+ACF+实时重识别结果。(a)摄像头 A;(b)摄像头 B

Fig. 5 Real-time reidentification results of kLFDA+ACF+. (a) Camera A; (b) camera B

人的特征距离分别为 4.74 和 5.22, 均小于阈值  $\beta$ , 所以判断为待识别目标, 用红色矩形框标记。检测结果与真实情况一致。

## 4 结 论

针对多摄像头监控中的实时行人重识别应用问题, 建立了一个行人检测与重识别相统一的框架; 基于特征复用原则和完备性原则, 提出一种增强 ACF, 并结合测度学习实现了快速准确的行人重识别算法。实验结果表明, 本文算法在保持很好的重识别精度的同时, 基本满足行人重识别的实时性需求, 具有一定的应用价值。但是, 目前行人检测器在复杂场景下的检测精度还有待改进, 下一步的工作将所提出的增强 ACF 用于训练行人检测器, 提高其检测性能。

## 参 考 文 献

- [1] Gheissari N, Sebastian T B, Hartley R. Person reidentification using spatiotemporal appearance[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, 2: 1528-1535.
- [2] Matsukawa T, Okabe T, Suzuki E, *et al.* Hierarchical gaussian descriptor for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1363-1372.
- [3] Das A, Chakraborty A, Roy-Chowdhury A K. Consistent re-identification in a camera network[C]. European Conference on Computer Vision, 2014: 330-345.
- [4] Köstinger M, Hirzer M, Wohlhart P, *et al.* Large scale metric learning from equivalence constraints[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2012, 8690: 2288-2295.
- [5] Weinberger K Q, Saul L K. Distance metric learning for large margin nearest neighbor classification[J]. Journal of Machine Learning Research, 2009, 10(2): 207-244.
- [6] Davis J V, Kulis B, Jain P, *et al.* Information-theoretic metric learning[C]. Proceedings of the 24<sup>th</sup> International Conference on Machine Learning, 2007: 209-216.
- [7] Guillaumin M, Verbeek J, Schmid C. Is that you? Metric learning approaches for face identification[C]. IEEE 12<sup>th</sup> International Conference on Computer Vision, 2009: 498-505.
- [8] Pedagadi S, Orwell J, Velastin S, *et al.* Local fisher discriminant analysis for pedestrian re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2013: 3318-3325.
- [9] Xiong F, Gou M R, Camps O, *et al.* Person re-identification using kernel-based metric learning methods[C]. European Conference on Computer Vision, 2014: 1-16.
- [10] Yi D, Lei Z, Liao S C, *et al.* deep metric learning for person re-identification[C]. 22<sup>nd</sup> International Conference on Pattern Recognition, 2014: 34-39.
- [11] Li W, Zhao R, Xiao T, *et al.* Deepreid: Deep filter pairing neural network for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 152-159.
- [12] Chen S Z, Guo C C, Lai J H. Deep ranking for person re-identification via joint representation learning[J]. IEEE Transactions on Image Processing, 2016, 25(5): 2353-2367.
- [13] Qin Jian, Wang Meihua. Fast pedestrian proposal generation algorithm using online Gaussian model[J]. Acta Optica Sinica, 2016, 36(11): 1115001.  
覃 剑, 王美华. 采用在线高斯模型的行人检测候选框快速生成方法[J]. 光学学报, 2016, 36(11): 1115001.
- [14] Yu Xiaqiong, Chen Xiangning. Robust method for moving object detection in dynamic background[J]. Laser & Optoelectronics Progress, 2011, 48(7): 071001.  
喻夏琼, 陈向宁. 运动序列中动目标检测的稳健性方法[J]. 激光与光电子学进展, 2011, 48(7): 071001.
- [15] Chen Yin, Ren Kan, Gu Guohua, *et al.* Moving object detection based on improved single Gaussian background model[J]. Chinese J Lasers, 2014, 41(11): 1109002.  
陈 银, 任 侃, 顾国华, 等. 基于改进的单高斯背景模型运动目标检测算法[J]. 中国激光, 2014, 41(11): 1109002.
- [16] Dollár 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.



- [17] Dollár P, Tu Z, Perona P, *et al.* Integral channel features[C]. Proceedings of the British Machine Vision Conference, 2009, 23(91): 1-11.
- [18] 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.
- [19] Liu C X, Gong S G, Loy C C, *et al.* Person re-identification: what features are important?[C]. European Conference on Computer Vision, 2012, 7583: 391-401.
- [20] Liao S C, Zhao G Y, Kellokumpu V, *et al.* Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2010: 1301-1306.
- [21] Tan X Y, Triggs B. Enhanced local texture feature sets for face recognition under difficult lighting conditions[J]. IEEE Transactions on Image Processing, 2010, 19(6): 1635-1650.
- [22] Gray D, Brennan S, Tao H. Evaluating appearance models for recognition, reacquisition, and tracking [C]. IEEE International Workshop on Performance Evaluation for Tracking and Surveillance, 2007, 3(5): 1-7.
- [23] Zheng W S, Gong S, Xiang T. Associating groups of people[C]. British Machine Vision Conference, 2009(23): 1-11.
- [24] Baltieri D, Vezzani R, Cucchiara R. 3DPeS: 3D people dataset for surveillance and forensics[C]. Proceedings of the 2011 Joint ACM Workshop on Human Gesture and Behavior Understanding, 2011: 59-64.
- [25] Cheng D S, Cristani M, Stoppa M, *et al.* Custom pictorial structures for re-identification[C]. British Machine Vision Conference, 2011, 2(5): 6.
- [26] Liao S C, Hu Y, Zhu X Y, *et al.* Person re-identification by local maximal occurrence representation and metric learning[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 2197-2206.