

基于特征融合与子空间学习的行人重识别算法

朱小波^{1,2}, 车进^{1,2*}

¹宁夏大学物理与电子电气工程学院, 宁夏 银川 750021;

²宁夏大学宁夏沙漠信息智能感知重点实验室, 宁夏 银川 750021

摘要 针对现存行人重识别算法不能较好地适应光照、姿态、遮挡等变化的问题, 提出一种基于特征融合与子空间学习的行人重识别算法。该算法对整幅行人图像提取方向梯度(HOG)直方图特征和 HSV(Hue, Saturation, Value)直方图特征作为整体特征, 再在滑动窗口内提取色彩命名(CN)特征和两个尺度的尺度不变局部三元模式(SILTP)特征。为了使算法具有更好的尺度不变性, 对原图像进行两次下采样, 再对采样后的图像提取上述特征。提取特征后, 采用核函数分别将原始特征空间转换到非线性空间, 在非线性空间内学习一个子空间, 同时在子空间内学习一个相似性度量函数。在 3 个公开数据集上进行了实验, 结果表明, 所提算法可以较好地提高重识别率。

关键词 机器视觉; 行人重识别; 特征融合; 子空间; 非线性; 核函数

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.021503

Person Re-Identification Algorithm Based on Feature Fusion and Subspace Learning

Zhu Xiaobo^{1,2}, Che Jin^{1,2*}

¹School of Physics and Electronic-Electrical Engineering, Ningxia University, Yinchuan, Ningxia 750021, China;

²Ningxia Key Laboratory of Intelligent Sensing for Desert Information, Ningxia University, Yinchuan, Ningxia 750021, China

Abstract Aiming at the problem that the existing person re-identification algorithm cannot be adapted well to the variances of illumination, attitude and occlusion, a novel person re-identification algorithm based on feature fusion and subspace learning is proposed, in which the Histogram of Oriented Gradient (HOG) feature and the Hue-Saturation-Value (HSV) histogram feature are first extracted from the entire pedestrian image as the overall feature and then the Color Naming (CN) feature and the two-scale Scale Invariant Local Ternary Pattern (SILTP) feature are extracted in a sliding window. In addition, in order to make this algorithm have better scale invariance, the original images are first down-sampled twice and then the above features are extracted from the sampled images. After the features are extracted, a kernel function is used to transform the original feature space into a nonlinear space, in which a subspace is learned. Simultaneously, in this subspace, a similarity function is learned. The experiments on three public datasets are conducted and the results show that the proposed algorithm can be used to improve the re-identification rate relatively well.

Key words machine vision; person re-identification; feature fusion; subspace; nonlinearity; kernel function

OCIS codes 150.0155; 110.2970; 150.0155

1 引言

行人重识别被普遍认为是一个图像检索的子问题, 主要解决的问题是从其他摄像头所拍摄的图像中集中检索是否存在当前摄像头中感兴趣的目标人物。然而, 如果摄像机的分辨率较低, 所处的位置不同, 会

导致所拍摄到的图像在光照、姿态、视角、遮挡等方面产生很大差异, 这些问题给重识别带来了很大的挑战。

行人图像的特征表达和相似性度量是行人重识别的两个关键问题。在特征表达方面, Farenzena 等^[1]通过分割法将图片中的背景剔除, 再对行人图像进行对称和不对称部件分割, 对分割后的不同区

收稿日期: 2018-07-09; 修回日期: 2018-07-30; 录用日期: 2018-08-08

基金项目: 国家自然科学基金(61861037)

* E-mail: koalache@126.com

域分别提取重复度高的结构区域、极大稳定颜色区域和带权重的颜色直方图 3 个特征。黄新宇等^[2]首先采用聚集信道特征(ACF)算法检测行人,再将直方图和纹理特征结合作为行人特征描述子。Zhao 等^[3]提出了一种 mid-level 滤波器特征,具有跨视野不变性。Liu 等^[4]将局部二值特征和 Gabor、Schmid 滤波器特征结合起来作为行人图像描述子。范彩霞等^[5]将图像的二阶空间直方图作为整体特征,将 HSV(Hue, Saturation, Value)颜色特征和面向梯度(HOG)直方图特征作为局部特征,再将整体特征和局部特征融合。袁立等^[6]通过提取颜色直方图特征、颜色特征和 HOG 特征作为行人图像描述子。行人重识别的另一个过程是寻找一个稳健的相似性度量函数。Zheng 等^[7]采用学习行人特征间最佳距离对不同行人特征向量进行相似性度量。齐美彬等^[8]提出将原始特征空间通过核函数转换到非线性空间后再进行度量的重识别算法。

近年来,深度学习^[9]也被应用于行人重识别算法的研究中,并在大数据集上取得了不错的识别效果。Zheng 等^[10]将校验模型和识别模型相结合,并采用交互熵损失函数得到一个更具判别性的行人特征描述子。陈兵等^[11]采用卷积神经网络,通过改善样本之间距离关系的约束和增强样本监督信息的使用来达到提高识别率的目的。但是,由于深度学习需要的数据集较大(在小数据集上会导致过拟合),对于不同问题所需要的模型不相同,在理论方面还缺乏一定的指导,因此具有较大的主观性。

为了解决目前行人重识别中存在的问题,本文在现有算法的基础上提出一种新的算法。在特征提取前,对图像进行预处理以增强图像的可处理性。预处理后,对行人图像分别提取局部特征和整体特征,再利用核函数将所提取的特征转换到非线性空间,在非线空间学习一个子空间,最后在子特征空间学习一个测度矩阵进行相似度量。选取 HOG 直方图作为整体特征,在局部部件上选取要提取的特征为颜色命名(CN)特征和两个尺度的尺度不变局部三元模式(SILBP)特征。核函数主要选取效果较好的 RFB- χ^2 (RBF: Radial Basis Function)。通过在 3 个数据集上进行实验,结果表明,所提出的行人重识别方法能够提高识别效果。

2 特征选择与提取

2.1 图像预处理

为了获得更好的颜色特征和视觉效果,在对图

像进行特征提取之前采用带颜色恢复的多尺度 Retinex(MSRCR)^[12]算法对行人图像进行预处理。MSRRCR 最早是由 Edwin 于 1971 年提出的一种基于色彩恒常理论得到的图像增强方法,该算法与传统的图像增强算法只能增强图像中的某一种特征相比,但 MSRRCR 算法可以在边缘增强、动态范围压缩和颜色恒常 3 个方面达到平衡,并且可以对各种不同类型的图像进行自适应增强。在尺度选择方面,这里选择了大小为 5 和 20 的两种尺度。MSRRCR 变换前后的图像如图 1 所示。其中同一列为同一个人,第一行为变换前的图片,第二行为变换后的图像。由图可以看出,处理后的图像在视觉效果方面得到了很大的改善。



图 1 MSRRCR 变换前后的图像(VIPeR 数据集)

Fig. 1 Images before and after MSRRCR transformation (VIPeR dataset)

2.2 整体与局部特征选取

首先对图像提取整体特征,采用 HOG 和 HSV 颜色特征作为行人图像的整体特征。由于 HOG 是在行人图像的局部单元格上提取特征,因此对图像的光学和几何形变都具有一定的稳健性,然而,上述两种形变只有在更大的空间上才会出现。除此之外,HOG 还具有较强的局部光学归一化性能。而 HSV 颜色空间的亮度和拍摄图像的彩色信息无关,饱和度和色调与人感受颜色的方式具有相似度关系,因此能够较好地反映人类对颜色的感知能力。HSV 颜色直方图描述的是不同颜色在整幅图像中所占的比例,实现简单,效果较好,因此选择 HSV 颜色直方图作为整体的颜色特征。

局部特征选取 SILTP 特征和 CN 特征。SILTP 特征是基于经典算法 LBP(Local Binary Pattern)的改进算法,对光照具有较好的稳健性。事实上,LBP 特征仅在灰度变换下具有很好的稳健性,但是对噪声比较敏感,SILTP 通过引入尺度不变局部比较容差改善 LBP,以此实现对图片噪声的稳健性。CN 特征是指人们使用一定的方法和准则对生活中出现的颜色以一种语义给予标签,目前在目

标分类领域具有一定的应用。由于在不同的光照下颜色的名称不会改变,因此 CN 特征不仅对形变、遮挡具有一定的稳健性,对光照也具有一定的稳健性。

2.3 整体与局部特征提取

为了精确地识别出目标行人,采用整体特征和局部特征相结合的办法。整体特征是对整幅图像提取的特征,因此不仅可以对行人进行快速识别,而且还考虑到了整幅行人图像各部件之间的联系。由于不同位置的摄像头拍摄到同一个行人时,受到遮挡、视角、几何形变的影响较大,因此这种情况下采用整体特征的匹配率较低,而采用局部特征可以很好地解决这个问题。在特征提取完成后,将整体特征和局部特征融合组成最终的特征描述子。

整体特征提取:对行人图像提取 HOG 特征时,第 1 步,对原始图像进行灰度化处理;第 2 步,采用 Gamma 校正法对图像颜色空间进行归一化处理,这样不仅可以调节图像的对比度,还可以减少噪声;第 3 步,计算图像每个像素点的梯度值;第 4 步,将图像分成 $n \times n$ 大小的小格子;第 5 步,统计每个小格子里的梯度直方图(每个小格子有 9 个维度的描述子);第 6 步,将每 4 个小格子组成一个小块,在图像上的重叠步伐为 $n/2$,再将所有小块级联组成最终的 HOG 描述子。由于颜色在不同的尺度下可以认为是相同的,因此在提取 HOG 特征后对整幅图像提取一个尺度的 HSV 直方图特征,HSV 特征一共由 $8 \times 6 \times 6$ 维向量组成。

为了使特征描述子对尺度变化具有稳健性,对图像进行两次下采样,得到 $1/2$ 原图大小和 $1/4$ 原图大小的两个图像,同时在上述第 4 步中,HOG 中的小格子也按图像的尺度缩小同样的倍数。初始图像大小为 128×48 ,两次下采样后图像的大小分别为 64×24 和 32×12 。初始的 $n \times n$ 为 8×8 ,减小后的大小分别为 4×4 和 2×2 。这样最后的特征描述子的维度为 $15 \times 5 \times 4 \times 9 \times 3 + 8 \times 6 \times 6 = 8388$ 维。

局部特征提取:目前很多文献给出了相关局部特征提取的方法,Zheng 等^[13]提出将行人图像在竖直方向平均分成 6 个水平条带。谭飞刚等^[14]将 4 个数据集里的图像叠加,找到了适合的分割比例,但是对图像进行部件分割会使图像丢失一部分信息。

这里选择在滑动窗口中提取局部特征。滑动窗口的大小选择为 16×16 ,以 8 个像素点为重叠步伐,在大小为 128×48 的图像上滑过。在每个窗口内提取两个尺度的 SILTP(SILTP_{4,4}^{0,3}, SILTP_{4,5}^{0,3},上标为尺度,下标分别为中心像素周围的数量和中心

像素)特征直方图和 11 个维度的 CN 特征。这样在每个窗口内都存在 SILTP 和 CN 融合的局部特征。为了解决视角变化,在水平方向上的所有窗口内选取最大值^[15],这样在水平方向上可以认为仅剩一个滑动窗口。与整体特征相同,对图像进行两次下采样,但是滑动窗口大小不变。这样得到局部特征描述子的维度为 $(3^4 \times 2 + 11) \times (15 + 7 + 3) = 4325$ 维。

3 多特征融合与相似性度量

3.1 相似性度量函数

假设行人 i 的图像使用特征向量 \mathbf{x}_i 表示,行人 j 的图像使用特征向量 \mathbf{x}_j 表示,由 Koestinger 等^[16]提出的 KISSME 算法可知,一对有序行人对之间的相似性可以表示为

$$S(\mathbf{x}_i, \mathbf{x}_j) = \log \frac{P_0(\mathbf{x}_i, \mathbf{x}_j)}{P_1(\mathbf{x}_i, \mathbf{x}_j)}, \quad (1)$$

式中: $P_0(\mathbf{x}_i, \mathbf{x}_j)$ 和 $P_1(\mathbf{x}_i, \mathbf{x}_j)$ 分别代表行人对 $(\mathbf{x}_i, \mathbf{x}_j)$ 为相同行人的概率和不同行人的概率。 $S(\mathbf{x}_i, \mathbf{x}_j)$ 的值越大表明这两个行人图像是同一个行人的概率越大。由文献^[16]可知, $P_0(\mathbf{x}_i, \mathbf{x}_j)$ 和 $P_1(\mathbf{x}_i, \mathbf{x}_j)$ 分别服从均值为零、协方差为 $|\boldsymbol{\Sigma}_s|$ 和 $|\boldsymbol{\Sigma}_D|$ 的多维高斯分布,可分别表示为

$$P_0(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\sqrt{2 \prod |\boldsymbol{\Sigma}_s|}} \exp \left[-\frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \boldsymbol{\Sigma}_s^{-1} (\mathbf{x}_i - \mathbf{x}_j) \right], \quad (2)$$

$$P_1(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\sqrt{2 \prod |\boldsymbol{\Sigma}_D|}} \exp \left[-\frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \boldsymbol{\Sigma}_D^{-1} (\mathbf{x}_i - \mathbf{x}_j) \right], \quad (3)$$

式中:

$$\boldsymbol{\Sigma}_s = \frac{1}{N_{S(i,j) \in S}} \sum (\mathbf{x}_i, \mathbf{x}_j) (\mathbf{x}_i, \mathbf{x}_j)^T, \quad (4)$$

$$\boldsymbol{\Sigma}_D = \frac{1}{N_{D(i,j) \in D}} \sum (\mathbf{x}_i, \mathbf{x}_j) (\mathbf{x}_i, \mathbf{x}_j)^T, \quad (5)$$

其中, N_S 和 N_D 分别表示相对行人对和不同行人对的数量, S 和 D 分别表示相对行人对和不同行人对。把(4)式和(5)式代入(1)式可以得到

$$S(\mathbf{x}_i, \mathbf{x}_j) = \log \frac{1}{\sqrt{2 \prod |\boldsymbol{\Sigma}_s|}} - \log \frac{1}{\sqrt{2 \prod |\boldsymbol{\Sigma}_D|}} + \frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \boldsymbol{\Sigma}_D^{-1} (\mathbf{x}_i - \mathbf{x}_j) - \frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \boldsymbol{\Sigma}_s^{-1} (\mathbf{x}_i - \mathbf{x}_j). \quad (6)$$

(6)式中的常数项不影响相似性函数,因此对(6)式进行简化可以得到相似性度量函数的最终表达式

$$S(x_i, x_j) = (x_i - x_j)^T (\Sigma_D^{-1} - \Sigma_S^{-1})(x_i - x_j). \quad (7)$$

再令 $M = \Sigma_D^{-1} - \Sigma_S^{-1}$, 则 M 为原始空间下的相似性度量矩阵。

由于特征向量在原始特征空间的维度较大,而低维度的特征向量对分类更有利,目前一些算法采用主成份分析(PCA)^[16]等方法降维,但是该方法在降维过程中没有考虑到距离度量问题,因此不是最好的。根据文献[15],在交叉数据集中学习一个子空间 $[W = (w_1, w_2, \dots, w_r) \in \mathbf{R}^{d \times r}]$ 的同时,在 r 维子空间里学习一个距离度量函数。假设有一个 c 个类的交叉训练数据集 (X, Z) , 其中 $X = (x_1, x_2, \dots, x_n) \in \mathbf{R}^{d \times n}$, 维度为 d , 包含 n 个行人, $Z = (z_1, z_2, \dots, z_m) \in \mathbf{R}^{d \times m}$ 与 X 具有相同的维度, 包含 m 个行人。假设要学习的子空间为 W , 则由(7)式算法可以得到在 r 维子空间下的相似性度量函数

$$d_w(x, z) = (x - z)^T W (\Sigma_S'^{-1} - \Sigma_D'^{-1}) W^T (x - z), \quad (8)$$

式中: $\Sigma_S' = W^T \Sigma_S W$, $\Sigma_D' = W^T \Sigma_D W$, 因此最终学习的函数为

$$M(w) = W (\Sigma_S'^{-1} - \Sigma_D'^{-1}) W^T. \quad (9)$$

由于 W 包含在两个求逆矩阵,因此直接优化 d_w 是比较困难的,文献[15]中采用求最大值的方法得到需要的核空间矩阵,表示为

$$\max_w w^T \Sigma_D w, \text{ s.t. } w^T \Sigma_S w = 1, \quad (10)$$

式中:s.t.表示约束条件。但是,文献[15]只是在原始特征空间中训练得到相似性度量矩阵,由于原始空间线性不可分,因此获得的识别效果不理想。这里,先将原始特征空间通过核函数转换到非线性空间,再学习相似性度量矩阵。核函数转换的作用就是将原始特征空间转换到易于区分的非线性空间。将行人特征 x_i 投影到非线性空间的特征为 $\phi(x_i)$, 接着,在非线性空间下同时学习子空间和距离度量函数,进而通过(8)式进行相似性度量。目前常用的核函数主要有3种,分别为 linear, χ^2 , RFB- χ^2 , 通过实验选取结果最好的 RFB- χ^2 核函数。

3.2 多特征融合的度量学习

单一特征无法很好地适应遮挡、光照、姿态等的变换,而多特征可以很好地适应多种变换,因此这里将整体特征和局部特征进行融合。假设提取的整体特征中的 HOG 特征用 f_1 表示, HSV 特征用 f_2 表示。提取的局部特征中的 SILTP 特征用 f_3 表示,

CN 特征用 f_4 表示。则将它们融合后的特征为 $F = [f_1 f_2 f_3 f_4]$ 。最后通过(8)式进行相似性度量。

4 实验测试与结果

4.1 实验数据和算法评价标准

实验中所有数据集中的图片大小都统一为 128×48 。实验的数据随机从3个数据集中选取 k 对行人图像作为测试集,剩下的作为训练集,其中测试集又分成查询集和行人图像库两部分。对每个数据集中的 k 分别随机选取10次,取平均值作为最终的实验结果。表1为不同算法在 VIPeR 数据集上的性能比较。表中所有的实验数据都是在 RFB- χ^2 核函数下得到的结果。在行人再识别算法中普遍采用累积匹配特性(CMC)曲线作为算法性能的评价标准,因此本实验也采用 CMC 曲线作为评价标准,累积匹配特征曲线主要体现了算法在图像库中的搜索能力,描述的是前 r 个搜索结果中包含待查询行人的比率。其中第一匹配率很重要,表征算法性能的好坏。

4.2 不同数据集上的算法比较

实验主要选择在 VIPeR^[17]、iLIDS^[18]、3DPeS^[19] 3个数据集上进行。这3个数据集在姿态、遮挡、光照方面都有很大的变化,而这也正是在实际应用中存在的问题。因此将所提出的方法在这3个数据集上进行测试具有实际的意义,同时也具有很大的挑战性。

VIPeR 数据集中的行人图片来自两个摄像头,行人的姿态、光照变化都比较大, VIPeR 数据集主要由632个人的1264张图片组成。随机从 VIPeR 数据集中选取 $k=316$ 对行人图像进行测试,余下的316对行人图像作为训练集。从表1可以看出,所提出的算法比 PCCA (Principal Component Analysis) 高了21.0%,比 Mid-level filter 高了12.4%,表明所提出的算法在 VIPeR 数据集上具有一定的优势。

表1 不同算法在 VIPeR 数据集上的性能比较

Table 1 Performance comparison of different algorithms on VIPeR dataset

Algorithm	Machining rate / %		
	Rank 1	Rank 10	Rank 20
SDALF ^[1]	19.9	49.4	65.7
Ref. [6]	28.9	59.6	73.1
Ref. [20]	26.2	71.8	82.3
Mid-level filter ^[3]	29.1	66.0	79.9
KISSME ^[16]	23.8	71.0	85.3
PCCA ^[21]	19.3	64.9	80.3
Our algorithm	41.5	81.7	91.8

iLIDS 数据集一共有 119 个行人的 476 张图片, 每个人的图片有 2~8 张不等, 由于这个数据集是在机场采集的, 因此行人图像受到遮挡比较严重。随机选取 $k=55$ 个行人图片用来测试, 然后在每个人的所有图片中随机选取一张组成测试集。在剩下每个行人图像中也随机选取一张组成训练集。表 2 为不同算法在 iLIDS 数据集上的性能比较。从表 2 可以看出, 所提算法在 iLIDS 数据集的测试中也具有一定的优势。

表 2 不同算法在 iLIDS 数据集上的性能比较

Table 2 Performance comparison of different algorithms on iLIDS dataset

Methods	Machining rate / %		
	Rank 1	Rank 10	Rank 20
LFDA ^[22]	32.2	68.7	81.6
rPCCA ^[23]	28.0	71.8	85.9
KISSME ^[16]	28.0	67.9	81.6
Literature ^[20]	40.8	85.3	90.6
Literature ^[14]	39.8	65.6	86.1
Our algorithm	45.3	80.1	91.9

3DPeS 数据集主要由 192 个人的 1011 张图片组成的, 视角变换较大, 每个人的图片有 2~26 张不等。不同算法在 3DPeS 数据集上的性能如表 3 所

示。比较选取的 k 值和 iLIDS 数据集中类似, 这里 $k=100$ 。由表 3 可以看出, 所提算法在 3DPeS 数据集数据集上表现最好。

表 3 不同算法在 3DPeS 数据集上的性能比较

Table 3 Performance comparison of different algorithms on 3DPeS data sets

Methods	Machining rate / %		
	Rank 1	Rank 10	Rank 20
LFDA ^[18]	39.1	71.8	82.6
Literature ^[20]	39.8	86.6	94.4
rPCCA ^[21]	43.5	81.8	91.0
KISSME ^[16]	34.2	69.6	80.2
PCCA ^[21]	39.7	79.6	89.5
Our algorithm	51.2	89.9	96.20

除了与现有算法进行比较外, 还在 VIPeR 数据集上进行了只有局部特征和只有整体特征的实验比较。识别率比较 CMC 曲线图如图 2 所示。只使用 CN 和 SILTP 特征和只使用 HOG 和 HSV 特征进行行人相似性度量以及整体特征和局部特征结合的实验结果相比较如图 2(a) 所示, 从图中可以看出, 当只有 SILTP 和 CN 特征时候效果最差, 而当整体和局部结合时候效果最好。

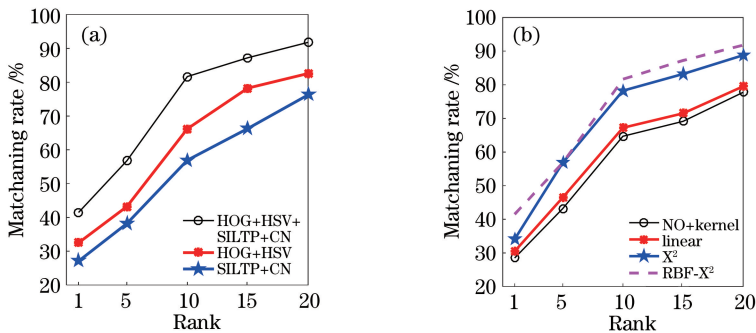


图 2 识别率比较 CMC 曲线图。(a) 特征比较; (b) 核函数比较

Fig. 2 CMC curves for recognition rate comparison. (a) Feature comparison; (b) kernel function comparison

最后进行了有无使用核函数的实验验证。由图 2(b) 可以看出, 使用核函数的实验效果明显优于不使用核函数的效果, 而在 3 个核函数中, RBF- χ^2 的效果最好, 证明了所提算法的有效性。

5 结 论

目前大部分的特征算法均采用局部特征融合或直接采用整体特征融合, 很少采用整体和局部相结合的办法, 而所提出的算法充分考虑了这些问题, 将整体特征和局部特征相结合。在相似性度量方面, 大部分算法采用在原始线性空间下进行度量, 而本算法将原始特征向量通过核函数转换到非线性空间

后, 在非线性空间下同时学习一个子空间和一个距离度量函数。采用学习子空间的方法可以降低特征向量的维度, 同时也避免了 PCA 等降维方法的缺点。从实验结果可以看出, 所提算法超过了大部分算法的第一识别率, 但是识别率还不是很高, 不能满足实际应用需要, 因此下一步将继续优化算法, 提高识别率。

参 考 文 献

- [1] Farenzena M, Bazzani L, Perina A, *et al.* Person re-identification by symmetry-driven accumulation of local features [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern

- Recognition, June 13-18, 2010, San Francisco. New York: IEEE, 2010: 2360-2367.
- [2] Huang X Y, Xu J L, Guo G, *et al.* Real-time pedestrian reidentification based on enhanced aggregated channel features [J]. *Laser & Optoelectronics Progress*, 2017, 54(9): 091001.
黄新宇, 许娇龙, 郭纲, 等. 基于增强聚合通道特征的实时行人重识别[J]. *激光与光电子学进展*, 2017, 54(9): 091001.
- [3] Zhao R, Ouyang W L, Wang X G. Learning mid-level filters for person re-identification [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 23-28, 2014, Columbus. New York: IEEE, 2014: 144-151.
- [4] Liu C X, Gong S G, Loy C C, *et al.* Person reidentification: what features are important? [C] // *Proceedings of the 12th European Conference on Computer Vision*, October 7-13, 2012, Florence. Berlin: Springer, 2012: 391-401.
- [5] Fan C X, Zhu H, Li G F, *et al.* Person re-identification based on multi-features[J]. *Journal of Image and Graphics*, 2013, 18(6): 711-717.
范彩霞, 朱虹, 蔺广逢, 等. 多特征融合的人体目标再识别[J]. *中国图象图形学报*, 2013, 18(6): 711-717.
- [6] Yuan L, Tian Z R. Person re-identification based on multi-feature fusion [J]. *Pattern Recognition and Artificial Intelligence*, 2017, 30(3): 269-278.
袁立, 田子茹. 基于融合特征的行人再识别方法[J]. *模式识别与人工智能*, 2017, 30(3): 269-278.
- [7] Zheng W S, Gong S G, Xiang T. Reidentification by relative distance comparison[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(3): 653-668.
- [8] Qi M B, Hong S S, Wang Y X, *et al.* Multi-feature Subspace and Kernel Learning for Person Re-identification[J]. *Acta Automatica Sinica*, 2016, 42(2): 299-308.
齐美彬, 檀胜顺, 王运侠, 等. 基于多特征子空间与核学习的行人再识别[J]. *自动化学报*, 2016, 42(2): 299-308.
- [9] Xu L, Zhao H T, Sun S Y. Monocular infrared image depth estimation based on deep convolutional neural networks[J]. *Acta Optica Sinica*, 2016, 36(7): 0715002.
许路, 赵海涛, 孙韶媛. 基于深层卷积神经网络的单目红外图像深度估计[J]. *光学学报*, 2016, 36(7): 0715002.
- [10] Zheng Z D, Zheng L, Yang Y. A discriminatively learned CNN embedding for person re-identification [J]. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 2018, 14(1): 13.
- [11] Cheng B, Zha Y F, Li Y Q, *et al.* Robust Convolutional Neural Networks Feature Learning for Person Re-identification [J]. *Acta Optica Sinica*, 2018, 38(7): 0720002.
陈兵, 查宇飞, 李运强, 等. 基于卷积神经网络全批量学习的行人重识别[J]. *光学学报*, 2018, 38(7): 0720002.
- [12] Jobson D J, Rahman Z, Woodell G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes [J]. *IEEE Transactions on Image Processing*, 1997, 6(7): 965-976.
- [13] Zheng W S, Gong S, Xiang T. Person re-identification by probabilistic relative distance comparison[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 20-25, 2011. Colorado Springs, CO, USA. New York: IEEE, 2011: 649-656.
- [14] Tan F G, Huang L, Zhai C, *et al.* A Cross-camera pedestrian re-identification algorithm for major transport hub [J]. *Journal of The China Railway Society*, 2017, 39(1): 76-82.
谭飞刚, 黄玲, 翟聪, 等. 一种用于大型交通枢纽的跨摄像机行人再识别算法研究[J]. *铁道学报*, 2017, 39(1): 76-82.
- [15] Liao S C, Hu Y, Zhu X Y, *et al.* Person re-identification by local maximal occurrence representation and metric learning [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 7-12, 2015, Boston. New York: IEEE, 2015: 2197-2206.
- [16] Koestinger M, Hirzer M, Wohlhart P, *et al.* Large scale metric learning from equivalence constraints [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 16-21, 2012, Providence. New York: IEEE, 2012: 2288-2295.
- [17] Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features [C] // *Proceedings of the 10th European Conference on Computer Vision*, October 12-18, 2008, Marseille. Berlin: Springer, 2008, 5302: 262-275.
- [18] Zheng W S, Gong S G, Xiang T. Associating groups of people [C] // *Proceedings of the 20th British*

- Machine Vision Conference, September 7-10, 2009, London. New York: IEEE, 2009: 1-11.
- [19] 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, December 01, 2011, Scottsdale. New York: ACM, 2011: 59-64.
- [20] Tian Z R, Wang H Y, Chen F H, *et al.* Person re-identification based on distance centralization and projection vectors learning [J]. Journal of Computer Research and Development, 2017, 54 (8): 1785-1794.
丁宗元, 王洪元, 陈付华, 等. 基于距离中心化与投影向量学习的行人重识别 [J]. 计算机研究与发展, 2017, 54(8): 1785-1794.
- [21] Mignon A, Jurie F. Pcca: A new approach for distance learning from sparse pairwise constraints [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence. New York: IEEE, 2012: 2666-2672.
- [22] Pedagadi S, Orwell J, Velastin S, *et al.* Local fisher discriminant analysis for pedestrian reidentification [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 25-27, 2013, Portland. New York: IEEE, 2013: 3318-3325.
- [23] Li Z, Chang S Y, Liang F, *et al.* Learning locally-adaptive decision functions for person verification [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland. New York: IEEE, 2013: 3610-3617.