基于多特征子空间与核学习的行人再识别

齐美彬 [1] 檀胜顺 [1] 王运侠 [1] 刘 皓 [1] 蒋建国 [1]

摘 要 行人再识别指的是在无重叠视域多摄像机监控系统中, 匹配不同摄像机视域中的行人目标. 针对当前基于距离测度学习的行人再识别算法中存在着特征提取复杂、训练过程复杂和识别效果差的问题, 我们提出一种基于多特征子空间与核学习的行人再识别算法. 该算法首先在不同特征子空间中基于核学习的方法得到不同特征子空间中的测度矩阵以及相应的相似度函数, 然后通过比较不同特征子空间中的相似度之和来对行人进行识别. 实验结果表明, 本文提出的算法具有较高的识别率, 其中在 VIPeR 数据集上, RANK1 达到了 40.7 %, 且对光照变化、行人姿态变化、视角变化和遮挡都具有很好的鲁棒性.

关键词 行人再识别, 特征空间, 测度学习, 核学习

引用格式 齐美彬, 檀胜顺, 王运侠, 刘皓, 蒋建国. 基于多特征子空间与核学习的行人再识别. 自动化学报, 2016, **42**(2): 299−308

DOI 10.16383/j.aas.2016.c150344

Multi-feature Subspace and Kernel Learning for Person Re-identification

QI Mei-Bin[1] TAN Sheng-Shun[1] WANG Yun-Xia[1] LIU Hao[1] JIANG Jian-Guo[1]

Abstract Person re-identification is to match pedestrian images observed from different camera views of non-overlapping multi-camera surveillance systems. The current person re-identification based on metric learning is complicated for feature extraction and training process, and has low performance. Therefore, we propose a multi-feature subspace and kernel learning based method for person re-identification. The distance metric and similarity functions can be achieved firstly in different feature subspaces by kernel learning. Then, the object can be recognized by comparing the sum of similarity of different feature subspaces. Experimental results show that the proposed method has a higher accuracy rate, achieving a 40.7 % rank-1 recognition rate on the VIPeR benchmark and that it is robust to different viewpoints, illumination changes, varying poses and the effects of occlusion.

Key words Person re-identification, feature space, metric learning, kernel learning

Citation Qi Mei-Bin, Tan Sheng-Shun, Wang Yun-Xia, Liu Hao, Jiang Jian-Guo. Multi-feature subspace and kernel learning for person re-identification. *Acta Automatica Sinica*, 2016, **42**(2): 299−308

行人再识别指的是在非重叠监控视频中, 检索某个监控视频中出现的目标是否出现在其他的监控视频中. 最近几年, 行人再识别问题引起了广大科研人员的兴趣与研究. 监控视频中的目标图像分辨率低、场景中存在着光照变化、视角变化、行人姿态变化以及摄像机自身属性的问题, 导致同一目标在不同的监控视频中外观区别很大, 使得行人再识别问题遇到了很大的挑战. 为了有效地解决这些挑战, 广大的研究者提出了很多解决方法. 目前的行人再识别算法可以简单概括为 4 种: 直接法[1−7]、基于深度学习的方法[8−9]、显著性学习的方法[10] 和间接法[11−19].

直接法利用视觉特征对行人建立一个鲁棒性和区分性的表示, 不需要通过学习直接利用行人特征, 然后根据传统的相似性度量算法 (欧氏距离、巴氏距离等) 来度量行人之间的相似度. 文献 [1] 首次提出了利用直方图特征来表征目标并通过度量直方图的相似度来识别不同的目标. 文献 [2] 将行人图像粗略分为头部、上半身和下半身 3 部分, 然后串联每一部分的颜色直方图对行人描述. 文献 [3] 采用分割的技术提取人的前景, 并利用行人区域的对称性和非对称性将人的前景划分成不同的区域. 对于每个区域, 提取带权重的颜色直方图特征、极大稳定颜色区域 (Maximally stable color regions) 特征和重复度高的结构区域 (Recurrent highly structured patches) 特征描述它们. 文献 [4] 提出了一种结合 Gabor 特征和协方差矩阵描述的 BiCov 描述子来对行人描述. 文献 [5] 采用图案结构 (Pictorial structure) 算法定位图像中人的各个部件所在的区域. 对于每个部件的区域, 提取与文献 [3] 类似的颜色直方图特

收稿日期 2015-05-28 录用日期 2015-11-11
Manuscript received May 28, 2015; accepted November 11, 2015
国家自然科学基金 (61371155), 安徽省科技攻关项目 (1301b042023) 资助
Supported by National Natural Science Foundation of China (61371155), Science and Technology Brainstorm Project of Anhui Province (1301b042023)
本文责任编委 王亮
Recommended by Associate Editor WANG Liang
1. 合肥工业大学计算机与信息学院 合肥 230009
1. School of Computer and Information, Hefei University of Technology, Hefei 230009

征、极大稳定颜色区域特征来描述它们. 文献 [6] 利用 LBP (Local binary pattern) 特征和 21 个滤波器特征 (8 个 Gabor 滤波器和 13 个 Schmid 滤波器) 来描述图像中的行人. 文献 [7] 通过共生矩阵对行人的形状和外形特征进行描述. 虽然直接法的算法模型简单, 但由于在光照变化、视角变化以及姿态变化等情况下, 同一个人的外观变化往往很大, 很难提取出鲁棒性和区分性的特征描述. 因此在光照变化、视角变化以及姿态变化等情况下, 直接法的效果很差.

深度学习近年来在计算机视觉中得到了广泛的应用, 因此不少学者研究并提出了基于深度学习的行人再识别算法. Li 等[8] 提出了一种 6 层的 FPNN (Filter pairing neural network) 神经网络, 它能有效解决行人再识别中出现的光照变化、姿态变化、遮挡和背景粘连等问题, 从而提高了识别率. Ahmed 等[9] 提出了一种深层卷积结构能够同时自动学习特征和相应的相似性测度函数. 但基于深度学习的行人再识别算法需要非常大的训练数据库, 导致训练时间长, 此外还需要针对特定问题搭建相应的模型, 因此不利于该方法的推广. 同时, 深度学习方法中还存在调置参数缺乏理论性指导, 具有较大主观性的问题.

基于显著性学习的行人再识别方法近年来也受到研究者的广泛兴趣. 人们可以通过行人的一些显著信息来识别行人, 但传统的方法在比较两张图片的相似性的时候, 往往忽略了行人身上的显著性特征. 对此, Zhao 等[10] 提出了一种通过学习行人图像的显著性信息来度量两张行人图像的相似性的方法. 但显著性学习的行人再识别算法在行人姿态变化的情况下, 显著性区域会出现偏移或者消失, 导致识别效果较差.

间接法主要是学习得到一个分类器或一个排序模型. 间接法代表性的算法有距离测度学习[11−16]、支持向量机[17]、迁移学习[18] 和流形排序算法[19]. 距离测度学习算法作为间接法中的一种, 最近几年在行人再识别中得到了广泛的应用. 本文提出的算法也是基于距离测度学习, 所以下面着重介绍基于距离测度学习的行人再识别算法.

与手动设计特征的直接法不同, 距离测度学习方法是一种利用机器学习的算法得出两张行人图像的相似度度量函数, 使相关的行人图像对的相似度尽可能高, 不相关的行人图像对的相似度尽可能低的方法. 代表性的测度学习算法有 Zheng 等[11] 把行人再识别问题当成距离学习问题, 提出了一种基于概率相对距离的行人匹配模型, 文献 [12] 提出了一种基于统计推断的方法学习测度矩阵来度量两张行人图像的相似度, 以及文献 [13−16] 中提出的相应测度学习算法.

距离测度算法是将原始特征空间投影到另一个更具区分性的特征空间. 与其他算法相比, 距离测度学习算法具有更好的效果. 距离测度学习算法即使只使用简单的颜色直方图作为特征, 算法的性能往往优于其他算法. 文献 [12] 提出了一种基于统计推断的方法学习测度矩阵来度量行人对的相似度, 但作者直接在原始特征空间训练得到测度矩阵, 进而得到样本之间的相似性函数. 原始特征空间的线性不可分性导致通过原始特征空间直接训练得到的测度矩阵不能很好地表征样本之间的相似性和差异性. 本文提出基于核学习的方法, 首先, 通过相应的核函数将原始特征空间投影到非线性空间; 然后, 在非线性空间中学习得到相应的测度矩阵. 投影后的非线性特征空间具有很好的可分性, 这样学习得到的测度矩阵能准确地表征样本之间的相似性和差异性. 另外, 基于测度学习的行人再识别算法一般是把多特征融合并建立特征模型, 然后基于相应的测度学习算法, 学习得到一个测度矩阵. 然而这种特征融合往往忽略了不同属性特征之间的差别, 这样学习得到的测度矩阵不能准确地表征样本之间的相似性与差异性. 对此, 本文提出在不同的特征空间中学习相应的测度矩阵, 进而得到表示不同特征空间的相似性函数, 最后根据不同的权重结合这些相似性函数来表征样本之间的相似性. 本文算法在公共实验数据集上的实验效果优于目前主流的行人再识别算法, 尤其是第一匹配率 (Rank1).

本文其余章节的组织安排如下. 第 1 节介绍本文提出的行人再识别算法; 第 2 节介绍本文算法在公共数据集上的实验; 第 3 节总结全文以及展望.

1 基于多特征子空间与核学习的行人再识别算法

1.1 基于核学习的相似度量函数的学习

文献 [13] 中提出了一种 KISSME 的算法, 文中指出, 从统计学角度考虑, 一对有序行人对 (i, j) 的相似度可以表示为式 (1).

$$S(i,j) = \log\left(\frac{P_0(i,j)}{P_1(i,j)}\right) \tag{1}$$

其中, $P_0(i,j)$ 表示一对有序行人对 (i,j) 属于相关行人对的概率, $P_1(i,j)$ 表示一对有序行人对 (i,j) 属于不相关行人对的概率. S 的值越大, 表示有序行人对 (i,j) 属于相关行人对的概率更大.

设行人 i 是用特征 $\boldsymbol{x}_i$ 表示, 行人 j 同样用特征

$\boldsymbol{x}_j$ 表示, 则式 (1) 投影到特征空间表示为式 (2).

$$S\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)=\log\left(\frac{P_0(\boldsymbol{x}_i, \boldsymbol{x}_j)}{P_1(\boldsymbol{x}_i, \boldsymbol{x}_j)}\right) \tag{2}$$

根据文献 [13] 可知, 相关的行人对和不相关的行人对在特征空间中分别服从均值为 0, 协方差矩阵 $\sum_S$ 和 $\sum_D$ 的多维正态分布. 它们分别表示为式 (3) 和式 (4).

$$P_0(\boldsymbol{x}_i, \boldsymbol{x}_j)=\frac{1}{\sqrt{2\Pi\left|\Sigma_S\right|}}\exp\Big(-\frac{1}{2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\Sigma_S^{-1}(\boldsymbol{x}_i-\boldsymbol{x}_j)\Big) \tag{3}$$

$$P_1(\boldsymbol{x}_i, \boldsymbol{x}_j)=\frac{1}{\sqrt{2\Pi\left|\Sigma_D\right|}}\exp\Big(-\frac{1}{2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\Sigma_D^{-1}(\boldsymbol{x}_i-\boldsymbol{x}_j)\Big) \tag{4}$$

其中

$$\Sigma_S=\frac{1}{N_S}\sum_{(i,j)\in S}(\boldsymbol{x}_i-\boldsymbol{x}_j)(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}$$

$$\Sigma_D=\frac{1}{N_D}\sum_{(i,j)\in D}(\boldsymbol{x}_i-\boldsymbol{x}_j)(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}$$

将式 (3) 和式 (4) 代入到式 (2), 得到式 (5).

$$S(\boldsymbol{x}_i, \boldsymbol{x}_j)=\log\left(\frac{1}{\sqrt{2\pi\left|\Sigma_S\right|}}\right)-\log\left(\frac{1}{\sqrt{2\pi\left|\Sigma_D\right|}}\right)+\frac{1}{2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\Sigma_D^{-1}(\boldsymbol{x}_i-\boldsymbol{x}_j)-\frac{1}{2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\Sigma_S^{-1}(\boldsymbol{x}_i-\boldsymbol{x}_j) \tag{5}$$

相似度函数 $S(\boldsymbol{x}_i, \boldsymbol{x}_j)$ 中常数不影响相似度的度量, 则省略相似度函数中的常数和常数系数, 将式 (5) 化简得到式 (6).

$$S(\boldsymbol{x}_i, \boldsymbol{x}_j)=(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\Sigma_D^{-1}(\boldsymbol{x}_i-\boldsymbol{x}_j)-(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\times\Sigma_S^{-1}(\boldsymbol{x}_i-\boldsymbol{x}_j)=(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\left(\Sigma_D^{-1}-\Sigma_S^{-1}\right)(\boldsymbol{x}_i-\boldsymbol{x}_j) \tag{6}$$

其中, 令 $M=\Sigma_D^{-1}-\Sigma_S^{-1}$.

文献 [13] 中提出的算法是直接在原始线性特征空间中训练得到测度矩阵, 进而得到表示样本之间相似性的相似度函数. 由于原始特征空间的线性不可分, 上述方法得到的测度矩阵不能准确表达样本之间的相似性和差异性, 导致识别效果差. 本文提出基于核学习的算法首先通过相应的核函数将原始特征空间投影到更易区分的非线性空间, 然后在非线性空间中训练得到测度矩阵 M. 这样得到的测度矩阵 M 具有很好的区分性, 能使同类样本之间的距离尽可能小, 异类样本之间的距离尽可能大.

核学习的主要思想是将原始线性特征空间投影到区分性好的非线性空间. 原始特征空间中的特征 $\boldsymbol{x}_i$ 通过函数 Φ 投影到非线性空间, 则非线性空间的特征表示为 $\Phi(\boldsymbol{x}_i)$. 非线性映射函数一般是隐性函数, 则很难得到显式表达式, 可以利用核函数求解特征空间中样本点的内积来解决, 如式 (7) 所示.

$$k(\boldsymbol{x}, \boldsymbol{z})=\langle\Phi(\boldsymbol{x}), \Phi(\boldsymbol{z})\rangle \tag{7}$$

其中, $\boldsymbol{x}, \boldsymbol{z}\in\mathbf{R}^d$ (d 表示原始特征空间特征的维度). 通过式 (7) 得式 (8), 式 (8) 把原始特征空间中任一特征 $\boldsymbol{x}_q$ 投影到易区分的非线性特征空间中得到非线性特征 $\boldsymbol{k}_q$.

$$\boldsymbol{k}_q=[k(\boldsymbol{x}_1, \boldsymbol{x}_q), k(\boldsymbol{x}_2, \boldsymbol{x}_q), \cdots, k(\boldsymbol{x}_n, \boldsymbol{x}_q)] \tag{8}$$

式中, $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 表示原特征空间中的 n 个样本特征. 式 (6) 通过核函数变换得到式 (9).

$$S(\boldsymbol{k}_i, \boldsymbol{k}_j)=(\boldsymbol{k}_i-\boldsymbol{k}_j)^{\mathrm{T}}\Sigma_D^{-1}(\boldsymbol{k}_i-\boldsymbol{k}_j)-(\boldsymbol{k}_i-\boldsymbol{k}_j)^{\mathrm{T}}\times\Sigma_S^{-1}(\boldsymbol{k}_i-\boldsymbol{k}_j)=(\boldsymbol{k}_i-\boldsymbol{k}_j)^{\mathrm{T}}M(\boldsymbol{k}_i-\boldsymbol{k}_j) \tag{9}$$

此时

$$\Sigma_S=\frac{1}{N_S}\sum_{(i,j)\in S}(\boldsymbol{k}_i-\boldsymbol{k}_j)(\boldsymbol{k}_i-\boldsymbol{k}_j)^{\mathrm{T}} \tag{10}$$

$$\Sigma_D=\frac{1}{N_D}\sum_{(i,j)\in D}(\boldsymbol{k}_i-\boldsymbol{k}_j)(\boldsymbol{k}_i-\boldsymbol{k}_j)^{\mathrm{T}} \tag{11}$$

$$M=\Sigma_D^{-1}-\Sigma_S^{-1} \tag{12}$$

上述流程首先将原始线性特征空间通过核函数 k 投影到易区分的非线性空间, 然后在投影后的特征空间训练得到测度矩阵.

常用的核函数有 linear、χ^2 和 RBF-χ^2 (Radial basic function-χ^2)[20]. 经过后面的实验论证, 三种核函数的实验效果为 RBF- $\chi^2>\chi^2>$ linear.

1.2　基于多特征子空间的测度学习

基于测度学习的行人再识别算法一般是把多特征融合并建立特征模型, 然后基于相应的测度学习算法得到测度矩阵. 这种方法忽略了不同属性特征之间的差别, 导致学习得到的测度矩阵不能准确地体现样本之间的相似性与差异性. 对此, 本文提出对于不同的特征空间单独学习相应的测度矩阵, 进而得到表示不同特征空间的相似性函数, 最后根据不同的权重结合这些相似性函数来表示样本之间的相似性.

假设有特征空间 C 和特征空间 T, 特征空间 C 中样本对 (i, j) 的特征表示为 $(\boldsymbol{c}_i, \boldsymbol{c}_j)$ 以及特征空间

T 中样本对 (i,j) 的特征表示为 $(\boldsymbol{t}_i, \boldsymbol{t}_j)$. 特征空间 C 中的任一特征 $\boldsymbol{c}_q$ 通过式 (8) 投影到核空间的特征为 $\boldsymbol{k}_{cq} = [k(\boldsymbol{c}_1, \boldsymbol{c}_q), k(\boldsymbol{c}_2, \boldsymbol{c}_q), \cdots, k(\boldsymbol{c}_n, \boldsymbol{c}_q)]$, 同理特征空间 T 中的任一特征 $\boldsymbol{t}_q$ 通过式 (8) 投影到核空间的特征为 $\boldsymbol{k}_{tq} = [k(\boldsymbol{t}_1, \boldsymbol{t}_q), k(\boldsymbol{t}_2, \boldsymbol{t}_q), \cdots, k(\boldsymbol{t}_n, \boldsymbol{t}_q)]$. 依据第 1.1 节中提出的算法, 在特征空间 C 对应的核空间和特征空间 T 对应的核空间分别训练得到测度矩阵 M_C 和 M_T, 进而分别得到两种核空间的相似度函数 S_C 和 S_T. 最后样本之间相似度函数为

$$S = aS_c + (1-a)S_T$$

a 和 $1-a$ 为不同特征空间相似度函数的权值, 通过在后面不同数据集上的实验论证, 当 a 满足条件 $a \in [0.5, 0.7]$ 时, 算法性能达到最优. 本文实验中权值统一设置为 $a = 0.5$.

1.3 行人图像的特征表示

本文采用颜色特征和 LBP 特征对行人目标进行描述, 生成两种特征子空间. 颜色空间有很多种, 用不同的色彩空间描述一图片的效果是不同的. 依据文献 [21−22], 本文采用的颜色特征从 RGS、HSV、YCbCr 和 CIE Lab 4 种颜色空间中提取. RGS 空间的定义为 $R = R/(R+G+B)$、$G = G/(R+G+B)$、$S = (R+G+B)/3$. 为了获取具有鲁棒性和区分性的颜色特征表示, 本文将 4 种颜色空间的特征融合. 融合后的颜色特征描述对于光照变化的场景具有良好的鲁棒性. 图 1 给出了一张被平均分成 6 个水平条带的行人图像, 对于每个水平条带的每种颜色空间的每个通道, 提取 16 维的颜色直方图特征, 将所有的颜色直方图特征串联形成行人图像的颜色特征, 从而得到行人图像颜色特征的维数为 1 152 维 $(4 \times 6 \times 3 \times 16)$. 原始特征空间的 1 152 维特征经过核函数投影后的特征维度较高且大部分信息都是冗余的. 因此实验中利用 PCA (Principle component analysis) 将核空间中的特征维数降到保持大于 90 % 的能量.

为了更好地描述图像的局部信息, 本文中提取的局部特征来自于 LBP 等价模式中邻域点数为 8 半径为 1 和邻域点数为 16 半径为 2 两种模式的特征. 上面两种 LBP 等价模式的维度分别为 59 维和 243 维. 类似于颜色特征的提取, 一张行人图像被平均分成 6 个水平条带, 对于每个水平条带提取 LBP 两种模式的特征, 将所有的 LBP 特征串联形成行人图像的局部特征, 从而得到行人图像局部特征的维数为 1 812 维 $((59+243) \times 6)$. 同理, 原始特征空间的 1 812 维特征经过核函数投影后的特征维度同样较高并且大部分信息也都是冗余的. 因此实验中同样利用 PCA 将核空间的特征维度降到保持 90 % 以上的能量.

图 1 行人图像分成 6 个无重叠的水平条带

Fig. 1 6 non-overlapping horizontal bands divided by pedestrians image

1.4 本文算法的具体操作步骤

步骤 1. 输入具有 N 对相关行人的训练集, 并提取每对相关行人的颜色特征 $(\boldsymbol{c}_{1i}, \boldsymbol{c}_{2i})$ 和 LBP 特征 $(\boldsymbol{t}_{1i}, \boldsymbol{t}_{2i})$, $i = 1, 2, \cdots, N$.

步骤 2. 将每对相关行人的颜色特征和 LBP 特征基于相应的核函数通过第 1.1 节中式 (8) 分别投影到可分的非线性空间. 经过核函数投影的颜色特征和 LBP 特征在投影后的空间中分别表示为 $(\boldsymbol{k}c_{1i}, \boldsymbol{k}c_{2i})$ 和 $(\boldsymbol{k}t_{1i}, \boldsymbol{k}t_{2i})$.

步骤 3. 按照第 1.1 节中式 (10) ∼ 式 (12) 学习得到颜色子空间和 LBP 子空间对应核空间中的测度矩阵.

步骤 4. 利用步骤 3 得到不同核空间中的测度矩阵计算样本在不同核空间的相似度, 然后将样本在不同核空间中的相似度按照一定的权值结合来表示样本之间的相似性.

2 实验测试与结果

本节首先介绍实验中所使用的测试数据和算法性能的评测准则, 其次介绍本文算法在不同公共实验集上与已有的行人再识别算法的性能比较, 然后在不同公共实验集上对比核映射前后的算法性能, 最后在不同公共实验集上分析权值不同时对算法性能的影响. 文中所有的实验是基于 vs2010+opencv2.4.9 实现的, 实验平台是 24 GB 内存的 Intel Xeon CPU E5506 2.13 GHz (2 处理器) PC 台式机.

2.1 测试数据和算法性能的评测准则

为了与已有算法公正比较, 实验中, 采用先前工作普遍采用的评价框架. 如文献 [12] 所述, 随机选择

p 对行人图像对作为测试集, 余下的行人图像对作为训练集. 测试集由查询集和行人图像库两部分组成. 每对行人图像, 随机选择一张图像并入查询集, 另一张则并入行人图像库. 当给定一个行人再识别算法, 衡量该算法在行人图像库中搜索待查询行人的能力来评测此算法的性能. 为了测试算法在只有少量的训练样本时的性能, p 分别取 316、432、532 进行实验. p 取值越大, 则测试样本越多 (匹配越困难), 训练样本越少 (训练越困难). 对于每组实验, 以上产生测试集和训练集的过程重复 10 次, 10 次实验结果的平均值作为本组实验的结果. 已有的行人再识别算法大部分采用累积匹配特性 (Cumulative match characteristic, CMC) 曲线评价算法性能. 给定一个查询集和行人图像库, 累积匹配特征曲线描述的是在行人图像库中搜索待查询的行人, 前 r 个搜索结果中找到待查询人的比率. 第 1 匹配率 ($r=1$) 很重要, 因为它表示的是真正的识别能力. 但是当 r 取值很小时, 第 r 匹配率也很有现实意义. 因为在实际应用中, 反馈的前 r 个搜索结果中, 可以通过人眼判断找到查询目标.

2.2　不同公共实验集上实验对比

本文算法选择在 VIPeR[23] 数据集、iLIDS[24] 数据集、ETHZ[25] 数据集和 CUHK01[26] 数据集上进行实验.

VIPeR 数据集中相关行人对来自两个不同的摄像机. 相关行人对的外观由于姿态、视角、光照和背景的变化而差异大. iLIDS 数据集包含 119 个行人的 476 张图像, 每个行人的图像从 2 张到 8 张不等. iLIDS 数据集是从机场收集的, 数据集有严重遮挡和光照变化的问题. ETHZ 数据集包含 146 个行人的 8 555 张图像. 数据集中的图像来自移动摄像机拍摄的三个视频序列. ETHZ 数据集中的图像由同一摄像机拍摄, 则数据集中行人姿态变化和视角变化的程度没有 VIPeR 数据集那么明显. 为了使 ETHZ 数据集具有挑战性, 实验中相关行人提取两张图像, 一张近景和一张远景. 远景图像含有大量无用的背景信息, 使选用的数据集有明显的遮挡情况和光照变化. CUHK01 数据集是由两个摄像机在校园环境中拍摄得到, 包含 971 个行人的 3 884 张图像. 每个行人包含 4 张图像, 前两张图像是摄像机拍摄的行人前后景图像, 后两张图像是摄像机拍摄的行人侧面图像. 每个行人的前两张图像只是姿态上有稍许的变化, 前两张图像与后两张图像在视角上、姿态上差异较大以及有明显的光照变化. 实验中, 每个行人前后景图像随机选择一张, 侧面图像随机选择一张, 这样得到的实验集具有显著的视角变化、姿态变化和光照变化.

2.2.1　**VIPeR 数据集**

VIPeR 数据集是由 632 对相关行人对图像组成. 如图 2 所示, 同一列中的行人图像为同一个人.

图 2　每列为来自不同摄像头场景的同一个人

Fig. 2　Each column describes the same person captured by different cameras

为了对比本文算法基于不同核函数的实验效果, 表 1 给出了该算法基于不同核函数的实验对比. 实验中测试样本集和训练样本集均为 316 对相关行人图像. VIPeR 数据集上的其他实验, 如果没明确表明测试样本集和训练样本集的个数, 都默认有 316 对相关行人图像.

表 1　本文算法基于不同核函数在 VIPeR 数据集上的识别率 (%)

Table 1　Mathing rates (%) of the proposed algorithm based on different kernel functions on the VIPeR dataset

Kernel	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
linear	25.1	53.4	67.3	80.1
χ^2	38.2	70.0	82.5	91.3
RBF-χ^2	40.7	72.37	83.95	92.08

从表 1 可知, 本文算法基于 RBF-χ^2 核函数的效果最优. 为了充分体现算法的效果, 在后面的实验效果对比中都是基于 RBF-χ^2 核函数. 表 2 给出了该算法与当前主流算法的效果对比.

从表 2 可知, 该算法性能有较大的提升, 尤其是 Rank1, Rank1 比表 2 中最好的结果提高了约 8 (%). 此外, 该算法与表中的 PCCA (Pairwise constrained component analysis)、rPCCA、kLFDA (Local Fisher discriminant analysis) 和 MFA (Marginal Fisher analysis) 等算法都是基于 RBF-χ^2 核函数; 但该算法整体效果明显优于它们. 值得一提的是, 对于 Rank1, 该算法的效果相对主流的行人再识别算法有了显著的提高. 第 1 匹配率很重要, 因为它表示的是真正的识别能力. 行人再识别技术一个典型的应用是刑事侦查; 若 Rank1 越高, 则在刑事侦查中, 搜集与嫌疑人有关的线索的效率就会提高.

表 2 不同算法在 VIPeR 数据集上的识别率 (%)

Table 2 Mathing rates (%) of different methods on the VIPeR dataset

Methods	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
PCCA[14]	19.6	51.5	68.2	82.9
LFDA[27]	19.7	46.7	62.1	77.0
SVMML[28]	27.0	60.9	75.4	87.3
KISSME[13]	23.8	54.8	71.0	85.3
文献 [12]	29.7	59.8	73.0	84.1
rPCCA[20]	22.0	54.8	71.0	85.3
kLFDA[20]	32.3	65.8	79.7	90.9
MFA[20]	32.2	66.0	79.7	90.6
RDC[29]	15.66	38.42	53.86	70.09
eSDC_knn[10]	26.31	46.61	58.86	72.77
eSDC_ocsvm[10]	26.74	50.70	62.37	76.36
Ours	40.7	72.37	83.95	92.08

为了充分说明本文算法的优越性, 表 3 给出该算法在仅用 HSV (Hue, saturation, value) 特征情况下与其他算法效果对比.

表 3 不同算法在 VIPeR 数据集上的识别率 (%)

Table 3 Mathing rates (%) of different methods on the VIPeR dataset

Methods	Features	Rank1	Rank10	Rank20
SDALF[3]	HSV, structures	20	49	56
PS[5]	HSV, structures	22	57	71
RDC[29]	HSV, YCbCr, texture	16	54	70
KISSME[13]	HSV, Lab, LBP	20	62	77
ITML[30]	HSV	15	50	66
Euclidean	HSV	7	23	34
NRDV[25]	HSV	25	65	78
KRMCA[31]	HSV	23.2	72.2	85.8
Ours	HSV	28.4	74.1	86.9

由表 3 可知, 该算法虽然只使用了 HSV 特征, 但是效果比采用多特征的算法 (SDALF (Symmetry-driven accumulation of local features)、PS (Pictorial structures)、RDC (Relative distance comparison) 和 KISSME) 更好. KISSME 融合了 HSV、LAB 和 LBP 等特征, Rank1 仅有 20 (%); 该算法只用 HSV 特征, Rank1 就达到了 28.4 (%). 另外, 该算法与同样只使用 HSV 特征的算法 (如 ITML (Information-theoretic metric learning)、Euclidean、NRDV (Nonlinear Ranking with difference vectors) 和 KRMCA (Kernelized relaxed Margin components analysis) 等) 相比, 仍然优于它们. 其中, ITML 和 Euclidean 算法整体效果都比该算法差. NRDV 算法虽然 Rank1 与该算法相近, 但是 Rank10 和 Rank20 较低, 且 NRDV 算法模型比该算法复杂得多. KRMCA 算法效果总体上都不如该算法且 KRMCA 的代价函数收敛很慢, 算法训练时间很长.

当测试集规模为 $P = 432$ 和 $P = 532$ 时, 该算法与已有行人再识别算法的性能比较如表 4 和表 5. 从表中可知, 在只有少量的训练样本情况下, 该算法性能同样优于已有算法. 由此可见, 该算法有效解决了学习相似度度量函数中出现的过拟合问题.

表 4 当 $P = 432$, 不同算法在 VIPeR 数据集上的识别率 (%)

Table 4 Mathing rates (%) of different methods at $P = 432$ on the VIPeR dataset

Methods	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
Euclidean	4.8	11.5	16.4	23.2
KISSME[13]	17.6	42.6	56.6	71.5
PRDC[11]	12.6	32	44.3	60
ITML[30]	8.4	24.5	36.8	52.3
LMNN[32]	5.1	13.1	20.3	33.9
文献 [12]	22.5	48.6	61.4	74.4
Ours	28.7	59.3	72.7	83.1

表 5 当 $P = 532$, 不同算法在 VIPeR 数据集上的识别率 (%)

Table 5 Mathing rates (%) of different methods at $P = 532$ on the VIPeR dataset

Methods	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
Euclidean	4.0	10.3	14.8	20.9
KISSME[13]	11.3	29.4	42.1	56.2
PRDC[11]	9.1	24.2	34.4	48.6
ITML[30]	4.2	11.1	17.2	24.6
LMNN[32]	4.0	9.7	14.2	21.2
文献 [12]	12.4	31.1	43.0	56.7
Ours	15.5	36.6	49.2	62.1

2.2.2 iLIDS 数据集

iLIDS 数据集中每个行人包含 2 张到 8 张照片不等. 实验中, 从每个行人所包含的图像中随机取两张作为实验集, 最后得到的实验集为 119 对相关行人图像. 最终实验效果是多次随机取得的实验集效果的平均值.

数据集中的图像尺寸是不尽相同的, 实验中统一把图像的尺寸设置为高 128 像素, 宽 48 像素. 实验中训练集为 59 对行人图像, 测试集为 60 对行人图像. 本文算法在 iLIDS 数据集上与其他算法的效果对比如表 6.

该算法与表 6 中基于测度学习的算法 PCCA、rPCCA、MFA 和 kLFDA 都是基于 RBF-χ^2 核函数. 但从表 6 可知, 该算法整体性能优于 PCCA、rPCCA 和 MFA 等算法; 该算法虽然与 kLFDA 算法性能接近, 但总体上还是优于 kLFDA 算法. 由此可见, 该算法比使用相同核函数的其他算法效果更好. 另外, 该算法整体性能也优于 KISSME、SVMML 和 LFDA 算法.

表 6 不同算法在 iLIDS 数据集上的识别率 (%)
Table 6 Mathing rates (%) of different methods on the iLIDS dataset

Methods	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
KISSME[13]	28.0	54.2	67.9	81.6
PCCA[14]	24.1	53.3	69.2	84.8
LFDA[27]	32.2	56.0	68.7	81.6
SVMML[28]	20.8	49.1	65.4	81.7
rPCCA[20]	28.0	56.5	71.8	85.9
kLFDA[20]	36.9	65.3	78.3	89.4
MFA[20]	32.1	58.8	72.2	85.9
ours	38.3	66.5	79.0	88.3

2.2.3 ETHZ 数据集

数据集中的图像尺寸是不相同的, 实验中统一把图像尺寸设置为高 128 像素, 宽 48 像素. 实验中训练集为 76 对行人图片, 测试集为 70 对行人图片. 表 7 分析了本文算法在 ETHZ 数据集上与其他算法的效果对比.

表 7 不同算法在 ETHZ 数据集上的识别率 (%)
Table 7 Mathing rates (%) of different methods on the ETHZ dataset

Methods	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
KISSME[13]	48.6	65.2	76.4	87.8
PCCA[14]	40.2	64.4	76.1	88.5
LFDA[27]	52.8	68.3	78.1	90.8
SVMML[28]	37.5	65.8	77.6	90.6
rPCCA[20]	45.5	65.6	76.3	90.1
kLFDA[20]	53.5	73.3	82.6	91.5
MFA[20]	52.6	70.2	79.3	90.1
Ours	61.09	74.77	81.96	91.76

从表 7 可知, 该算法在 ETHZ 数据集上的整体性能优于同样基于 RBF-χ^2 核函数的 PCCA、rPCCA、MFA 和 kLFDA 等算法. 在 ETHZ 数据集上, 同样证明了该算法比使用相同核函数的其他算法效果更好. 值得一提的是, 该算法的 Rank1 较于其他算法显著提升了. 另外, 该算法整体性能也优于 KISSME、SVMML 和 LFDA 算法.

2.2.4 CUHK01 数据集

数据集中的图片的尺寸是不相同的, 实验中统一把图片的尺寸设置为高 128 像素, 宽 48 像素大小. 实验中训练集为 486 对行人图片, 测试集为 485 对行人图片. 本文算法在 CUHK01 数据集上与其他算法的效果对比如下表 8.

从表 8 可知, 该算法在 CUHK01 数据集上的整体性能同样优于 KISSME 和 SVMML 算法以及基于测度学习的算法 PCCA、LFDA、rPCCA、MFA 和 kLFDA. 该算法与 MidLevel 算法效果接近, 但 MidLevel 算法模型复杂. 该算法与其他算法效果对比可知, 该算法可以学习得到具有良好区分性的相似性度量函数. 通过在 CUHK01 数据集上的效果对比, 进一步说明该算法与使用相同核函数的其他算法相比效果更好.

表 8 不同算法在 CUHK01 数据集上的识别率 (%)
Table 8 Mathing rates (%) of different methods on the CUHK01 dataset

Methods	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
KISSME[13]	12.5	31.5	42.5	54.9
PCCA[14]	17.8	42.4	55.9	69.1
LFDA[27]	13.3	31.1	42.2	54.3
SVMML[28]	18.0	42.3	55.4	68.8
rPCCA[20]	21.6	47.4	59.8	72.6
kLFDA[20]	29.1	55.2	66.4	77.3
MFA[20]	29.6	55.8	66.4	77.3
MidLevel[26]	34.30	55.74	64.52	74.97
Ours	36.10	62.68	72.61	81.90

2.3 特征核映射前后算法性能的比较

为了分析特征经过核映射后对算法的影响, 表 9～表 12 分别给出在 4 种数据集上特征经过核映射前后算法效果的对比实验.

表 9 特征核映射前后在 VIPeR 实验集上的对比效果
Table 9 Performance comparison between before and after the kernel map on the VIPeR dataset

Kernel	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
Before	25.66	53.43	67.4	80.5
After	40.7	72.37	83.95	92.08

表 10 特征核映射前后在 iLIDS 实验集上的对比效果
Table 10 Performance comparison between before and after the kernel map on the iLIDS dataset

Kernel	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
Before	30.4	56.4	67.4	78.3
After	38.3	66.5	79.0	88.3

表 11 特征核映射前后在 ETHZ 实验集上的对比效果
Table 11 Performance comparison between before and after the kernel map on the ETHZ dataset

Kernel	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
Before	57.0	72.97	81.5	90.4
After	61.09	74.77	81.96	91.76

表 12 特征核映射前后在 CUHK01 实验集上的对比效果
Table 12 Performance comparison between before and after the kernel map on the CUHK01 dataset

Kernel	Rank1 (%)	Rank5 (%)	Rank10 (%)	Rank20 (%)
Before	17.92	38.10	48.0	58.84
After	36.10	62.68	72.61	81.90

通过在 4 个公共数据集上实验对比可知, 特征经过核映射后算法效果在 VIPeR、iLIDS 和 CUHK01 数据集上整体上都得到了显著的提升, 在 ETHZ 数据集上虽然提高不明显, 但还是优于已有算法. 总的来说, 该算法在特征经过核映射后, 学习得到的相似度度量函数更具有区分性, 能够得到较好的识别效果.

2.4 权值 a 取值不同时算法的性能比较

为了分析权值 a 对算法性能的影响, 图 3～图 6 分别给出在 4 种数据集上不同的权值 a 下, 本文算法性能的对比实验.

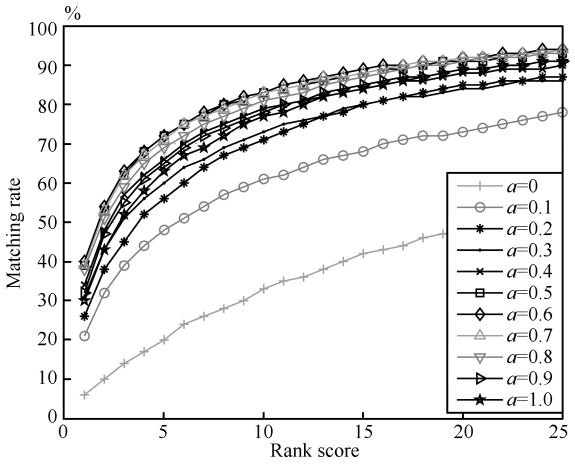

图 3 算法在 VIPeR 数据集上在不同权值 a 下的性能比较

Fig. 3 Performance comparison at different weights a on the VIPeR dataset

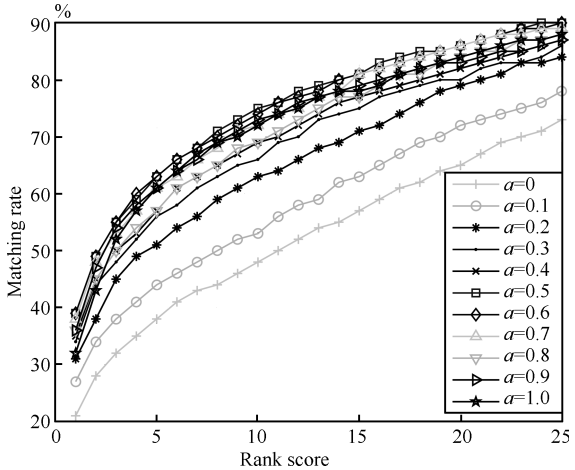

图 4 算法在 iLIDS 数据集上在不同权值 a 下的性能比较

Fig. 4 Performance comparison at different weights a on the iLIDS dataset

本文算法学习得到的相似度函数为

$$S = aS_C + (1-a)S_T$$

其中, S_C 为颜色特征子空间的相似度函数, S_T 为 LBP 特征子空间的相似度函数. 权值 a 越大, 代表相似度函数中颜色特征子空间的相似度函数比重越大.

通过在 4 种数据集上不同的权值 a 下算法性能的对比实验可知, a 取值对算法效果的影响较大. 当 a 取值很小时, 算法效果不是很理想, 当 a 增大时, 算法性得到一定程度的提升, 当 a 在 0.5～0.7 范围内取相应的值时, 算法性能能达到最优, 当 a 继续增大后, 算法性能有一定程度的下降.

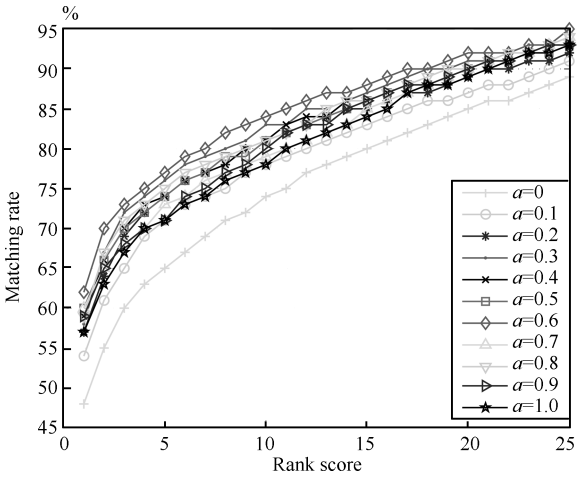

图 5 算法在 ETHZ 数据集上在不同权值 a 下的性能比较

Fig. 5 Performance comparison at different weights a on the ETHZ dataset

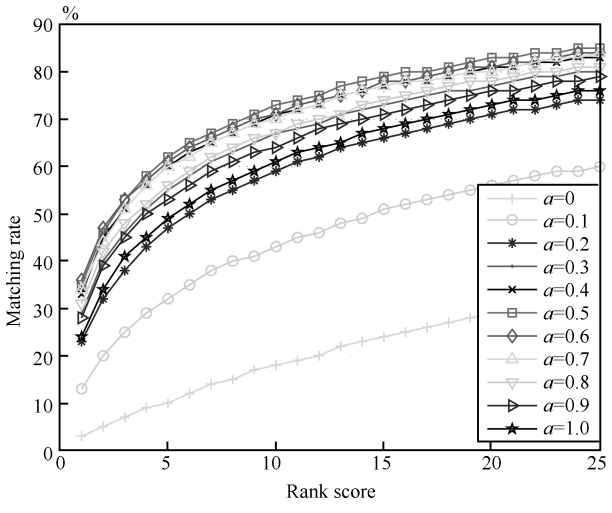

图 6 算法在 CUHK01 数据集上在不同权值 a 下的性能比较

Fig. 6 Performance comparison at different weights a on the CUHK01 dataset

3 结论

已有基于测度学习的行人再识别算法一般是把多特征融合并建立特征模型, 然后基于相应的测度

学习算法在原始特征空间学习得到测度矩阵. 但是将不同的属性特征融合在一起, 却忽略了不同属性特征之间的差别以及由于原始特征空间的线性不可分, 这样学习得到的测度矩阵不能准确地表示样本之间的相似性与差异性, 从而导致识别效果较差. 对此, 本文提出基于多特征子空间与核学习的行人再识别算法. 该算法首先将不同特征子空间投影到可分性好的核空间; 其次, 在不同核空间中学习得到测度矩阵和相应的相似性函数; 最后, 将样本在不同核空间的相似性按照一定的权值结合来表示样本之间最终的相似性.

总的来说, 4 个公共数据集里的图像有着光照变化、行人姿态变化、视角变化以及遮挡等情况, 但该算法在 4 个数据集上的识别效果都比已有算法更好. 由此可见, 算法对光照变化、行人姿态变化和视角变化以及遮挡都具有较好的鲁棒性.

基于距离测度学习的行人再识别算法其实就是要寻找一个最优的测度矩阵 M, 使得同类样本之间的距离尽可能小, 异类样本之间的距离尽可能大. 为了得到一个最近似于最优的测度矩阵 M, 下一步可以考虑如何融合不同方法得到的测度矩阵.

References

1 Swain M J, Ballard D H. Indexing via color histograms. In: Proceedings of the 3rd International Conference on Computer Vision. Osaka, Japan: IEEE, 1990. 390−393

2 Gray D, Brennan S, Tao H. Evaluating appearance models for recognition, reacquisition, and tracking. In: Proceedings of the 2007 IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. Rio de Janeiro, Brazil: IEEE, 2007. 41−47

3 Farenzena M, Bazzani L, Perina A, Murino V, Cristani M. Person re-identification by symmetry-driven accumulation of local features. In: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, USA: IEEE, 2010. 2360−2367

4 Ma B P, Su Y, Jurie F. BiCov: a novel image representation for person re-identification and face verification. In: Proceedings of the 23rd British Machine Vision Conference. Surrey, UK: IEEE, 2012. 1−6

5 Cheng D S, Cristani M, Stoppa M, Bazzani L, Murino V. Custom pictorial structures for re-identification. In: Proceedings of the 22nd British Machine Vision Conference. Dundee, UK: IEEE, 2011. 749−760

6 Liu C X, Gong S G, Loy C C, Lin X G. Person re-identification: what features are important? In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: IEEE, 2012. 391−401

7 Wang X G, Doretto G, Sebastian T, Rittscher J, Tu P. Shape and appearance context modeling. In: Proceedings of the 11th International Conference on Computer Vision. Rio de Janeiro, Brazil: IEEE, 2007. 1−8

8 Li W, Zhao R, Xiao T, Wang X G. DeepReID: deep filter pairing neural network for person re-identification. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, USA: IEEE, 2014. 152−159

9 Ahmed E, Jones M, Marks T K. An improved deep learning architecture for person re-identification. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 3908−3916

10 Zhao R, Ouyang W L, Wang X G. Unsupervised salience learning for person re-identification. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Portland, USA: IEEE, 2013. 3586−3593

11 Zheng W S, Gong S G, Xiang T. Person re-identification by probabilistic relative distance comparison. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA: IEEE, 2011. 649−656

12 Du Yu-Ning, Ai Hai-Zhou. A statistical inference approach for person re-identification. *Journal of Electronics and Information Technology*, 2014, **36**(7): 1612−1618 (杜宇宁, 艾海舟. 基于统计推断的行人再识别算法. 电子与信息学报, 2014, **36**(7): 1612−1618)

13 Kostinger M, Hirzer M, Wohlhart P, Roth P M, Bischof H. Large scale metric learning from equivalence constraints. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, Rhode island, USA: IEEE, 2012. 2288−2295

14 Mignon A, Jurie F. Pcca: a new approach for distance learning from sparse pairwise constraints. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, Rhode island, USA: IEEE, 2012. 2666−2672

15 Dikmen M, Akbas E, Huang T S, Ahuja N. Pedestrian recognition with a learned metric. In: Proceedings of the 10th Asian Conference on Computer Vision. Queenstown, New Zealand: Springer-Verlag, 2010, **6495**: 501−512

16 Hirzer M, Roth P M, Köstinger M, Bischof H. Relaxed pairwise learned metric for person re-identification. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: Springer-Verlag, 2012, **7577**: 780−793

17 Prosser B, Zheng W S, Gong S G, Xiang T. Person re-identification by support vector ranking. In: Proceedings of the 22th British Machine Vision Conference (BMVC). Aberystwyth, UK: IEEE, 2010. 1−11

18 Li W, Zhao R, Wang X G. Human re-identification with transferred metric learning. In: Proceedings of the 11th Asian Conference on Computer Vision. Daejeon, Korea: Springer-Verlag, 2012, **7724**: 31−44

19 Loy C C, Liu C X, Gong S G. Person re-identification by manifold ranking. In: Proceedings of the 20th IEEE International Conference on Image Processing (ICIP). Melbourne, VIC, Australia: IEEE, 2013. 3567−3571

20 Xiong F, Gou M R, Camps O, Sznaier M. Person re-identification using kernel-based metric learning methods. In: Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer-Verlag, 2014, **8695**: 1−16

21 Gevers T, Smeulders A W M. Color-based object recognition. *Pattern Recognition*, 1999, **32**(3): 453−464

22 Du Y N, Ai H Z, Lao S H. Evaluation of color spaces for person re-identification. In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR). Tsukuba, Japan: IEEE, 2012. 1371−1374

23 Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In: Proceedings of the 10th European Conference on Computer Vision. Marseille, France: Springer-Verlag, 2008, **5302**: 262−275

24 Zheng W S, Gong S G, Xiang T. Associating groups of people. In: Proceedings of the 20th British Machine Vision Conference (BMVC). London, UK: IEEE, 2009. 1−11

25 Zhou T F, Qi M B, Jiang J G, Wang X, Hao S J, Jin Y L. Person re-identification based on nonlinear ranking with difference vectors. *Information Sciences*, 2014, **279**: 604−614

26 Zhao R, Ouyang W L, Wang X G. Learning mid-level filters for person re-identification. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, USA: IEEE, 2014. 144−151

27 Pedagadi S, Orwell J, Velastin S, Boghossian B. Local fisher discriminant analysis for pedestrian re-identification. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Portland, USA: IEEE, 2013. 3318−3325

28 Li Z, Chang S Y, Liang F, Huang T S, Cao L L, Smith J R. Learning locally adaptive decision functions for person verification. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Portland, USA: IEEE, 2013. 3610−3617

29 Zheng W S, Gong S G, Xiang T. Re-identification by relative distance comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(3): 653−668

30 Davis J V, Kulis B, Jain P. Information-theoretic metric learning. In: Proceedings of the 24th International Conference on Machine Learning. Corvallis, USA: IEEE, 2007. 209−216

31 Liu H, Qi M B, Jiang J G. Kernelized relaxed margin components analysis for person re-identification. *IEEE Signal Processing Letters*, 2015, **22**(7): 910−914

32 Weinberger K Q, Saul L K. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 2009, **10**(2): 207−244

齐美彬 合肥工业大学计算机与信息学院教授. 主要研究方向为视频编码, 运动目标检测与跟踪和 DSP 技术.
E-mail: qimeibin@163.com
(**QI Mei-Bin** Professor at the School of Computer and Information, Hefei University of Technology. His research interest covers video coding, moving target detection and tracking, and DSP technology.)

檀胜顺 合肥工业大学计算机与信息学院硕士研究生. 主要研究方向为计算机视觉, 图像处理, 行人再识别. 本文通信作者.
E-mail: tss901118@mail.hfut.edu.cn
(**TAN Sheng-Shun** Master student at the School of Computer and Information, Hefei University of Technology. His research interest covers computer vision, image processing, and person re-identification. Corresponding author of this paper.)

王运侠 合肥工业大学计算机与信息学院硕士研究生. 主要研究方向为计算机视觉和图像检索.
E-mail: wangyunxia0807@163.com
(**WANG Yun-Xia** Master student at the School of Computer and Information, Hefei University of Technology. Her research interest covers computer vision and image retrieval.)

刘 皓 合肥工业大学计算机与信息学院博士研究生. 2014 年获得合肥工业大学硕士学位. 主要研究方向为图像检索和行人再识别.
E-mail: hfut.haoliu@gmail.com
(**LIU Hao** Ph. D. candidate at the School of Computer and Information, Hefei University of Technology. He received his master degree from Hefei University of Technology in 2014. His research interest covers image retrieval and person re-identification.)

蒋建国 合肥工业大学计算机与信息学院教授. 主要研究方向为数字图像分析与处理, 分布式智能系统和 DSP 技术及应用. E-mail: jgjiang@hfut.edu.cn
(**JIANG Jian-Guo** Professor at the School of Computer and Information, Hefei University of Technology. His research interest covers digital image analysis and processing, distributed intelligent systems, DSP technology and applications.)