

基于双向参考集矩阵度量学习的行人再识别

陈莹* 许潇月

(江南大学轻工过程先进控制教育部重点实验室 无锡 214000)

摘要: 针对行人再识别中由于外观差异不显著导致特征描述不准确的问题, 该文提出一种基于双向参考集矩阵度量学习(BRM²L)的行人再识别算法。首先通过互近邻算法获得每个摄像头下的互近邻参考集, 为保证参考集的鲁棒性, 联合考虑各摄像头下的互近邻参考集获得双向参考集。通过双向参考集挖掘出困难样本进行特征描述, 从而得到准确的外观差异描述。最后利用该特征描述进行更有效的矩阵度量学习。在多个公开数据集上的实验结果证明了该算法比现有算法具有更好的行人再识别性能。

关键词: 行人再识别; 外观差异; 矩阵度量; 互近邻; 双向参考集

中图分类号: TN911.73; TP391

文献标识码: A

文章编号: 1009-5896(2020)02-0394-09

DOI: 10.11999/JEIT190159

Matrix Metric Learning for Person Re-identification Based on Bidirectional Reference Set

CHEN Ying XU Xiaoyue

(Key Laboratory of Advanced Control Light Process, Jiangnan University, Wuxi 214000, China)

Abstract: To solve the problem of inaccurate feature representation caused by indistinctive appearance difference in person re-identification domain, a new Matrix Metric Learning algorithm based on Bidirectional Reference (BRM²L) set is proposed. Firstly, reciprocal-neighbor reference sets in different camera views are respectively constructed by the reciprocal-neighbor scheme. To ensure the robustness of reference sets, the reference sets in different camera views are jointly considered to generate the Bidirectional Reference Set (BRS). With hard samples which are mined by the BRS to represent feature descriptors, accurate appearance difference representations could be obtained. Finally, these representations are utilized to conduct more effective matrix metric learning. Experimental results on several public datasets demonstrate the superiority of the proposed method.

Key words: Person re-identification; Appearance difference; Matrix metric; Reciprocal neighbor; Bidirectional reference set

1 引言

行人再识别用于识别出无重叠视域中不同摄像头视角下处于不同位置的同一个行人^[1]。作为多摄像头监控系统中的重要组成部分, 行人再识别在近几年来引起了学术界的广泛关注。然而, 由于不同摄像头视角间剧烈的光照变化, 行人的姿态差异以及遮挡等因素的存在, 限制了行人特征的表征能力。因此, 在已知行人特征的基础上, 如何通过距离度量学习得到准确的相似度度量便是需要解决的

关键问题^[2]。

众多度量学习方法多数根据行人自身的外观, 用一个人工设计或深度学习的特征向量表示每个行人。随后, 用特征向量间的距离表示图像对之间的相似性^[3]。然而, 正如文献^[4]所说, 仅通过每个行人自身的特征来描述该行人是不够全面的, 也应当考虑该行人与其他行人间的外观关系。为此, 基于参考集的方法得到了关注。Chen等人^[5]建立不同摄像头视角下轨迹的参考集, 根据参考集完成轨迹的外观建模。An等人^[6]也利用参考集表达行人特征, 该方法侧重于重构参数以产生更准确的特征向量描述子。为更好地消除因每个摄像头里成像条件不同带来的特征上的偏差, 文献^[4]在向量描述子的基础上提出了基于参考集的矩阵描述子。其中, 参考集是采用随机选取样本的策略生成的, 该策略容易选

收稿日期: 2019-03-18; 改回日期: 2019-05-24; 网络出版: 2019-07-03

*通信作者: 陈莹 chenying@jiangnan.edu.cn

基金项目: 国家自然科学基金(61573168)

Foundation Item: The National Natural Science Foundation of China (61573168)

出许多简单样本进行训练, 导致了学习所得的度量矩阵的判别力较弱。

为避免随机策略下过多的简单样本构成参考集, 保证度量学习性能的稳定性和有效性, 本文提出了基于双向参考集矩阵度量学习的行人再识别算法(Bidirectional Reference set based Matrix Metric Learning for person re-identification, BRM²L)。该算法关注于寻找困难样本构建参考集, 从而保证矩阵描述子的表征力, 首先构造出每个行人在各摄像头视角下的互近邻参考集。其次, 保证不同摄像头视角下参考集样本标签的一致性, 利用联合互近邻参考集获得双向参考集。然后, 根据所得的双向参考集计算每个行人的差异矩阵, 随后通过轮换寻优和梯度下降算法进行矩阵度量学习, 以实现不同视角下的行人再识别。

2 问题分析

文献[4]中提出了矩阵描述子的概念并将其运用于基于视频的行人再识别。首先从训练集中, 随机选取多个不同视角下的相同行人构成参考集。对每个行人, 用其本身的特征与当前视角下的参考集中每个行人特征作差, 从而获得该行人的矩阵描述子, 如图1所示。

尽管矩阵描述子考虑了行人与其他行人间的外观关系, 使特征描述更为准确, 但仍存在以下问题:

(1) 不稳定性: 各摄像头下的参考集是由随机选取策略生成的。对于摄像头下某个行人来说, 其参考集由同一视角下除本身之外任意数个行人构成。随机选择策略使参考集中的样本构成具有很大的不确定性。

(2) 弱判别性: 同一视角下, 负样本的数量远大于正样本的数量, 多数负样本与正样本外观差异显著^[7]。因此, 随机选择策略容易选出许多简单样本组成参考集。由于简单样本与当前行人差异

很大, 两者作差所得的差异向量不能反映当前行人有识别力的特征, 使得矩阵描述子的表征能力不够显著。

为解决以上问题, 本文采用了基于互近邻模型的双向参考集。互近邻模型致力于挖掘同一视角下与每个行人的困难样本。考虑到相机差异导致不同视角下同一行人的互近邻集差异较大, 本文提出基于互近邻的双向参考集, 以保证选出当前行人在不同视角下标签一致的困难样本, 即不同视角下均位于邻近集中的行人作为最终的困难样本, 形成参考集。通过量化当前行人与其困难样本的差异, 可以消除相机视角下偏差并凸显出当前行人显著区别于其他行人的特征。

3 算法框架

本文采用的基于双向参考集矩阵度量学习的行人再识别算法框架如图2所示。

为提高矩阵描述子的表征能力, 首先利用互近邻模型分别构建摄像头A和摄像头B下的双向参考集, 如图2(a)所示, 灰色虚线框内包含当前行人的互近邻集, 带有蓝色框的行人表示同一行人在两个摄像头所得的互近邻集里标签相同的行人, 由这些行人分别构成不同视角下的双向参考集。其次, 根据双向参考集生成行人的矩阵描述子。如图2(b)所示, 在当前摄像头下, 行人 i 关于双向参考集中行人1的差异描述用标号为1的三角形表示。因此, 虚线圆中的所有三角形即为行人 i 在该摄像头下的差异矩阵描述子。

在随后的矩阵度量学习部分, 本文采用的是文献[4]中的学习方法, 该方法最终获得内部差异投影矩阵和外部差异投影矩阵。两个投影矩阵的推拉力度分别用红色箭头和蓝色箭头表示。在图2(c)中, 蓝色和绿色三角分别表示来自摄像头A和B的同一行人, 橙色菱形表示摄像头B下的不同行人。

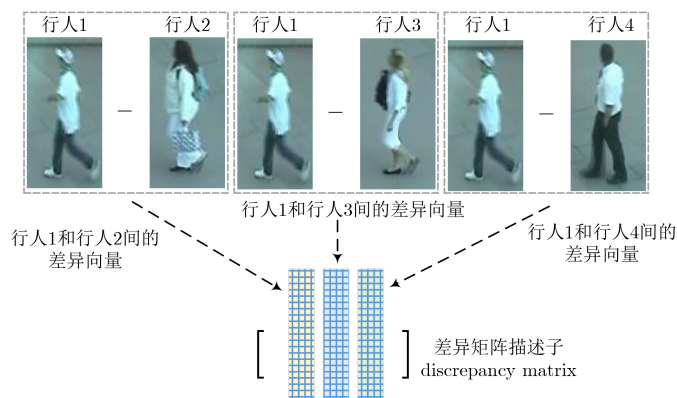


图1 差异矩阵描述子

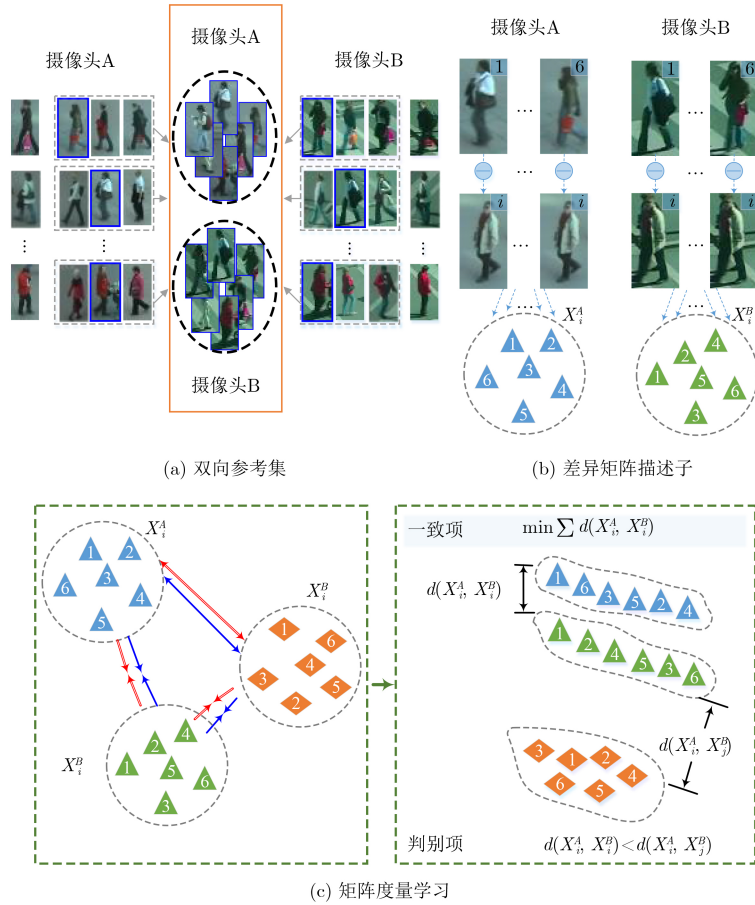


图2 算法框架图

4 基于互近邻的双向参考集

双向参考集是根据不同摄像头下的互近邻集联合产生的。首先，互近邻算法获取行人在该摄像头下的困难样本来构建当前视角下每个行人的互近邻集。随后联合考虑各摄像头下的互近邻集，获得双向参考集。

4.1 互近邻模型

互近邻模型通常是以摄像头A中的样本作为待查询集(probe set)，摄像头B中的样本作为候选行人库(gallery set)^[8]。但本文的目的在于选出同一摄像头下的困难样本，以生成每个样本在摄像头下的参考集。因此，不同于文献^[8]，本文在各摄像头内构建待查询集和候选行人库。以摄像头A为例，给出该视角下待查询集里标签为 p 的行人，其特征表示为 $\mathbf{x}_p^A$ ，并记该视角下的候选行人库为 $\mathcal{Gal} = [g_1^A, g_2^A, \dots, g_{N_g}^A]$ 。其中， N_g 为候选行人库里的样本数。首先，通过欧式距离度量得到行人 p 的原始排序列表(ranking list)，记位于原始排序列表的前 k 个样本为 $\text{top}_k(p) = [g_{p,1}, g_{p,2}, \dots, g_{p,k}]$ 。当前的近邻关系只考虑了从 p 到 $\mathcal{Gal}$ 的相似度关系，即

$$g_{p,i} \in \text{top}_k(p) \nRightarrow p \in \text{top}_k(g_{p,i}), i \in [1, k] \quad (1)$$

为使邻近关系更为可靠，该模型随后对 $g_{p,i}$ 在 $\mathcal{Gal}$ 中进行反向查询，获得 $\text{top}_k(g_{p,i})$ 。此时，可根据式(2)获得 p 在摄像头A下的互近邻样本

$$\mathbf{R}_{k,A}(p) = \text{top}_k(p) \cap \text{top}_k(g_{p,i}) \quad (2)$$

然而，由于摄像头内部和摄像头之间的变化，一些与 p 外观和行为相似的困难样本可能未被包含在 $\text{top}_k(p)$ 中，因而没有在 $\mathbf{R}_{k,A}(p)$ 中出现。为此，该模型采用了扩充策略来保留这些困难样本。具体来说，对于每个 $g_{p,i} \in \mathbf{R}_{k,A}(p)$ ，如果满足式(3)，则用 $\mathbf{R}_{k,A}(g_{p,i})$ 的一半扩充 $\mathbf{R}_{k,A}(p)$ ，得到 p 最终的互近邻集 $\tilde{\mathbf{R}}_{k,A}(p)$ ，如果 $|\mathbf{R}_{k,A}(p) \cap \mathbf{R}_{k/2,A}(g_{p,i})| \geq \frac{2}{3} |\mathbf{R}_{k/2,A}(g_{p,i})|$ ，则

$$\tilde{\mathbf{R}}_{k,A}(p) \leftarrow \mathbf{R}_{k,A}(p) \cup \mathbf{R}_{k/2,A}(g_{p,i}) \quad (3)$$

因此，由互近邻模型可以分别得到不同摄像头下每个行人的互近邻集合。然而，不同摄像头下，每个行人互近邻集合较大，若合并所有行人的互近邻集并依靠随机选择策略^[4]选取样本，选择范围太大，从而导致难以定位出同属于不同摄像头下的困难样本。

4.2 双向参考集

获得了不同摄像头下的互近邻集后，需要基于

互邻近集生成多摄像头下的参考集。由于摄像头间的差异,不同视角下,同一行人所得的邻近集里所包含行人样本标签的差异比较显著。为了清晰反应不同视角下同一行人邻近集的关系,本文作了先验实验。随机取4个行人,将这3个行人在摄像头A和摄像头B下生成邻近集的重叠率 σ 记录在表1。可以看出,邻近集的重叠率较低。

表1 两个摄像头下参考集里的样本标签的重叠率 σ (%)

行人	行人1	行人2	行人3	行人4
重叠率 σ	20	50	10	25

因此,为保证能够选出每个行人的困难样本构建参考集,本文在所得邻近集的基础上,提出了双向参考集模型。已知在多个摄像头视角下,捕捉到了 M 个行人。以两个摄像头的情况为例,对于行人 p ,其在两个摄像头下的互近邻集分别为: $\tilde{\mathbf{R}}_{k,A}(p)$ 和 $\tilde{\mathbf{R}}_{k,B}(p)$ 。那么,行人 p 的双向参考集为

$$\mathcal{R}(p) = \tilde{\mathbf{R}}_{k,A}(p) \cap \tilde{\mathbf{R}}_{k,B}(p) \quad (4)$$

因此, M 个行人构成的参考样本候选池(reference candidate pool)里样本的标签集合为

$$\mathcal{R} = \text{unique}[\mathcal{R}(1)\mathcal{R}(2) \cdots \mathcal{R}(M)] \quad (5)$$

最终,从 $\mathcal{R}$ 中随机选取 N_r 个样本构成每个摄像头下的双向参考集,即为

$$\mathbf{F}^c = [\mathbf{x}_{f_1}^c \mathbf{x}_{f_2}^c \cdots \mathbf{x}_{f_{N_r}}^c] \quad (6)$$

其中, $c \in [A, B, C, \dots]$ 。

双向参考集模型兼顾了每个行人在不同摄像头下生成的邻近集,将它们共有的邻近样本放入双向参考集中,从而保证了所选样本是与该行人外观相似的困难样本。

为进一步反映双向参考模型有利于行人再识别性能的提升,实验过程的6.3.2节进行了分析讨论。

5 基于双向参考集的矩阵度量学习

在行人再识别领域,表征一个行人外观的最常用方式是利用该行人本身的特点来描述。常用的特征描述子便是将每个行人的图片转换成一个有识别力的特征向量。然而,每个行人不仅可以通过自身的特征来描述,也可以通过该行人与其他行人间的外观差异来描述。同时,由于不同摄像头下的比例缩放、光照变化、姿态差异以及成像条件的不同,使得原始的特征描述容易受到摄像头偏差的影响,进而影响其鲁棒性。

5.1 差异矩阵描述子

给出一组参考图像的特征 $\mathbf{F} = [\mathbf{x}_{f_1} \mathbf{x}_{f_2} \cdots \mathbf{x}_{f_{N_r}}]$,

$\mathbf{x}_f \in \mathbb{R}^{N_f \times 1}$, 其中, N_f 为特征向量维度。对于任意一张图像,其特征用 $\mathbf{x}$ 表示,那么,该图像特征的特征差异矩阵描述子将由其与 $\mathbf{F}$ 中的每个向量的特征差异表示为

$$\mathbf{X} = [\mathbf{x} - \mathbf{x}_{f_1} \quad \mathbf{x} - \mathbf{x}_{f_2} \cdots \mathbf{x} - \mathbf{x}_{f_{N_r}}] \quad (7)$$

那么,由式(6)可知,摄像头A下的双向参考集为 $\mathbf{F}^A = [\mathbf{x}_{f_1}^A \mathbf{x}_{f_2}^A \cdots \mathbf{x}_{f_{N_r}}^A]$,摄像头B下的双向参考集为 $\mathbf{F}^B = [\mathbf{x}_{f_1}^B \mathbf{x}_{f_2}^B \cdots \mathbf{x}_{f_{N_r}}^B]$ 。因此,根据式(7),对于来自摄像头A的行人 p 的图像 I_p^A ,其差异矩阵描述子为 $\mathbf{X}_p^A = [\mathbf{x}_p^A - \mathbf{x}_{f_1}^A \quad \mathbf{x}_p^A - \mathbf{x}_{f_2}^A \cdots \mathbf{x}_p^A - \mathbf{x}_{f_{N_r}}^A]$ 。同样,对于来自摄像头B的行人 p 的图像 I_p^B ,其差异矩阵描述子为 $\mathbf{X}_p^B = [\mathbf{x}_p^B - \mathbf{x}_{f_1}^B \quad \mathbf{x}_p^B - \mathbf{x}_{f_2}^B \cdots \mathbf{x}_p^B - \mathbf{x}_{f_{N_r}}^B]$,由此,特征描述完成了从行人本身的向量描述到行人差异的矩阵描述的转换。

5.2 矩阵距离度量

给出两个摄像头下的特征描述集合: $[\mathbf{X}_1^A \cdots \mathbf{X}_i^A \cdots \mathbf{X}_M^A]$ 和 $[\mathbf{X}_1^B \cdots \mathbf{X}_j^B \cdots \mathbf{X}_M^B]$,度量学习的本质在于在监督信息下,通过相似度约束和不相似度约束,求得最优的投影矩阵。为更好地保留差异矩阵的结构且引入更少的参数,参照文献[4],本文也采用基于双投影矩阵的矩阵度量学习算法进行行人再识别。该算法首先定义了图像对的距离度量方式。具体地,对于任意图像对 $(\mathbf{X}_p^A, \mathbf{X}_q^B)$,采用矩阵度量计算图像对间的距离

$$d(\mathbf{X}_p^A, \mathbf{X}_q^B) = \|\mathbf{L}_1(\mathbf{X}_p^A - \mathbf{X}_q^B)\mathbf{L}_2\|_F^2 \quad (8)$$

其中, $\mathbf{L}_1 \in \mathbb{R}^{N_1 \times N_f}$ 为内部差异投影矩阵(intradiscrepancy projection matrix),其侧重于对每个差异矩阵的不同维度进行加权,体现同一差异矩阵中不同维度的重要性。 $\mathbf{L}_2 \in \mathbb{R}^{N_r \times N_2}$ 为外部差异投影矩阵(interdiscrepancy projection matrix),其侧重于反映不同差异矩阵的重要性,选出双向参考集中更有效的参考图像进行距离度量。

最终的目标函数由类内一致项 F_1 ,类间判别项 F_2 和稀疏项 F_3 3部分组成

$$F(\mathbf{L}_1, \mathbf{L}_2) = F_1(\mathbf{L}_1, \mathbf{L}_2) + F_2(\mathbf{L}_1, \mathbf{L}_2) + \lambda F_3(\mathbf{L}_2) \quad (9)$$

其中, $F_1(\mathbf{L}_1, \mathbf{L}_2) = \frac{1}{M} \sum_{i=1}^M d(\mathbf{X}_i^A, \mathbf{X}_i^B)$, $F_1(\mathbf{L}_1, \mathbf{L}_2) =$

$$\frac{1}{M} \sum_{i=1}^M d(\mathbf{X}_i^A, \mathbf{X}_i^B), F_1(\mathbf{L}_1, \mathbf{L}_2) = \frac{1}{M} \sum_{i=1}^M d(\mathbf{X}_i^A, \mathbf{X}_i^B),$$

$$F_2(\mathbf{L}_1, \mathbf{L}_2) = \frac{1}{T} \sum_{k=1}^T \frac{1}{\alpha} \ln(1 + e^{\alpha I(t_k)}), F_3(\mathbf{L}_2) = \|\mathbf{L}_2\|_{2,1},$$

λ 为稀疏项的权重。

获得了目标函数后, 矩阵度量学习问题便等价于为求解式(10)所示最优问题

$$(\mathbf{L}_1^*, \mathbf{L}_2^*) = \arg \min_{\mathbf{L}_1, \mathbf{L}_2} F(\mathbf{L}_1, \mathbf{L}_2) \quad (10)$$

为高效求解该模型, 本文采用交替优化^[9]的过程中采用简单的梯度下降算法迭代地学习 $\mathbf{L}_1$ 和 $\mathbf{L}_2$ 。

6 实验结果及分析

实验仿真均在Matlab R2016a, Window10系统, Inter(R) Core(TM) i5-6500 CPU处理器, 3.20 GHz主频和8 GB内存的计算机环境下编程实现。

6.1 实验数据集与评价标准

6.1.1 实验数据集

(1) VIPeR数据集^[10]: VIPeR数据集包含两个不重叠视角的摄像头视角下632个行人的1264张图片, 每个行人都有一对分别从两个视角下拍摄的图片。所有行人的图片被归一化为 128×48 像素。引起行人外观变化的最主要因素是视角的变化。此外, 该数据集也包含了光照变化, 姿态变化以及图像质量的变化。

(2) PRID 450S数据集^[11]: PRID 450S数据集包含450个行人分别在两个视角不相交的摄像头里的900张图片。每个行人在两个视角下均有一张图片。所有图片的尺寸被归一化为 168×80 像素。不同于VIPeR数据集, 该数据集有着显著并且一致的光照变化。

(3) CUHK01数据集^[12]: CUHK01数据集是一个更大的数据集, 它包含了来自两个不重叠摄像头视角下的971个行人, 一共有3884张图片。所有图片尺寸都被归一化为 160×60 像素。摄像头A中包含更多视角和姿态变化, 摄像头B中则主要包含从正面和背面拍摄的行人图像。考虑到本文中每个摄像头下的每个行人只需要一张代表性的图片, 因此, 进行实验时, 本文从该数据集的每个视角下每个行人的两个样本里随机选取一张图片。

6.1.2 评价标准

本文采用行人再识别领域最常使用的两个性能评价标准: 累计匹配特性(Cumulative Match Characteristic, CMC)^[13]曲线和Rank- N 表格。

累计匹配特性首先获得待查询行人检索结果, 随后计算出前 r 个检索结果中包含正确匹配的比率。该评估指标广泛应用于行人再识别性能评估中, 其计算公式为

$$\text{CMC}(k) = \sum_{r=1}^k p(r), k = 1, 2, \dots, m \quad (11)$$

式(11)表明CMC曲线反映了前 k 个匹配结果中

获得正确匹配的概率。该曲线的横坐标是 k , 纵坐标是识别率。纵坐标数值越大, 说明算法的识别性能越好。随着候选目标数的增多, 识别率必然上升。因此, CMC曲线呈递增趋势。

Rank- N 表格是以数值的形式给出, 因而更能直观地体现识别精度的差异。该评估标准给出了关键匹配点处的累计匹配率。通常选择Rank-1, Rank-5, Rank-10和Rank-20处的累计匹配率。

6.2 实验设置

本文利用两种特征来验证BRM²L算法的有效性: 27622维的GoG^[14]描述子和4096维的FTCNN^[15]描述子。首先利用PCA将特征降至70维, 即 $N_f = 70$ 。实验过程中, 设定 $N_1 = N_f$ 且 $N_2 = N_r$ 。同时, 为每个数据集构造3个子集: 训练集, 测试集和初始参考集。每个子集均从总的数据集里随机选出。具体地:

(1) VIPeR数据集: 随机选出232个样本对作为初始参考集以得到双向参考集。设定双向参考集大小 $N_r = 100$ 。训练集和测试集大小相同, 均由200个样本对构成, 即 $M = 200$ 。

(2) PRID 450S数据集: 随机选出100个样本对作为初始参考集, 设定 $N_r = 60$, $M = 150$ 。

(3) CUHK01数据集: 随机选取1342个样本作为初始参考集, 设定 $N_r = 100$, $M = 300$ 。

在互近邻模型中, 依照文献^[11], 取 $k = 10$ 。

6.3 实验结果及讨论

6.3.1 自身评估

引入参考集后, 图像的特征不再由单一的特征向量描述, 而是由多个特征差异构成的矩阵来描述。为评价所采用的基于差异矩阵的矩阵度量学习的性能, 如表2所示, 该部分在3个数据集上对比了 L_2 范数度量下的特征向量, F 范数下的差异矩阵以及矩阵度量下的差异矩阵的Rank-1(%)和Rank-5(%)的匹配精度。

由表2可知, 本文的BRM²L度量在3个数据集上的效果更为出色。

6.3.2 双向参考集性能评估

为证明双向参考集的有效性, 该部分采用基于互近邻的双向参考集模型与传统的随机选择策略^[4,5]所得到的匹配精度。实验结果如图3所示。纵轴表示匹配精度, 横轴表示Rank值。图3(a), 3(b), 3(c)为采用GoG描述子的实验结果, 图3(d), 3(e), 3(f)为采用FTCNN描述子的实验结果。由图3可以看出, 基于互近邻的双向参考集在3个数据集上的匹配精度均更出色。这是由于其兼顾不同摄像头下生成的邻近集, 选出了与训练集中行人外观相似的

表2 在3个数据集上采用不同特征的匹配精度(%)

方法	VIPeR		CHUK01		PRID450S	
	Rank-1	Rank-5	Rank-1	Rank-5	Rank-1	Rank-5
L_2 范数(GoG)	19.00	38.00	24.17	51.33	32.44	60.00
F范数(GoG)	20.17	41.83	34.50	69.83	52.22	80.22
BRM ² L(GoG)	38.33	69.17	45.33	70.50	54.44	80.67
L_2 范数(FCTNN)	29.00	46.00	37.44	58.00	31.73	57.96
F范数(FCTNN)	30.00	49.83	46.56	72.11	44.40	72.84
BRM ² L(FCTNN)	41.33	68.17	47.42	77.44	45.51	72.96

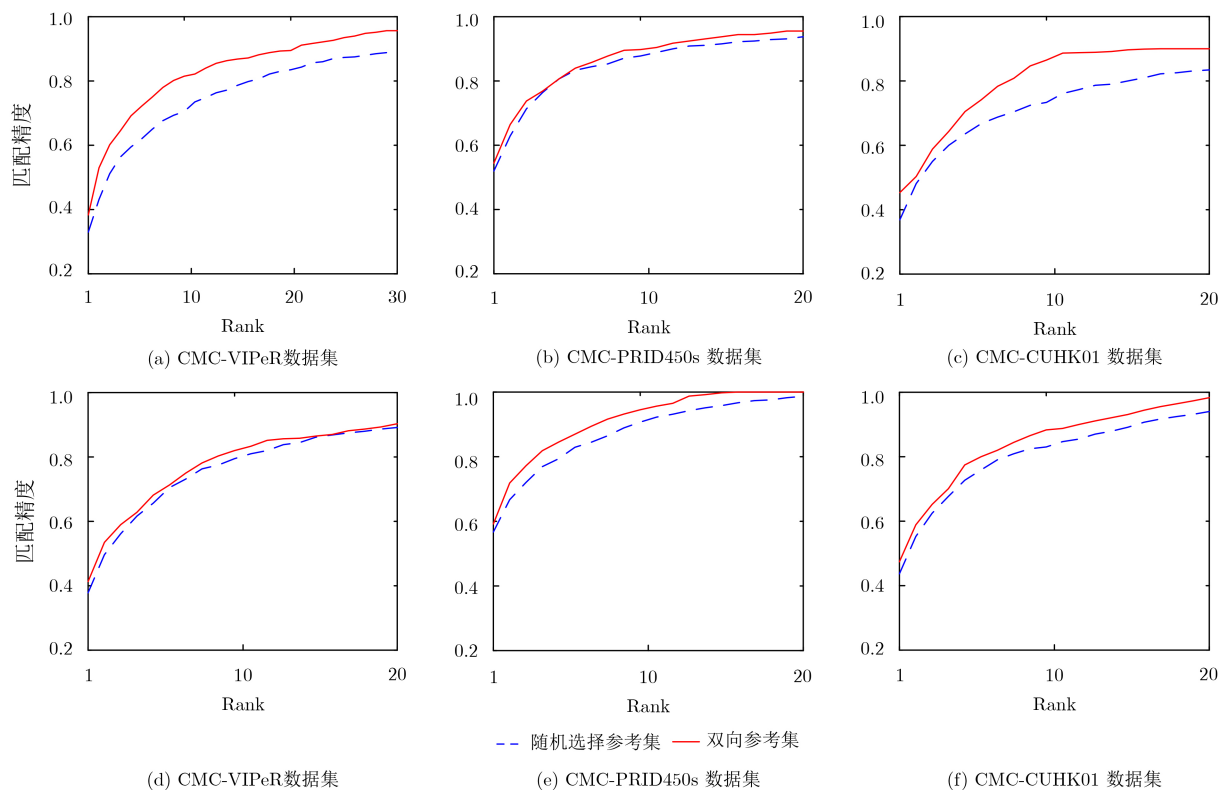


图3 双向参考集与随机参考集结果对比图

困难样本。因而在矩阵度量所得的投影矩阵更有判别力，提升了行人再识别的匹配精度。

6.3.3 与现有算法匹配精度对比

本节将本文采用的BRM²L算法在不同数据集上同当前的行人再识别方法进行对比。

(1) VIPeR数据集：本文的算法将与如下这些方法进行比较：PCCA^[16]，BiCov^[17]，KISSME^[18]，eSDC^[19]，LADF^[20]，Midfilter^[21]，SCNCD^[22]，LOMO+XQDA^[23]，RD^[6]，DeepMetric^[24]，DeepRanking^[25]，GoG^[14]+XQDA，FTCNN^[15]+XQDA，DeepList^[26]以及DM³^[4]。其中，RD算法提取了基于特征集的特征描述，但该特征集中没有考虑特征向量的差异性。而DM³算法中采用的参考集虽然考虑了特征向量间的差异，但没有进一步进行困难样本

挖掘。所有实验结果在表3中列出，加粗表示最优结果，下划线表示次优结果。

(2) PRID 450S数据集：本文算法将与如下算法进行比较：KISSME^[18]，SCNCD^[22]，CBRA^[27]，CSL^[28]，Mirror^[29]，DRML^[30]以及DM³^[4]。所有的实验结果在表4中列出，由表4可知，本文算法所得实验结果优于其他算法的实验结果。

(3) CUHK01数据集：除了在上述两个数据集上进行评估，本文还在CUHK01数据集上评估采用的BRM²L算法。如表5所示，本文的BRM²L算法与如下的算法进行对比：SDALF^[1]，TML^[12]，MidFilter^[21]，ImprovedDeep^[31]，RD^[6]以及DM³^[4]。从表5中可以看出，本文算法所得结果优于其他算法结果。

表 3 VIPeR数据集上的结果

方法	Rank-1	Rank-5	Rank-10	Rank-20
PCCA ^[16]	19.3	48.9	64.9	80.3
KISSME ^[18]	19.6	48.0	62.2	77.0
BiCov ^[17]	20.6	43.2	56.1	68.0
eSDC ^[19]	26.3	46.4	58.6	72.8
DeepMetric ^[24]	28.2	59.3	73.4	86.4
Midfilter ^[21]	29.1	52.5	65.9	79.9
LADF ^[20]	30.0	64.0	80.0	92.0
FTCNN ^[15] +XQDA	31.2	59.8	74.0	83.5
RD ^[6]	33.3	65.1	78.3	88.5
GoG ^[14] +XQDA	37.3	67.4	77.2	89.6
SCNCD ^[22]	37.8	68.5	81.2	90.4
DM ³ ^[4]	38.3	67.2	77.0	89.3
DeepRanking ^[25]	38.4	69.2	81.3	90.4
LOMO+XQDA ^[23]	40.0	68.5	80.5	91.0
DeepList ^[26]	<u>40.5</u>	69.1	80.1	<u>91.2</u>
BRM ² L (GoG)	38.33	<u>69.17</u>	<u>81.50</u>	89.50
BRM ² L (FTCNN)	41.33	68.17	82.00	90.33

表 4 PRID 450S数据集上的结果

方法	Rank-1	Rank-5	Rank-10	Rank-20
KISSME ^[18]	33.0	59.8	71.0	79.0
CBRA ^[27]	26.4	57.1	71.0	83.2
CSL ^[28]	44.4	71.6	82.2	89.8
Mirror ^[29]	55.4	79.3	87.8	93.9
DRML ^[30]	56.4	—	82.2	90.2
DM ³ ^[4]	<u>56.7</u>	<u>83.1</u>	88.4	94.7
BRM ² L (GoG)	54.44	80.67	<u>89.78</u>	<u>95.56</u>
BRM ² L (FTCNN)	59.20	84.53	94.53	99.78

表 5 CUHK01数据集上的结果

方法	Rank-1	Rank-5	Rank-10	Rank-20
SDALF ^[1]	9.9	22.6	30.3	41.0
TML ^[12]	20.0	43.5	56.0	69.3
MidFilter ^[21]	34.3	55.1	65.0	74.9
ImprovedDeep ^[31]	47.5	<u>71.0</u>	80.0	—
RD ^[6]	31.1	—	68.5	79.1
DM ³ ^[4]	43.7	70.1	77.4	88.7
BRM ² L (GoG)	<u>45.33</u>	70.50	<u>86.50</u>	<u>90.00</u>
BRM ² L (FTCNN)	47.42	77.44	88.33	98.33

7 结论

本文提出了基于双向参考集矩阵度量学习的行人再识别算法,该算法通过联合考虑各摄像头下的互近邻参考集获得双向参考集。双向参考集中包含

了多个摄像头下的困难样本,表达出了相似行人间的特征差异。基于该特征差异的度量学习过程提高了对相似行人的识别力,因而可以获得更有效的度量矩阵,从而提高了行人再识别性能。在VIPeR数据集,PRID 450S数据集和CUHK01数据集上验证了该算法的有效性。

参 考 文 献

- [1] FARENZENA M, BAZZANI L, PERINA A, *et al.* Person re-identification by symmetry-driven accumulation of local features[C]. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010: 2360–2367. doi: [10.1109/CVPR.2010.5539926](https://doi.org/10.1109/CVPR.2010.5539926).
- [2] 李幼蛟, 卓力, 张菁, 等. 行人再识别技术综述[J]. 自动化学报, 2018, 44(9): 1554–1568. doi: [10.16383/j.aas.2018.c170505](https://doi.org/10.16383/j.aas.2018.c170505).
- [3] LI Youjiao, ZHUO Li, ZHANG Jing, *et al.* A survey of person re-identification[J]. *Acta Automatica Sinica*, 2018, 44(9): 1554–1568. doi: [10.16383/j.aas.2018.c170505](https://doi.org/10.16383/j.aas.2018.c170505).
- [4] ZHENG Lilei, DUFFNER S, IDRISSE K, *et al.* Pairwise identity verification via linear concentrative metric learning[J]. *IEEE Transactions on Cybernetics*, 2018, 48(1): 324–335. doi: [10.1109/TCYB.2016.2634011](https://doi.org/10.1109/TCYB.2016.2634011).
- [5] WANG Zheng, HU Ruimin, CHEN Chen, *et al.* Person reidentification via discrepancy matrix and matrix metric[J]. *IEEE Transactions on Cybernetics*, 2018, 48(10): 3006–3020. doi: [10.1109/TCYB.2017.2755044](https://doi.org/10.1109/TCYB.2017.2755044).
- [6] CHEN Xiaojing, AN Le, and BHANU B. Reference set based appearance model for tracking across non-overlapping cameras[C]. 2013 International Conference on Distributed Smart Cameras, Palm Springs, USA, 2013: 1–6. doi: [10.1109/ICDSC.2013.6778236](https://doi.org/10.1109/ICDSC.2013.6778236).
- [7] AN Le, KAFI M, YANG Songfan, *et al.* Person reidentification with reference descriptor[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016, 26(4): 776–787. doi: [10.1109/TCSVT.2015.2416561](https://doi.org/10.1109/TCSVT.2015.2416561).
- [8] LIAO Shengcai and LI S Z. Efficient PSD constrained asymmetric metric learning for person re-identification[C]. 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 3685–3693. doi: [10.1109/ICCV.2015.420](https://doi.org/10.1109/ICCV.2015.420).
- [9] ZHONG Zhun, ZHENG Liang, CAO Donglin, *et al.* Re-ranking person re-identification with k-reciprocal encoding[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 3652–3661. doi: [10.1109/CVPR.2017.389](https://doi.org/10.1109/CVPR.2017.389).

- [9] SUN Yipeng, TAO Xiaoming, LI Yang, *et al.* Robust two-dimensional principal component analysis via alternating optimization[C]. 2013 IEEE International Conference on Image Processing, Melbourne, Australia, 2013: 340–344. doi: [10.1109/ICIP.2013.6738070](https://doi.org/10.1109/ICIP.2013.6738070).
- [10] GRAY D, BRENNAN S, and TAO Hai. Evaluating appearance models for recognition, reacquisition, and tracking[C]. The 10th IEEE International Workshop on Performance Evaluation for Tracking and Surveillance, Rio de Janeiro, 2007: 1–7.
- [11] ROTH P M, HIRZER M, KÖSTINGER M, *et al.* Mahalanobis Distance Learning For Person Re-identification[M]. London: Springer, 2014: 247–267. doi: [10.1007/978-1-4471-6296-4_12](https://doi.org/10.1007/978-1-4471-6296-4_12).
- [12] LI Wei, ZHAO Rui, and WANG Xiaogang. Human reidentification with transferred metric learning[C]. The 11th Asian Conference on Computer Vision, Daejeon, Korea, 2012: 31–44. doi: [10.1007/978-3-642-37331-2_3](https://doi.org/10.1007/978-3-642-37331-2_3).
- [13] WANG Xiaogang, DORETTO G, SEBASTIAN T, *et al.* Shape and appearance context modeling[C]. The 11th IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, 2007: 1–8. doi: [10.1109/ICCV.2007.4409019](https://doi.org/10.1109/ICCV.2007.4409019).
- [14] MATSUKAWA T, OKABE T, SUZUKI E, *et al.* Hierarchical gaussian descriptor for person re-identification[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1363–1372. doi: [10.1109/CVPR.2016.152](https://doi.org/10.1109/CVPR.2016.152).
- [15] MATSUKAWA T and SUZUKI E. Person re-identification using CNN features learned from combination of attributes[C]. The 23rd International Conference on Pattern Recognition, Cancun, Mexico, 2016: 2428–2433. doi: [10.1109/ICPR.2016.7900000](https://doi.org/10.1109/ICPR.2016.7900000).
- [16] MIGNON A and JURIE F. PCCA: A new approach for distance learning from sparse pairwise constraints[C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 2666–2672. doi: [10.1109/CVPR.2012.6247987](https://doi.org/10.1109/CVPR.2012.6247987).
- [17] MA Bingpeng, SU Yu, and JURIE F. Bicov: A novel image representation for person re-identification and face verification[C]. British Machine Vision Conference, Surrey, UK, 2012: 57. 1–57. doi: [10.5244/C.26.57](https://doi.org/10.5244/C.26.57).
- [18] KÖSTINGER M, HIRZER M, WOHLHART P, *et al.* Large scale metric learning from equivalence constraints[C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 2288–2295. doi: [10.1109/CVPR.2012.6247939](https://doi.org/10.1109/CVPR.2012.6247939).
- [19] ZHAO Rui, OUYANG Wanli, and WANG Xiaogang. Unsupervised salience learning for person re-identification[C]. 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 3586–3593. doi: [10.1109/CVPR.2013.460](https://doi.org/10.1109/CVPR.2013.460).
- [20] LI Zhen, CHANG Shiyu, LIANG Feng, *et al.* Learning locally-adaptive decision functions for person verification[C]. 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 3610–3617. doi: [10.1109/CVPR.2013.463](https://doi.org/10.1109/CVPR.2013.463).
- [21] ZHAO Rui, OUYANG Wanli, and WANG Xiaogang. Learning mid-level filters for person re-identification[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 144–151. doi: [10.1109/CVPR.2014.26](https://doi.org/10.1109/CVPR.2014.26).
- [22] YANG Yang, YANG Jimei, YAN Junjie, *et al.* Salient color names for person re-identification[C]. The 13th European Conference on Computer Vision, Zurich, Switzerland, 2014: 536–551. doi: [10.1007/978-3-319-10590-1_35](https://doi.org/10.1007/978-3-319-10590-1_35).
- [23] LIAO Shengcai, HU Yang, ZHU Xiangyu, *et al.* Person re-identification by local maximal occurrence representation and metric learning[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 2197–2206. doi: [10.1109/CVPR.2015.7298832](https://doi.org/10.1109/CVPR.2015.7298832).
- [24] YI Dong, LEI Zhen, LIAO Shengcai, *et al.* Deep metric learning for person re-identification[C]. 2014 International Conference on Pattern Recognition, Stockholm, Sweden, 2014: 34–39. doi: [10.1109/ICPR.2014.16](https://doi.org/10.1109/ICPR.2014.16).
- [25] CHEN Shizhe, GUO Chaochun, and LAI Jianhuang. Deep ranking for person re-identification via joint representation learning[J]. *IEEE Transactions on Image Processing*, 2016, 25(5): 2353–2367. doi: [10.1109/TIP.2016.2545929](https://doi.org/10.1109/TIP.2016.2545929).
- [26] WANG Jin, WANG Zheng, GAO Changxin, *et al.* DeepList: Learning deep features with adaptive listwise constraint for person reidentification[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(3): 513–524. doi: [10.1109/tcsvt.2016.2586851](https://doi.org/10.1109/tcsvt.2016.2586851).
- [27] DE CARVALHO PRATES R F and SCHWARTZ W R. CBRA: Color-based ranking aggregation for person re-identification[C]. 2015 IEEE International Conference on Image Processing, Quebec City, Canada, 2015: 1975–1979. doi: [10.1109/ICIP.2015.7351146](https://doi.org/10.1109/ICIP.2015.7351146).
- [28] SHEN Yang, LIN Weiyao, YAN Junchi, *et al.* Person re-identification with correspondence structure learning[C]. 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 3200–3208. doi: [10.1109/ICCV.2015.7351146](https://doi.org/10.1109/ICCV.2015.7351146).

- [2015.366](#).
- [29] CHEN Yingcong, ZHENG Weishi, and LAI Jianhuang. Mirror representation for modeling view-specific transform in person re-identification[C]. The 24th International Conference on Artificial Intelligence, Buenos Aires, Argentina, 2015: 3402–3408.
- [30] YAO Wenbin, WENG Zhenyu, and ZHU Yuesheng. Diversity regularized metric learning for person re-identification[C]. 2016 IEEE International Conference on Image Processing, Phoenix, USA, 2016: 4264–4268. doi: [10.1109/ICIP.2016.7533164](#).
- [31] AHMED E, JONES M, and MARKS T K. An improved deep learning architecture for person re-identification[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 3908–3916. doi: [10.1109/CVPR.2015.7299016](#).
- 陈 莹: 女, 1976年生, 教授, 博士生导师, 研究方向为模式识别、信息融合.
- 许潇月: 女, 1994年生, 硕士生, 研究方向为行人再识别.