

基于双向关系相似度函数学习的行人再识别^①

张娜, 张福星, 王强, 胡玲玲, 桂江生

(浙江理工大学 信息学院, 杭州 310018)

通讯作者: 桂江生, E-mail: jsgui@zstu.edu.cn

摘要: 当前的行人再识别在度量学习上采用马氏距离相似度函数, 该相似度函数只与特征差分空间有关, 忽略了一对行人图像中每个个体的外观特征, 针对上述问题, 提出了通过学习一个双向关系相似度函数 (Bidirectional Relationship Similarity Function, BRSF), 来计算一对行人图像的相似度. BRSF 不但描述了一对行人图像的互相关关系, 而且关联了一对行人图像的自相关关系. 该文利用 KISSME (Keep It Simple and Straightforward Metric) 算法的思想进行相似度函数学习, 把一对样本特征的自相关关系和互相关关系用高斯分布来表示, 通过把最终高斯分布的比值转换为 BRSF 的形式, 得到一个对背景、视角、姿势的变化具有鲁棒性的相似度函数. 在 VIPeR, QMUL GRID 两个行人再识别数据集上的实验结果表明, 本文算法具有较高的识别率, 其中在 VIPeR 数据集上, Rank1 达到了 53.21%.

关键词: 行人再识别; 距离度量学习; 双向关系相似度函数; 滑动分块

引用格式: 张娜, 张福星, 王强, 胡玲玲, 桂江生. 基于双向关系相似度函数学习的行人再识别. 计算机系统应用, 2018, 27(5): 33-40. <http://www.c-s-a.org.cn/1003-3254/6354.html>

Learning Bidirectional Relationship Similarity Function for Person Re-Identification

ZHANG Na, ZHANG Fu-Xing, WANG Qiang, HU Ling-Ling, GUI Jiang-Sheng

(School of Information and Electronics, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: These dominant algorithms to learn a similarity is the metric learning that learns a Mahalanobis Similarity Function (MSF) to estimate the similarity of a pair of persons. However, the MSF only projects a pair of persons into feature difference space and ignores the appearance of each individual. In this study, we proposed to learn a Bidirectional Relationship Similarity Function (BRSF) that greatly strengthens the modeling ability of the similarity function. BRSF not only represents the cross correlation relationship of a pair of persons, but also describes the auto correlation relationship. We use the ideal of the Keep It Simple Straightforward Metric (KISSME) algorithm to learn a similarity function. Specifically, the auto correlation relationship and cross correlation relationship of a pair of sample features are expressed by Gaussian distribution. Finally, by converting the ratio of the final Gaussian distribution into the form of BRSF, we get a similarity function which is robust to the change of background, viewpoint, and posture. The proposed method is demonstrated on two public benchmark datasets including VIPeR and QMUL GRID, and experimental results show that the proposed method achieves excellent re-identification rates compared with other similar algorithms. Moreover, the re-identification results on the VIPeR dataset with half of dataset sampled as training samples are quantitatively analyzed, and the performance of the proposed method achieves a 53.21% at Rank1 (represents the correct matched pair).

Key words: person re-identification; metric learning; Bidirectional Relationship Similarity Function (BRSF); slide block

^① 基金项目: 国家自然科学基金(61379036, 61502430); 国家自然科学基金委中丹合作项目(61361136002); 浙江省重大科技专项重点工业项目(2014C01047); 浙江理工大学 521 人才培养计划(20150428)

收稿时间: 2017-09-06; 修改时间: 2017-09-27; 采用时间: 2017-10-12; csa 在线出版时间: 2018-03-12

行人再识别是指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.行人再识别在目前的视频监控应用中是一个非常重要的任务,该技术在行人检索^[1]、多摄像机行人跟踪^[2]等视频监控领域都有着重要的应用.但由于在不同摄像机视域中,行人受视角、姿势变化、光照条件、背景变化、摄像头设置等因素的影响,导致同一行人的图像外貌存在着很大的差异,使得行人再识别问题遇到了很大的挑战.为了应对这些挑战,近年来的广大研究者提出了许多的解决方法.在行人图片特征提取方法中,其最主要的是设计对不同行人具有区分性和对视角、光照、背景的变化具有鲁棒性的描述特征.许多已经存在的行人再识别算法试图通过建立一个特有的,健壮的代表特征来描述在各种变化环境下的行人外观.文献[3]利用行人区域的对称性和不对称性将人的前景划分成不同的区域,每个区域通过提取局部累积特征来处理姿态、视角和光照的变化.文献[4]也提出了类似的方法,用一个自适应的身体外形结构表示行人图像,包括头、胸、大腿和小腿,然后提取每个部分的颜色特征.文献[5]提出了对行人图像的同一直线位置的滑动窗口进行扫描,取所有滑动窗口的直方图相同 bin 的最大值作为该水平组的直方图特征.最近,显著性信息也出现在行人再识别上的应用研究^[6-9],文献[9]采用4种方向显著性加权融合学习的方法来度量一对行人图像的相似度.

除了有代表性的特征外,距离度量学习也是行人再识别的另一个研究方面.文献[10]首次提出距离测度学习的问题,根据样本的类别,将具有相同类别标签的样本组成正样本对,不同类别标签的样本构成负样本对,然后利用这些样本对作为约束条件来训练得到一个马氏矩阵.在文献[11]提出一种 PRDC 算法,算法的原理是最优化相关距离的比较.文献[12]提出了一种叫 KISSME (Keep It Simple and Straightforward Metric) 的度量学习算法,该算法将样本之间的差向量看作是高斯分布中的一个点,同类样本的差向量分布在同一个高斯分布中,而不同样本之间的差向量分布在另一个高斯分布中,然后用概率的比值来度量样本之间的距离.文献[13]在文献[12]的基础上提出了核学习的思想,把低维线性不可分的特征空间投影到高维可分的非线性空间中再进行 KISSME 度量学习.这些算法通过学习一个马氏距离相似度函数 (Mahalanobis

Similarity Function, MSF) 来度量一对行人图像的相似度.然而,MSF 只与特征差分空间有关,忽略了每个个体的独立特征,因此对于同一个人不同摄像机中的很大外观变化的捕捉能力有限.

基于上述情况,提出了一种基于双向关系相似度函数学习方法 (Bidirectional Relationship Similarity Function, BRSF) 来计算一对行人图像的相似度.利用 KISSME 算法的思想进行相似度函数学习,把一对样本特征的自相关关系和互相关关系用高斯分布来表示,最终高斯分布的比值转换为 BRSF 的形式,进而得到一个对背景、视角、姿势变化有着更好的鲁棒性的相似度函数.

1 双向关系相似度函数的学习

1.1 BRSF 算法

近年来,行人再识别问题的研究者也尝试设计合适的相似度学习算法,更好地计算一对行人图像的相似度.当给定一对来源于非重叠摄像机视角的行人图像,行人图像对 (i, j) 的特征向量分别为 X_i 和 X_j , $X_i, X_j \in \mathbb{R}^d$, d 是特征维度.这些算法通过学习一个马氏距离相似度函数 $f_M(X_i, X_j)$ 来度量一对行人图像的相似度. MSF 形式如下式 (1) 所示:

$$f_M(X_i, X_j) = -(X_i - X_j)^T M (X_i - X_j) \quad (1)$$

其中, M 是一个半正定矩阵,用于保证相似度函数的对称性.

正如引言所述,MSF 只与特征差分空间有关,忽略了每个个体的独立特征,因此对于同一个人不同摄像机中的很大的外观变化的捕捉能力有限.本文通过学习一个双向关系相似度函数 $f_B(X_i, X_j)$ 来计算一对行人图像的相似度. BRSF 是 MSF 的泛化形式, BRSF 的形式如下式 (2) 所示:

$$f_B(X_i, X_j) = X_i^T A X_i + X_j^T A X_j + X_i^T B X_j + X_j^T B X_i \quad (2)$$

其中 $A, B \in \mathbb{R}^{d \times d}$. 因为一对行人图像的相似度有着对称性,所以 A 和 B 不需要是半正定矩阵.在 BRSF 中, $X_i^T A X_i + X_j^T A X_j$ 描述了一对行人图像的自相关关系, $X_i^T B X_j + X_j^T B X_i$ 描述了一对行人图像的互相关关系.从式 (2) 中可以看出 BRSF 是 MSF 的泛化形式,当 $B = -A$ 时,可以从 BRSF 得到 MSF 的形式. $B = -A$ 表示 BRSF 的自相关矩阵和互相关矩阵互为负矩阵.因此,

MSF 仅与一对行人图像的特征差分空间有关, 忽略了一对行人图像中每个个体的独立特征, 而 BRSF 增强了一对行人图像中独立个体的特征, 捕捉用于行人再识别的区分性信息的能力更强.

为了说明上述观点, 这里用一个一维空间的二分类例子来证明. 有两个类的样本分别服从 $[-10, -2]$ 的均匀分布和 $[2, 10]$ 的均匀分布. 当给定一个样本 $X=-3$ 和另一个样本 Y , 图 1 显示了采用 MSF 和 BRSF 两种相似度量函数, 样本 X 和样本 Y 的相似度的分布. 其中, $S(X)=S(Y)$ 表示 X 和 Y 属于同一个类, $S(X) \neq S(Y)$ 表示 X 和 Y 属于不同的类. 当采用 MSF, 即 $f_M(X, Y)=-m(X-Y)^2$, 如图 1(a) 所示, 同类样本和不同类样本之间的相似度取值范围存在重叠的部分. MSF 并不能完全分开上述两类样本, 在这里, $m=1$, 因为 m 并不改变两类样本之间相似度的相对关系. 当相似度量函数采用 BRSF, 相似度量函数中的自相关矩阵和互相关矩阵不再互为负矩阵, 增强了自相关关系的影响, 即 $f_B(X, Y)=-(X-Y)^2+3XY$, 如图 1(b) 所示, 上述两类样本被完全分开. 上述例子说明 BRSF 比 MSF 更具有区分性.

1.2 相似度量函数的学习

文献[12]中提出了一种 KISSME 的算法, 该方法没有使用到复杂的迭代优化的过程, 学习相似度量函数的时间花销小. 因此把 KISSME 度量学习算法的思想应用到 BRSF 中. 一对行人图像 (i, j) 的特征向量为 (X_i, X_j) , $X_i, X_j \in \mathbb{R}^d$, 则行人图像对 (i, j) 的相似度可以表示为式 (3):

$$\delta(X_i, X_j) = \log \left(\frac{p(X_i, X_j|H_1)}{p(X_i, X_j|H_0)} \right) \quad (3)$$

其中, $p(X_i, X_j|H_1)$ 表示一对行人图像对在假设 H_1 是相关行人图像对的条件下 (X_i, X_j) 相似的概率. $p(X_i, X_j|H_0)$ 表示一对行人图像对在假设 H_0 是不相关行人图像对条件下 (X_i, X_j) 相似的概率. 如果 $\delta(X_i, X_j)$ 的值是比较大的, 则说明假设 H_1 是成立的. 相反, $\delta(X_i, X_j)$ 的值比较小, 则说明假设 H_1 是不成立的, 则这对行人图像对是不相关行人图像对.

文献[12]假设行人图像对是高斯分布中的一个点, 则行人图像对的自相关关系和互相关关系分别用高斯分布来表示如下式 (4) 和式 (5):

$$\delta_A(X_i, X_j) = \log \left(\frac{\frac{1}{\sqrt{2\pi|\sum_{A_{y_{ij}=1}}|}} \exp \left[-\frac{1}{2} \left(X_i^T \sum_{A_{y_{ij}=1}}^{-1} X_i + X_j^T \sum_{A_{y_{ij}=1}}^{-1} X_j \right) \right]}{\frac{1}{\sqrt{2\pi|\sum_{A_{y_{ij}=0}}|}} \exp \left[-\frac{1}{2} \left(X_i^T \sum_{A_{y_{ij}=0}}^{-1} X_i + X_j^T \sum_{A_{y_{ij}=0}}^{-1} X_j \right) \right]} \right) \quad (4)$$

$$\delta_C(X_i, X_j) = \log \left(\frac{\frac{1}{\sqrt{2\pi|\sum_{C_{y_{ij}=1}}|}} \exp \left[-\frac{1}{2} \left(X_i^T \sum_{C_{y_{ij}=1}}^{-1} X_j + X_j^T \sum_{C_{y_{ij}=1}}^{-1} X_i \right) \right]}{\frac{1}{\sqrt{2\pi|\sum_{C_{y_{ij}=0}}|}} \exp \left[-\frac{1}{2} \left(X_i^T \sum_{C_{y_{ij}=0}}^{-1} X_j + X_j^T \sum_{C_{y_{ij}=0}}^{-1} X_i \right) \right]} \right) \quad (5)$$

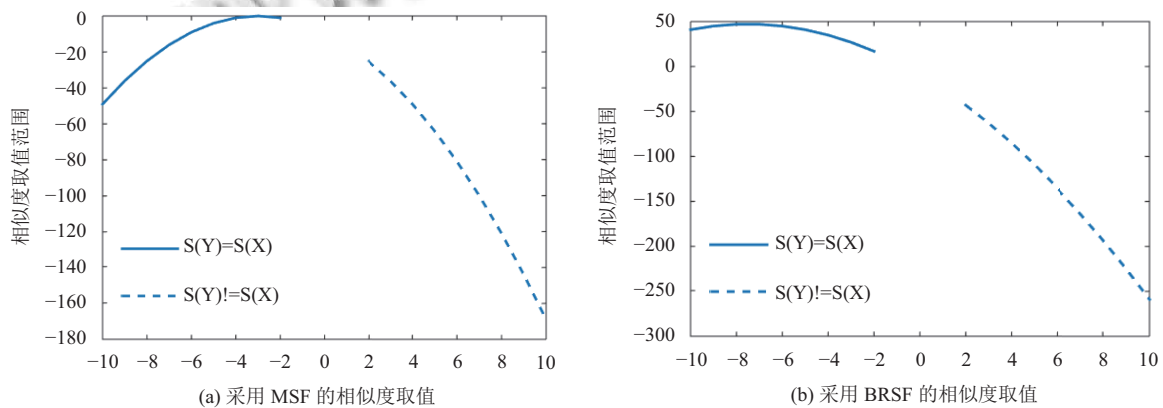


图 1 MSF 与 BRSF 相似度取值的对比

其中, $y_{ij}=1$ 表示行人图像对 (i, j) 属于相关行人图像对. 反之, $y_{ij}=0$ 表示行人图像对 (i, j) 属于不相关行人图像对. 相关行人图像对的样本分布在同一个高斯分布中, 不相关行人图像对的样本分布在另一个高斯分布中. 在这里,

$$\sum_{A_{y_{ij}=1}} = \sum_{y_{ij}=1} X_i X_i^T + X_j X_j^T \quad (6)$$

$$\sum_{A_{y_{ij}=0}} = \sum_{y_{ij}=0} X_i X_i^T + X_j X_j^T \quad (7)$$

$$\sum_{C_{y_{ij}=1}} = \sum_{y_{ij}=1} X_i^T X_j + X_j^T X_i \quad (8)$$

$$\sum_{C_{y_{ij}=0}} = \sum_{y_{ij}=0} X_i^T X_j + X_j^T X_i \quad (9)$$

把式 (4) 的右边进行 \log 简化, 我们重写式 (4) 得到式 (10).

$$\begin{aligned} \delta_A(X_i, X_j) = & X_i^T \sum_{A_{y_{ij}=0}}^{-1} X_i + X_j^T \sum_{A_{y_{ij}=0}}^{-1} X_j \\ & + \log\left(\left|\sum_{A_{y_{ij}=0}}\right|\right) - X_i^T \sum_{A_{y_{ij}=1}}^{-1} X_i \\ & - X_j^T \sum_{A_{y_{ij}=1}}^{-1} X_j - \log\left(\left|\sum_{A_{y_{ij}=1}}\right|\right) \end{aligned} \quad (10)$$

在式 (10) 中常数项只是提供了一个补偿, 不影响最后的结果, 我们可以把式中常数项去掉, 将式 (10) 简化得到式 (11):

$$\begin{aligned} \delta_A(X_i, X_j) = & X_i^T \left(\sum_{A_{y_{ij}=0}}^{-1} - \sum_{A_{y_{ij}=1}}^{-1} \right) X_i \\ & + X_j^T \left(\sum_{A_{y_{ij}=0}}^{-1} - \sum_{A_{y_{ij}=1}}^{-1} \right) X_j \end{aligned} \quad (11)$$

同理, 式 (5) 也可以简化重写得到式 (12):

$$\begin{aligned} \delta_C(X_i, X_j) = & X_i^T \left(\sum_{C_{y_{ij}=0}}^{-1} - \sum_{C_{y_{ij}=1}}^{-1} \right) X_i \\ & + X_j^T \left(\sum_{C_{y_{ij}=0}}^{-1} - \sum_{C_{y_{ij}=1}}^{-1} \right) X_j \end{aligned} \quad (12)$$

最后, 我们通过把式 (11) 和式 (12) 合并得到式 (13), 式 (13) 就是我们的 BRSF 的形式.

$$\delta(X_i, X_j) = X_i^T A X_i + X_j^T A X_j + X_i^T B X_j + X_j^T B X_i \quad (13)$$

在这里, 我们得到了:

$$A = \sum_{A_{y_{ij}=0}}^{-1} - \sum_{A_{y_{ij}=1}}^{-1}, B = \sum_{C_{y_{ij}=0}}^{-1} - \sum_{C_{y_{ij}=1}}^{-1} \quad (14)$$

2 基于 BRSF 学习的行人再识别系统

基于 BRSF 学习的行人再识别系统包括 3 个步骤: (1) 行人图像特征的提取; (2) BRSF 学习; (3) 查询目标行人图像. 整个系统流程如图 2 所示, 其中 BRSF 是利用数据集离线学习得到.

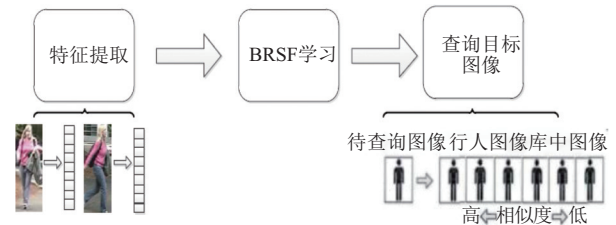


图 2 基于 BRSF 学习的行人再识别系统框架

2.1 行人图像特征的提取

颜色和纹理是描述行人图像非常重要的两个特征. 本文采用 HSV、YCbCr 和 Lab 三种颜色空间的颜色直方图和 SILTP^[14] (Scale Invariant Local Ternary Pattern) 纹理直方图来描述一张行人图像. 不同的颜色空间采用不同的衡量标准, 所以对同一张行人图像描述的侧重点也各不相同. SILTP 是著名的 LBP (Local Binary Pattern) 纹理描述算子的改进算法, LBP 算法的缺点是对图像噪声比较敏感, 所以 SILTP 算法还结合了 LTP (Local Ternary Patterns) 算法取得了对于图像噪声和光照变化具有更强的鲁棒性. 我们把一张行人图像归一化为 128×48 像素, 使用大小为 10×10 像素的滑动子窗口去描述行人图像的局部细节. 滑动子窗口的滑动方向是水平方向的, 每次滑动的步长为 5 个像素. 在每个子窗口中, 提取 HSV、YCbCr 和 Lab 三种颜色空间中每个通道的颜色直方图和 SILTP 纹理直方图. 为了解决不同摄像机下行人图像视角变化的问题, 在同一个水平位置上, 扫描所有子窗口, 取所有直方图相同 bin 的最大值作为该水平组的直方图特征. 这样得到的直方图特征不但对行人图像视角变化具有很好的鲁棒性, 而且可以捕捉到行人图像的局部细节特征. 图 3 显示滑动子窗口提取特征的过程.

在提取特征的过程, 我们考虑到了行人图像的多尺度信息. 在原始图像 128×48 像素的基础上进行两次金字塔降采样分别获得 64×24 像素的尺度和 32×12 的尺度图像. 在每个尺度的行人图像上重复上述的特征提取过程. 最后我们把所有特征级联形成行人图像的特性描述, 该特征的维数为 9000 维 ($3 \times 3 \times 16$ 维颜色特征 + 3^4 维 SILTP 特征) * ($24 + 11 + 5$ 水平组).

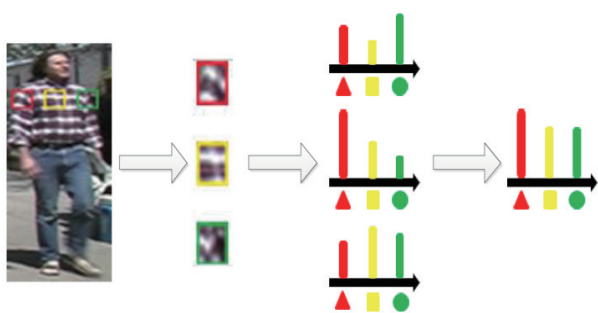


图3 滑动子窗口提取特征示意图

不同摄像机中行人视角的变化主要集中在水平方向上,而在垂直方向上并不很明显.因此,该文提取的特征对于水平方向的移动具有一定的不变性.为了说明本文提出的通过学习一个双向关系相似度函数,来计算一对行人图像的相似度的优越性.即使采用比较简单的特征来描述行人图像,算法性能也能优于一般的度量学习算法,后续的实验结果也会有所证明.

2.2 BRSF 学习

为了提高特征的有效性,便于 BRSF 的学习,实验中首先利用 PCA (Principle Component Analysis) 降维,在两个数据集的行人图像的特征维数进行降维,都保持 95% 以上的能量.在训练阶段,利用数据集离线学习得到一个 BRSF.当计算出 BRSF 的自相关矩阵 A 和互相关矩阵 B ,则根据式 (2) 计算一对行人图像 (i, j) 的相似度.

2.3 查询目标行人图像

行人再识别的研究目的是在不重叠的多摄像机视频监控系统中,根据输入的目标行人图像,检索该行人是否出现在这些视频中.该文假设通过行人检测技术,已经将监控视频的行人转换为行人图像库.利用离线学习得到的 BRSF,度量目标行人图像和行人图像库中所有图像的相似度.行人图像库中相似度高的图像为潜在的目标行人图像.

3 实验分析

算法通过大量的实验来测试,分别在 VIPeR 数据集^[15]和 QMUL GRID 数据集^[16]上进行了实验对比.实验中,随机选择 t 对行人图像对作为训练样本集,余下的行人图像对作为测试样本集.相机 A 中的行人图像作为查询目标库,相机 B 中的行人图像作为候选目标库.每对行人图像,任意选择一张图像加入查询目标库,

另一张则加入候选目标库.每个查询目标库与候选目标库中的每张行人图像都要有匹配.然后,再交换查询目标库和候选目标库.为了得到稳定的实验结果,以上过程重复 10 次,并将 10 次实验的平均值作为最终的实验结果.实验结果采用标准的累积匹配特性 (Cumulative Match Characteristic, CMC) 曲线评价算法性能,CMC 是指在候选目标库中匹配待查询的目标,前 r 个匹配结果中包含正确匹配结果的比率.其中,第一匹配率 (Rank=1) 为真正的识别能力,所以比较重要.但是当 Rank 值很小时,也是可以通过人眼进行辅助识别查找目标,因此也具有现实意义.

3.1 VIPeR 数据集的实验结果

VIPeR 数据集是行人再识别领域最常用的且最具有挑战性的数据集之一.它包含了 632 对行人图像,每对行人图像都是由两个不重叠视角的摄像机在不同的室外环境下获取的.在 632 对行人图像中存在着许多视角,亮度,背景有着较大变化的图像对.图 4 展示了在该数据集上,待查询图像和候选目标库中与其匹配的图像.在这两个实验中,最左边的行人图像为目标图像,右边的行人图像为候选目标库中根据相似度高低从左到右排序,其中虚线框内的行人图像为正确的匹配图像.图 5 展示了基于 MSF 和基于 BRSF 在候选目标库匹配待查询目标的实验结果,从图 5 中可以看出基于 BRSF 在相似度排序中的正确匹配图像比基于 MSF 的更加靠前的.

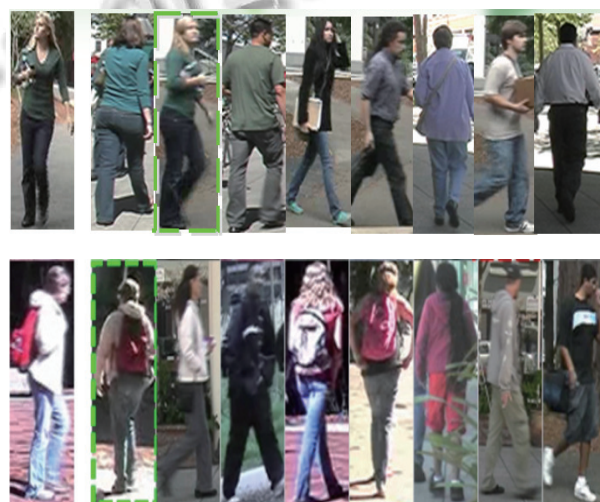


图4 本文算法在 VIPeR 数据集上的实验结果

实验中测试样本集和训练样本集均为 316 对行人图像. VIPeR 数据集上的其他实验,如果没有明确说明

测试集和训练集的个数,则都默认为316对行人图像。表1给出了本文算法与已有算法的性能比较的实验结果的数据。



(a) 基于 MSF 在候选目标库匹配待查询目标的实验结果



(b) 基于 BRSF 在候选目标库匹配待查询目标的实验结果

图5 基于 MSF 和基于 BRSF 在 VIPeR 数据集上的实验结果

表1 VIPeR 数据集上本文算法与已有算法性能对比实验结果(训练集样本规模 $t=316$ 对图像)(单位: %)

Method	Rank1	Rank10	Rank20
BRSF	53.21	92.43	96.56
XQDA ^[5]	40.00	80.51	91.08
RPLM ^[17]	27.00	69.10	83.00
rPCCA ^[18]	22.03	71.08	85.30
KISSME ^[12]	19.60	62.20	77.00
PCCA ^[19]	18.27	64.91	80.28
PRDC ^[11]	15.66	63.86	72.09

从表1可知,本文算法性能有较大的提升,尤其是Rank1比表中排第二的XQDA算法提升了约13.2%,并且在Rank20内都有着较高的识别率.在一定程度上,本文算法的效果已经能够应用到工程实践中,尤其是在刑事侦查等方面,刑侦人员可以在行人再识别返回的前 r 个结果中快速搜索出目标行人,大大提高侦查办案效率.

当训练集规模为 $t=200$ 时,本文算法与已有算法的性能对比结果如表2.

从表2中可知,在只有少量训练样本情况下,本文算法同样优于已有算法.由此可见,本文算法有效的解决了学习相似度度量函数中出现过拟合的问题.

表2 当训练集样本规模为 $t=200$ 时,本文算法与已有算法在VIPeR数据集上性能对比的实验结果(单位: %)

Methods	Rank1	Rank10	Rank20
BRSF	35.43	81.33	90.66
XQDA ^[5]	24.00	70.42	81.19
RPLM ^[17]	20.10	59.32	73.41
KISSME ^[12]	17.65	52.40	70.68
rPCCA ^[18]	12.53	41.73	65.21
PRDC ^[11]	12.36	40.36	63.12
PCCA ^[19]	10.27	36.61	57.54

为了说明本文提出的BRSF算法的优越性,表3给出了本文算法仅用一种特征情况下与其他算法效果对比.由于表中对比的算法都有HSV颜色空间特征和LBP纹理特征,因此实验中分别选用HSV和LBP作为BRSF算法提取的特征.

表3 仅用一种特征情况下的BRSF算法与已有算法性能比较实验结果(单位: %)

Methods	Features	Rank1	Rank10	Rank20
BRSF	HSV	47.40	88.25	94.50
BRSF	LBP	43.21	85.33	91.61
RPLM ^[17]	HSV, Lab, LBP	27.00	69.10	83.00
rPCCA ^[18]	RGB, YUV, HSV, LBP	22.03	71.08	85.30
KISSME ^[12]	HSV, Lab, LBP	19.60	62.20	77.00
PRDC ^[11]	YCbCr, HSV, LBP	15.66	63.86	72.09

由表3可知,该算法虽然只使用了一种特征,但是效果比其他采用多特征的算法更好.其中,RPLM算法融合了HSV、Lab和LBP等特征,Rank1仅有27.0%.而本文算法只用了HSV颜色特征,Rank1就达到了47.4%.当本文算法用到多特征时,算法的识别率又提升了一些,但是继续增加特征,算法识别率提升的幅度会越来越小,而算法的时间复杂度会越来越高.因此,本文算法在最终的特征选择上只选用了三种颜色空间特征和一种纹理特征.

评价度量学习算法性能好坏的一个重要因素是训练时间的长短,本文提出的BRSF算法是在基于KISSME度量学习算法上进行训练的,省略了迭代优化的时间,从而减少了训练时间.本文实验平台是8GB内存,Intel i5-2400 3.1 GHz CPU的PC台式机,实验环境是基于Matlab 2014b实验的.表4给出了本文算法与已有度量学习算法的训练时间比较实验结果,实验结果是取10组在VIPeR数据集上的训练时间的平均值.

表4 本文算法与已有算法的训练时间对比实验结果

Methods	BRSF	KISSME	XQDA	PCCA	PRDC
Time (s)	2.18	1.84	3.26	4.43	12.59

从表4可知,本文算法训练时间比其他大部分算法的耗时要短,只比KISSME算法耗时多了一些,因为KISSME训练得到MSF的一个矩阵,而本文算法基于KISSME上训练得到BRSF的两个矩阵,一个自相关矩阵,一个互相关矩阵。

3.2 QMUL GRID 数据集的实验结果

QMUL GRID数据集是目前行人再识别领域难度最高的数据集。QMUL GRID数据集上的图片是从地铁站的8个不相交的摄像机中拍摄获取的,其中包含了250对行人图像。除此之外,该数据集中还包含了额外的775张行人图像,这775张行人图像不属于250对行人图像中的任何一个行人,扩展了候选目标库。上述这些都体现了该数据集测试行人再识别的难度所在。图6展示了QMUL GRID数据集上的一些行人图像对样本。



图6 QMUL GRID数据集上的5对行人图像对

从图6中可以看到这些行人图像的图像质量,分辨率低,并包含了大量的照明和视角的变化。在QMUL GRID数据集的实验上,我们随机选择125对行人图像作为训练集,剩下的125对和额外的775张行人图像作为测试集。表5给出了本文算法在QMUL GRID数据集上与其他算法的比较实验结果。从表5可知,所有算法的识别率都不是特别高,但是本文算法的Rank1比表中排第二的XQDA算法提升了约7%,并且本文

算法的Rank10和Rank20比其他算法都有着显著的提升。

表5 QMUL GRID数据集本文算法与已有算法性能对比实验结果(单位:%)

Methods	Rank1	Rank10	Rank20
BRSF	23.96	54.56	63.24
XQDA ^[5]	16.56	41.84	52.40
rPCCA ^[18]	12.22	35.66	47.50
RPLM ^[17]	10.48	38.64	52.56
KISSME ^[12]	10.24	36.87	50.18
PRDC ^[11]	9.68	32.96	44.32
PCCA ^[19]	8.40	28.33	45.80

4 结论

行人再识别是一个具有很高应用价值同时也更具有挑战性的技术。提出了一种通过学习一个双向关系相似度函数,来计算一对行人图像的相似度。BRSF是MSF的泛化形式,与MSF相比,BRSF引入了一对行人的自相关关系和互相关关系,从而可以更好地捕捉同一个人在不同摄像机下的外观变化。为了学习BRSF,借助了KISSME的度量学习思想,把最终高斯分布的比值转换为BRSF的形式,得到一个对背景、视角、姿势的变化具有鲁棒性的相似度函数。由行人再识别的公共数据集VIPeR和QMUL GRID上的实验结果表明,基于双向关系度量学习得到有效的相似度函数,识别性能优于已有的行人再识别算法。

参考文献

- 1 Li Z, Chang SY, Liang F, *et al.* Learning locally-adaptive decision functions for person verification. Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA. 2013. 3610–3617.
- 2 Gong SG, Cristani M, Yan SC, *et al.* Person Re-identification. London: Springer, 2013. 301–313.
- 3 Farenzena M, Bazzani L, Perina A, *et al.* Person re-identification by symmetry-driven accumulation of local features. Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, CA, USA. 2010. 2360–2367.
- 4 Wang XY, Yang M, Zhu SH, *et al.* Regionlets for generic object detection. Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia. 2013. 17–24.
- 5 Liao SC, Hu Y, Zhu XY, *et al.* Person re-identification by

- local maximal occurrence representation and metric learning. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 2197–2206.
- 6 Liu Y, Shao Y, Sun FC. Person re-identification based on visual saliency. Proceedings of the 12th International Conference on Intelligent Systems Design and Applications. Kochi, India. 2012. 884–889.
- 7 Zhao R, Ouyang WL, Wang XG. Person re-identification by saliency matching. Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia. 2013. 2528–2535.
- 8 Luo Y, Jiang M, Wong YK, *et al.* Multi-camera saliency. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(10): 2057–2070. [doi: [10.1109/TPAMI.2015.2392783](https://doi.org/10.1109/TPAMI.2015.2392783)]
- 9 陈莹, 霍中花. 多方向显著性权值学习的行人再识别. 中国图象图形学报, 2015, 20(12): 1674–1683. [doi: [10.11834/jig.20151212](https://doi.org/10.11834/jig.20151212)]
- 10 Xing EP, Ng AY, Jordan MI, *et al.* Distance metric learning, with application to clustering with side-information. Proceedings of the 15th International Conference on Neural Information Processing Systems. Cambridge, MA, USA. 2003. 505–512.
- 11 Zheng WS, Gong SG, Xiang T. Person re-identification by probabilistic relative distance comparison. Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, CO, USA. 2011. 649–656.
- 12 Köstinger M, Hirzer M, Wohlhart P, *et al.* Large scale metric learning from equivalence constraints. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA. 2012. 2288–2295.
- 13 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. 自动化学报, 2016, 42(2): 299–308.
- 14 Liao SC, Zhao GY, Kellokumpu V, *et al.* Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes. Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, CA, USA. 2010. 1301–1306.
- 15 Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features. Proceedings of the 10th European Conference on Computer Vision. Marseille, France. 2008. 262–275.
- 16 Loy CC, Xiang T, Gong SG. Multi-camera activity correlation analysis. Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA. 2009. 1988–1995.
- 17 Hirzer M, Roth PM, Köstinger M, *et al.* Relaxed pairwise learned metric for person re-identification. Proceedings of the 12th European Conference on Computer Vision. Florence, Italy. 2012. 780–793.
- 18 Xiong F, Gou MR, Camps O, *et al.* Person re-identification using kernel-based metric learning methods. Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland. 2014. 1–16.
- 19 Mignon A, Jurie F. PCCA: A new approach for distance learning from sparse pairwise constraints. Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA. 2012. 2666–2672.