

基于稀疏学习的行人重识别算法

张文文 王洪元 万建武 孙金玉 丁宗元

(常州大学信息科学与工程学院, 常州, 213164)

摘要: 行人重识别问题是计算机视觉的重要研究内容之一, 旨在将多个非重叠相机中的目标行人准确加以识别。当将某摄像机中的行人图像视为目标行人在该摄像机视图上的一种表示时, 行人重识别可被认为是一种多视图学习问题。在此基础上提出的基于典型相关分析的行人重识别算法仅是一种线性降维算法, 很难从复杂的重识别系统(如目标行人图像受低分辨率、光照及行人姿态变化等因素影响)中提取有效的高层语义信息, 用于行人重识别。为此, 本文提出了一种基于稀疏学习的行人重识别算法(Sparsity learning based person re-identification, SLR)。SLR 首先通过稀疏学习获取目标行人在每一相机视图上的高层语义表示, 然后将高层特征映射到一个公共的隐空间, 使不同视图间的特征距离可比较。SLR 算法的优点在于通过学习鲁棒的行人图像特征表示, 能够获得更具判别性的公共隐空间, 以提高算法的行人重识别性能。在 VIPeR、CUHK 数据集上的实验结果表明了本文算法的有效性。

关键词: 计算机视觉; 行人重识别; 稀疏学习; 多视图学习

中图分类号: TP391.41 **文献标志码:** A

A Sparsity-Learning-Based Person Re-identification Algorithm

Zhang Wenwen, Wang Hongyuan, Wan Jianwu, Sun Jinyu, Ding Zongyuan

(School of Information Science and Engineering, Changzhou University, Changzhou, 213164, China)

Abstract: Person re-identification is one of the important issues addressed in computer vision. Existing recognition system concerns the matching of pedestrians across over-lapping cameras. When assuming pedestrian images as one representation of the camera view, person re-identification can be considered as a multi-view learning problem directly. On the basis of this assumption, a pedestrian recognition algorithm is proposed via canonical correlation analysis. Since the canonical correlation analysis is a linear dimensionality reduction algorithm, it is hard to extract semantic information for person re-identification (such as low resolution of images, changing illumination and other factors). A sparsity learning based person re-identification algorithm (SLR) is proposed. First, SLR obtained the semantic information of each camera view by the sparse learning, and then mapped the high-level features t into a public hidden space in order to make characteristic distance between different views can be compared. SLR aims to obtain more discriminable public hidden space. Finally, improve the matching rate of person re-identification across disjoint camera views. Comparing the proposed method and other common methods on the VIPeR dataset and CUHK campus dataset, experimental results show that the proposed method has

higher recognition efficiency.

Key words: computer vision; person re-identification; sparse learning; multi-view learning

引 言

行人重识别问题是计算机视觉的重要研究内容之一。它旨在从非重叠多摄像机监控系统中,匹配出不同摄像机视域中的目标行人。目前,在刑侦工作、图像检索等领域,行人重识别引起了学者们的极大关注^[1-2]。

在行人重识别系统中,由于监控视频中的拍摄场景存在着光照变化、目标图像分辨率低、拍摄视角变化、行人姿态各异及摄像机自身属性问题,导致同一行人在不同视角中拍出的行人图像差异很大,这为行人重识别带来了很大挑战。为解决这些难点问题,学者们提出了基于行人特征表示的方法^[3-7]和基于度量学习的方法^[8-12]。其中,基于行人特征表示的方法,主要关注提取更具有鉴别性的行人图像特征,仅利用标准距离(如欧氏距离和巴氏距离等)进行相似性度量。例如,文献[3]提出基于行人身体对称性的特征提取方法,将人划分头、躯干、腿部3个部分,然后提取除头部以外的累积颜色特征和纹理特征对行人进行描述。文献[5]将绘画结构应用于行人重识别,用类似身体外形的结构来表示行人,再提取行人颜色特征精确匹配。然而,在实际生活中,行人在经过非重叠的摄像机时,受到视角、光照等因素的影响,拍出的行人图像质量差。如果采用标准的距离度量方法,即等贡献考虑每一个特征,可能导致不同类样本间的度量距离小,而同类样本间度量距离大。为解决该问题,研究者们关注设计有效的距离度量函数,提出了基于度量学习的方法。代表性的工作有:文献[8]提出最大近邻分类间隔算法(Large margin nearest neighbor, LMNN),使得投影之后的同类点向内部紧缩,不同类点向外扩张。文献[11]将图像对投影到共同的特征空间,再对投影之后的图像特征进行相似性距离度量。

最近,Le等人认为,如果将不同视角下的行人图像视为该行人在不同相机视角上的多个描述,行人重识别也可被认为是一种多视图学习问题^[13-16]。为此,他们提出基于典型相关分析的行人重识别算法^[17-19]。典型相关分析(Canonical correlation analysis, CCA)^[17]是一种经典的多视图学习方法,它利用互补原理,最大化不同视图数据之间的相关性,减少数据之间的不确定性,从而达到增强识别能力的目的,最终提高行人重识别的准确率。目前,CCA 主要在多视图学习的特征融合方面有着广泛的应用,也常用于解决行人重识别的问题。

但是,CCA 是一种线性模型,它只关注成对样本之间的相关性,而实际中存在大量特征非线性的情况,当仅用 CCA 这样的简单线性模型来学习这些非线性相关现象时,将不可避免地出现欠拟合现象。为解决该问题,本文提出一种新的多视图学习方法,基于稀疏学习的行人重识别算法(Sparsity-learning-based person re-identification, SLR)。近年来,稀疏学习被广泛运用于人脸识别^[20]和行人重识别领域^[21-23]。稀疏学习能够通过特征选择去除与当前任务无关的特征,减少训练过程中的矩阵维数,降低任务难度,减少计算和存储开销。本文 SLR 算法的优点在于通过学习鲁棒的行人图像特征表示,能够获得更为判别的公共隐空间,以提高算法的行人重识别性能。SLR 通过稀疏表示获取目标行人在每一相机视图上的高层语义表示,再将高层特征的数据映射到一个公共隐空间,使不同视图间的特征距离可比较,实现了对多视角数据的充分利用。

1 基于典型相关分析的行人重识别

CCA^[20]是一种多元统计分析方法,目的是将多视图数据投影到一个公共的空间,使得视图间的相

关性最大化。给定多视图数据 $\mathbf{D}^A = \{d_i^A \in \mathbf{R}^m, i=1, 2, \dots, N\}$ 和 $\mathbf{D}^B = \{d_i^B \in \mathbf{R}^n, i=1, 2, \dots, N\}$, 其中 m 和 n 表示样本特征维数, N 代表了样本数。CCA 旨在找到两个投影矩阵 $\mathbf{W}_A \in \mathbf{R}^m$ 和 $\mathbf{W}_B \in \mathbf{R}^n$ 来最大化 \mathbf{D}^A 和 \mathbf{D}^B 之间的相关性, 其目标函数表达式如下

$$\max_{\mathbf{W}_A, \mathbf{W}_B} \frac{\mathbf{W}_A^T \mathbf{C}_{AB} \mathbf{W}_B}{\sqrt{\mathbf{W}_A^T \mathbf{C}_{AA} \mathbf{W}_A} \sqrt{\mathbf{W}_B^T \mathbf{C}_{BB} \mathbf{W}_B}} \quad (1)$$

式中, $\mathbf{C}_{AA} = E[\mathbf{D}^A \mathbf{D}^{A^T}]$, $\mathbf{C}_{BB} = E[\mathbf{D}^B \mathbf{D}^{B^T}]$, $\mathbf{C}_{AB} = E[\mathbf{D}^A \mathbf{D}^{B^T}]$, $E[\cdot]$ 为数学期望。

如果将行人图像的 Probe 集 $\{\mathbf{F}_i^p, i=1, 2, \dots, M\}$ 和 Gallery 集 $\{\mathbf{F}_i^g, i=1, 2, \dots, M\}$, 分别视作式(1)中的视图 A 和 B 的原始特征。经过式(1)求得的投影矩阵 \mathbf{W}_A 和 \mathbf{W}_B 投影之后变成 $\{f_i^p, i=1, 2, \dots, M\}$ 和 $\{f_i^g, i=1, 2, \dots, M\}$, M 为测试集样本数。利用余弦相似性度量函数来实现行人的重识别过程, 当余弦值接近 1 时, 样本间相似性最大; 余弦值越小, 样本间相似性最小。行人的重识别过程就是要从 Gallery 集的 M 个行人中找到第 k 个行人使该行人与 Probe 集中目标行人的余弦距离最大, 即第 k 个行人极有可能与目标行人是同一个人。具体计算如下

$$\operatorname{argmax}_k \frac{(f^p)^T f_k^g}{\|f^p\| \|f_k^g\|} \quad (2)$$

2 基于稀疏学习算法的行人重识别

假设: $\{\mathbf{I}_i^X, i=1, 2, \dots, N\}$ 和 $\{\mathbf{I}_i^Y, i=1, 2, \dots, N\}$ 是来自两个不同计算机视角的 N 个行人图像。本文提出的 SLR 算法在训练阶段, 提取每一张行人图像的特征(如颜色特征和纹理特征等), 分别记作: $\mathbf{X} = [x_1, x_2, \dots, x_N] \in \mathbf{R}^{d_x \times N}$ 和 $\mathbf{Y} = [y_1, y_2, \dots, y_N] \in \mathbf{R}^{d_y \times N}$, 其中 d_x 和 d_y 分别为样本 \mathbf{X} 和 \mathbf{Y} 的特征维数。

2.1 高层语义特征的提取

CCA 作为一种简单线性模型, 只关注成对样本之间的相关性, 而实际中存在大量特征非线性的情况, 当仅用 CCA 模型来学习这些非线性相关现象时, 将不可避免地出现欠拟合现象。文献[21]利用稀疏学习得到的字典将行人图像特征线性表达出来, 对输入样本按照得到的字典在指定约束条件下线性展开, 再对样本进行分类, 有效提高了行人匹配率。文献[22]将行人重识别问题作为一个块稀疏恢复问题, 并使用交替方向框架解决了相关优化问题。

近年, 稀疏学习被广泛应用于人脸识别和行人重识别领域, 也是信号处理领域的热点之一。稀疏学习通过尽可能少的训练样本找到对应测试样本的稀疏表示, 这样的突出优点是计算速度更快。更重要的是, 稀疏学习能够提取带语义信息的高层特征, 有利于接下来的分类、识别和检索。

本文提出的 SLR 算法通过将行人样本特征稀疏表示出来, 使行人样本在该模型下的稀疏字典只存在少量的非零元素, 使得特征: (1) 拥有更强大表达能力; (2) 拥有识别和约束能力。SLR 通过提取样本带语义信息的高层特征, 克服现有方法只抽取行人图像低层特征的缺点, 因此在面临多视角数据的处理时, 忽视了多视角数据本身蕴含的内在关联属性, 从而无法对多视角数据实现有效处理的不足。表达式如下

$$O_1(\mathbf{U}_X, \mathbf{V}_X, \mathbf{U}_Y, \mathbf{V}_Y) = \|\mathbf{X} - \mathbf{U}_X \mathbf{V}_X\|_F^2 + \lambda \sum_{i=1}^N |\mathbf{V}_X(:, i)|_1 + \|\mathbf{Y} - \mathbf{U}_Y \mathbf{V}_Y\|_F^2 + \lambda \sum_{i=1}^N |\mathbf{V}_Y(:, i)|_1 \quad (3)$$

式中, $\mathbf{V}_X = [v_{x1}, v_{x2}, \dots, v_{xN}] \in \mathbf{R}^{t \times N}$, $\mathbf{V}_Y = [v_{y1}, v_{y2}, \dots, v_{yN}] \in \mathbf{R}^{t \times N}$ 分别是样本 \mathbf{X} 和 \mathbf{Y} 的稀疏表示, \mathbf{X} 和 \mathbf{Y} 对应的稀疏字典为 $\mathbf{U}_X = [u_{x1}, u_{x2}, \dots, u_{xt}] \in \mathbf{R}^{d_x \times t}$ 和 $\mathbf{U}_Y = [u_{y1}, u_{y2}, \dots, u_{yt}] \in \mathbf{R}^{d_y \times t}$, t 是小于样本数 N 的稀疏系数。表达式中 λ 为平衡参数, λ 所在项为稀疏项, λ 值越大, 稀疏约束越大。

2.2 隐空间映射

实际情况中,提取出的高层特征并不在同一空间中,无法直接比较样本两个视角特征之间的相关性。本文通过隐空间映射,将显著的行人特征进行线性重构,即将特征旋转到同一个空间中,使得特征之间可比较,让两视角特征能够自动反映出行人样本之间的隐含关系,从而提高行人重识别的匹配率。具体就是将一个样本两个视角的稀疏表示 \mathbf{V}_X 和 \mathbf{V}_Y 通过式(4)得到 $\mathbf{R} \in \mathbf{R}^{t \times t}$,再利用参数 \mathbf{R} 对这两种特征稀疏表示进行矫正。

$$O_2(\mathbf{V}_X, \mathbf{V}_Y, \mathbf{R}) = \|\mathbf{V}_X - \mathbf{R}\mathbf{V}_Y\|_F^2 \quad (4)$$

2.3 目标函数

结合上面两个方面,SLR 算法的目标函数表达式如下

$$\begin{aligned} \min_{\mathbf{U}_X, \mathbf{V}_X, \mathbf{U}_Y, \mathbf{V}_Y, \mathbf{R}}: & O_1 + \mu O_2 \\ \text{s. t.} & \|\mathbf{U}_X(:, i)\|^2 \leq 1, \|\mathbf{U}_Y(:, i)\|^2 \leq 1, \|\mathbf{R}(:, i)\|^2 \leq 1, \forall i \end{aligned} \quad (5)$$

式中 μ 为平衡参数。

2.4 最优化求解

式(5)非凸,无法直接求出变量 $\mathbf{V}_X, \mathbf{V}_Y, \mathbf{U}_X, \mathbf{U}_Y, \mathbf{R}$ 的值。但若固定其他变量,式(5)对于剩余的一个变量则是凸函数。为此,本文采用交替迭代方法进行求解。

第 1 步:先随机生成 5 个变量 $(\mathbf{V}_X, \mathbf{V}_Y, \mathbf{U}_X, \mathbf{U}_Y, \mathbf{R})$ 的初始值。

第 2 步:为更新式(3)中 \mathbf{V}_X ,先固定 $\mathbf{V}_Y, \mathbf{U}_X, \mathbf{U}_Y, \mathbf{R}$ 4 个参数的值,即式(3)中样本 \mathbf{Y} 的两项均为已知项,得到

$$\min_{\mathbf{V}_X} \|\mathbf{X} - \mathbf{U}_X \mathbf{V}_X\|_F^2 + \lambda \sum_{i=1}^N |\mathbf{V}_X(:, i)|_1 \quad (6)$$

第 3 步:将式(6)改写成式(7),结合 Liu 等人提出的稀疏学习 SLEP 软件包(Sparse learning package)^[24],用 SLEP 软件包来更新 \mathbf{V}_X

$$\min_{\mathbf{V}_X} \left\| \begin{bmatrix} \mathbf{X} \\ \sqrt{\mu} \mathbf{R} \mathbf{V}_Y \end{bmatrix} - \begin{bmatrix} \mathbf{U}_X \\ \sqrt{\mu} \mathbf{I} \end{bmatrix} \mathbf{V}_X \right\|_F^2 + \lambda \sum_{i=1}^N |\mathbf{V}_X(:, i)|_1 \quad (7)$$

第 4 步:同理,为更新 \mathbf{V}_Y 的值,固定其余 4 个参数的值,得到

$$\min_{\mathbf{V}_Y} \|\mathbf{Y} - \mathbf{U}_Y \mathbf{V}_Y\|_F^2 + \lambda \sum_{i=1}^N |\mathbf{V}_Y(:, i)|_1 \quad (8)$$

第 5 步:将式(8)改写成式(9),用 SLEP 软件包来更新 \mathbf{V}_Y

$$\min_{\mathbf{V}_Y} \left\| \begin{bmatrix} \mathbf{Y} \\ \sqrt{\mu} \mathbf{V}_X \end{bmatrix} - \begin{bmatrix} \mathbf{U}_Y \\ \sqrt{\mu} \mathbf{R} \end{bmatrix} \mathbf{V}_Y \right\|_F^2 + \lambda \sum_{i=1}^N |\mathbf{V}_Y(:, i)|_1 \quad (9)$$

第 6 步:利用式(10)更新 $\mathbf{U}_X, \mathbf{U}_Y$ (仅写出关于样本 \mathbf{X} 的表达式,样本 \mathbf{Y} 的表达式与式(10)类似)

$$\begin{aligned} \min_{\mathbf{U}_X} & \|\mathbf{X} - \mathbf{U}_X \mathbf{V}_X\|_F^2 \\ \text{s. t.} & \|\mathbf{U}_X(:, i)\|^2 \leq 1, \forall i \end{aligned} \quad (10)$$

每一次的迭代计算,都得到式(5)中的目标函数值,重复多次迭代,直到目标函数值趋于稳定收敛,表示得到一定意义下的最好字典矩阵,用它能将样本特征稀疏表示出来。SLR 算法的具体描述如下:

输入:特征向量 \mathbf{X} 和 \mathbf{Y} ,随机生成 5 个参数($\mathbf{U}_X, \mathbf{V}_X, \mathbf{U}_Y, \mathbf{V}_Y$ 和 \mathbf{R})的初始值

循环:

Step 1: 利用公式(7)和公式(9)更新 \mathbf{V}_X 和 \mathbf{V}_Y 的值

Step 2: 利用公式(10)更新 \mathbf{U}_X 和 \mathbf{U}_Y 的值

Step 3: 利用公式(4)更新 \mathbf{R} 的值

Step 4: 目标函数未达到收敛条件,返回 Step 1, 否则输出

输出:

$\mathbf{U}_X, \mathbf{U}_Y$ 和 \mathbf{R}

3 行人的重识别

行人的重识别就是正确地从 Gallery 集中找出与 Probe 集中相对应的那个行人。假设 B 摄像头中样本 $\mathbf{I}_p \in \mathbf{R}^{d_p \times M}$ 为 Probe 集,摄像头 A 中样本 $\mathbf{I}_g \in \mathbf{R}^{d_g \times M}$ 为 Gallery 集, d_p 和 d_g 分别是 Probe 集和 Gallery 集中样本特征的维数, M 为样本数。

利用第 2 节中样本集 \mathbf{X} 和 \mathbf{Y} 的稀疏字典 \mathbf{U}_X 和 \mathbf{U}_Y 分别计算出 Gallery 集和 Probe 集的稀疏表示 \mathbf{V}_g 和 \mathbf{V}_p 。 \mathbf{V}_g 的具体求解公式如下(求 \mathbf{V}_p 的表达式与式(11)类似):

$$\min_{\mathbf{U}_x} \|\mathbf{I}_g - \mathbf{U}_x \mathbf{V}_g\|_F^2 \tag{11}$$

式(12)所示的最优化问题的解,就是 M 个 Gallery 集行人中与 Probe 集中目标行人最为相似的第 i 个行人。其中,矫正矩阵 \mathbf{R} 由 SLR 算法第 3 步迭代计算得到。

$$\min_i \|\mathbf{V}_g - \mathbf{R} \mathbf{V}_p\|_F^2 \tag{12}$$

4 实验

4.1 数据集

VIPeR 数据集:VIPeR 数据集是用于行人重识别研究的重要数据集之一^[17]。它包含 632 个行人图像对,每个行人在不同的视角下的身体姿态变化大,且都只有一张图像。因此,该数据中的行人图像总计 1 264 张。每一张图像的大小为 128 像素 \times 48 像素。

CUHK01 数据集:CUHK 数据集包含 3 个子集。本文实验选取 CUHK01 子集来做实验, CUHK01 数据集包含 971 个行人^[14],每个人在每个视角中有两张图像,共 3 884 张图。在实验中,任意选择其中的一张来组成 Gallery 集。同 VIPeR 数据集一样, CUHK01 数据集中的图像大小为 128 像素 \times 48 像素。

4.2 特征提取和实验参数设置

仿照文献[25]中的特征提取方法,将每张行人图像平均分成 18 个水平条块,分别提取每个水平条块的 RGB, HSV, YCbCr, Lab, YIQ 和 Gabor6 种特征,前 5 个提取的是图像的颜色空间特征,即直方图特征;Gabor 特征是一种纹理特征,根据不同波长、方向、带宽等分别取 16 组不同的 Gabor 滤波器。每个水平条块的每种特征空间的每个通道分为 16 维直方图统计特征,RGB 和 HSV 特征有 3 个通道, YCbCr, Lab 和 YIQ 特征有 2 个通道, Gabor 特征有 16 个通道,于是,每个水平条块的特征为 448 维。因此,每张行人图像在特征空间中被表示为 $(3+3+2+2+2+16) \times 16 \times 18 = 8\ 064$ 维特征向量,再将所有特征组合成一个列向量来代表一个行人图像作为实验的原始特征。

实验中选择 N 人作为测试集,其余用来训练。每个测试集中包含一个 Probe 集和一个 Gallery 集。为了让实验结果更具有说服力,整个实验过程进行 10 次,然后取平均值作为最终结果。

根据经验设置参数的值,例如在 VIPeR 数据集上实验时,本文取 $\lambda=0.04, \mu=0.5$ 。在 CUHK01 数据集上实验时,取 $\lambda=0.05, \mu=0.4$ 。在 4.3.3 节,对算法的参数敏感性进行了分析。

为了评判本文方法,用累计匹配特征曲线(Cumulative matching characteristic, CMC)来展示实验最终的匹配率。CMC 曲线表示在 Gallery 集中前 r 个搜索结果中(横向坐标)正确找到 Probe 集中待查询人的比率(垂直坐标)。

4.3 实验结果

本文实验在 MATLAB 2010b 平台下完成,实验环境为 CPU Intel Core(TM) i5-4460T 1.90 GHz,内存 8 GB。

4.3.1 算法收敛性分析

在 VIPeR, CUHK01 数据集上,采用 4.2 节的参数设置,对算法的收敛性进行了分析。实验结果如图 1 所示。由图 1 可知,本文算法在两个数据集上,通过 10 次迭代,即可收敛,目标函数的值几乎不变。这说明,本文算法只需少量的迭代次数,即可收敛,收敛速度快。在实验过程中,将迭代次数统一设为 30。

通过迭代 30 次,得到恰当字典矩阵 $U_X(U_Y)$ 和样本 $X(Y)$ 的稀疏表示 $V_X(V_Y)$,并利用这 4 个参数值来进行下一步实验。

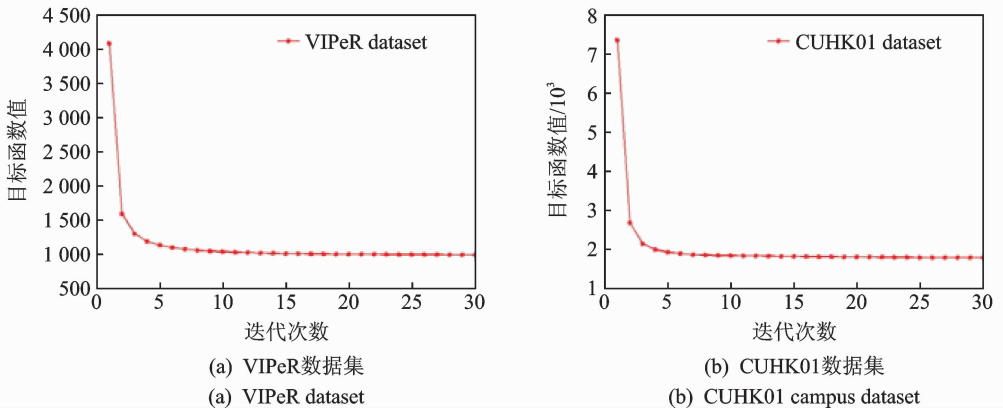


图 1 收敛性分析

Fig. 1 Convergence analysis

4.3.2 结果分析

为了评判 SLR 算法的好坏,本文将 SLR 算法与正则化典型相关分析算法(Regularization canonical correlation analysis, RCCA)^[24]和 L1-norm 算法^[11]在 VIPeR 数据集和 CUHK01 数据集上分别进行比较,并分别画出前 20 个搜索结果的曲线图。RCCA 方法是基于典型相关分析的改进算法,通过在典型相关分析基础上增加一个正则化项,来解决实际存在的图像特征维数远大于图像数目导致特征向量不存在的问题,能并有效防止过拟合。其次, RCCA 也是一种多视图学习算法,选取 RCCA 算法与 SLR 进行比较更能直观看出,两种多视图学习方法的性能。L1-norm 算法是最基础的距离度量方法,通过直接计算特征之间距离绝对值大小来判断两视图间的相关性,常用于跟其他算法进行对比实验。在训练集和测试集设置相同的情况下,具体方法比较结果如图 2 所示。

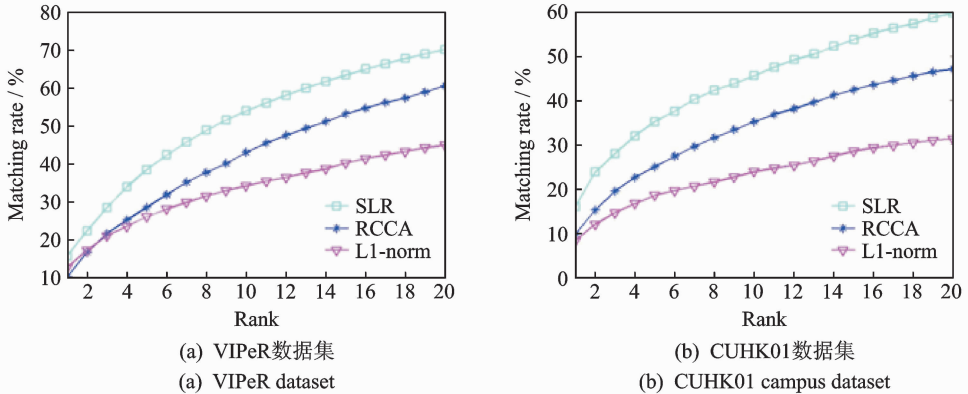


图2 几种算法的结果比较

Fig. 2 Performance comparison using CMC curves

VIPeR 数据集中每个行人只有 2 张图像,行人姿态的不确定性给行人重识别造成很大困难,通过图 2(a)可以看出: SLR 算法的第一匹配率达到 15.66%,而且随着排名等级的增大,SLR 算法的匹配率有了明显的提升,在 Rank20 时,SLR 算法匹配率达到 70.09%,明显高于其他两个算法。

CUHK01 数据集中每个行人都有 4 张图像,随机选取图像作为 Gallery 集存在很大偶然性,大量的数据和高维的图像特征使得很多方法在此数据集上效果均不佳。从图 2(b)可以看出,SLR 算法结果较另两个方法有很大的提高,第一匹配率达到 15.98%,比其他两个方法高出约 5%,整体匹配率也明显高于其他。

4.3.3 参数敏感性分析

参数的设置对实验结果的影响不容忽视,通过设置不同的参数查看参数变化对实验结果的影响很必要。因此,在实验中,通过查看参数 λ 和 μ 在不同取值下得到的最终实验结果来寻找最优参数值,使算法效果达到最佳。

为了检验参数 λ 和 μ 取值对算法性能的影响程度,设置几组不同取值的 λ 和 μ 在两个数据集上分别进行测试,结果如图 3,4 所示。

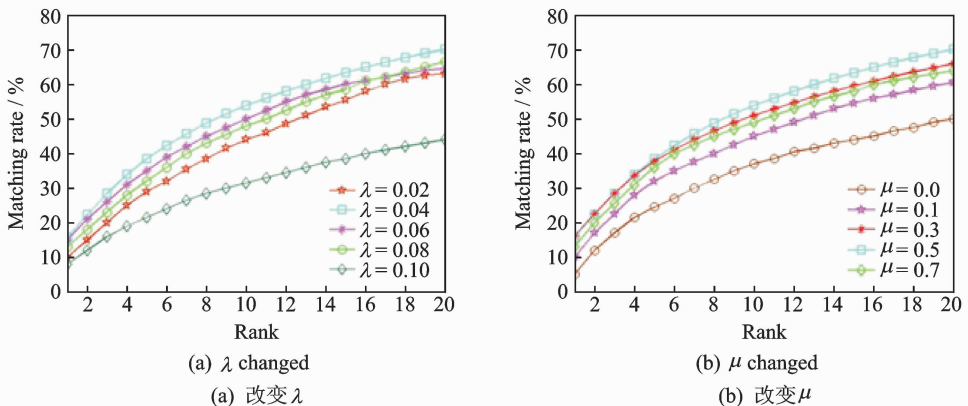


图3 VIPeR 数据集中不同参数值下算法结果比较

Fig. 3 Performance comparison on different values in the VIPeR dataset

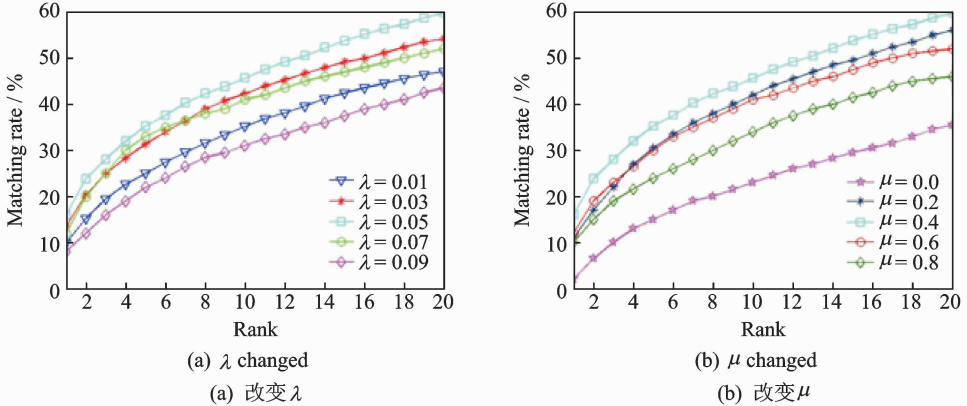


图4 CUHK01数据集中不同参数值下算法结果比较

Fig. 4 Performance comparison on different values in the CUHK01 dataset

对 VIPeR 数据集来说,从图 3(a)可知,当 $\mu=0.5$ 不变时, λ 值从 0.02 增加到 1,性能逐渐由小幅提升转为大幅下降,在 $\lambda=0.04$ 时性能最好, $\lambda=0.1$ 时性能最差;图 3(b)中,当 $\lambda=0.04$ 不变时, μ 值从 0.1 增加到 1,性能由小幅提升转为小幅下降,当 $\mu=0.5$ 时达到最大值。因此,在 VIPeR 数据集上,取 $\lambda=0.04, \mu=0.5$ 。

对 CUHK01 数据集来说,从图 4(a)可知,当 $\mu=0.4$ 不变时, λ 值从 0.01 增加到 0.1,性能由提升转为下降,在 $\lambda=0.05$ 时性能最好;图 4(b)中,当 $\lambda=0.05$ 不变时, μ 值从 0.2 增加到 1,在 $\mu=0.4$ 时匹配率最高,当 μ 值继续增大时,匹配率开始减小。因此,在 CUHK01 数据集上,取 $\lambda=0.05, \mu=0.4$ 。

5 结束语

在已有行人重识别领域的算法中,不乏有研究者从稀疏学习的角度来解决行人重识别问题,例如文献[26]结合度量学习和稀疏表示提出新的特征距离排序方法,利用稀疏表示对行人重识别的图像特征进行距离度量排序。与之不同,本文提出一种基于稀疏学习的行人重识别算法。SLR 首先通过稀疏学习获取目标行人在每一相机视图上的高层语义表示,然后将高层特征映射到一个公共的隐空间,使得不同视图间的特征距离可比较。SLR 算法的优点在于通过学习鲁棒的行人图像特征表示,能够获得更具判别性的公共隐空间,以提高算法的性能。在 VIPeR, CUHK01 数据集上的实验结果表明了本文算法的有效性。

本文 SLR 算法是一个无监督的学习方法。在未来将更多考虑融入监督信息,通过提高公共隐空间的判别能力,提升算法的行人重识别能力。

参考文献:

- [1] Zhu Jing, Roy S. MAC for dedicated short range communications in intelligent transport system [M]. Piscataway: IEEE Press, 2003; 60-67.
- [2] Blum J J, Eskandarian A, Hoffman L J. Challenges of intervehicle ad hoc network[M]. Piscataway: IEEE Press, 2004; 347-351.
- [3] Farenzena M, Bazzani L, Perina A, et al. Person re-identification by symmetry-driven accumulation of local features[C]// Computer Vision and Pattern Recognition. New York: IEEE, 2010; 2360-2367.

- [4] Shet V, Khamis S, Kuo C H. Person re-identification using semantic color names and rankboost[C]//Proceeding of the IEEE Workshop on Applications of Computer Vision. Clearwater Beach; IEEE Computer Society, 2013; 281-287.
- [5] Dong S C, Cristani M, Stoppa M, et al. Custom pictorial structures for re-identification[C]//Proceeding of the British Machine Vision Conference. Nottingham; University of Nottingham, 2014; 1-11.
- [6] Hirzer M, Beleznai C, Ranzato L, et al. Person re-identification by descriptive and discriminative classification [C]//Proceeding of the Scandinavian Conference on Image analysis. Berlin; Springer-Verlag, 2011; 91-102.
- [7] 赵猛, 张化祥, 刘丽. 基于颜色和边缘方向的图像检索方法[J]. 数据采集与处理, 2016, 31(3):577-583.
Zhao Meng, Zhang Huaxiang, Liu Li. Image retrieval method based on color and edge direction[J]. Journal of Data Acquisition and Processing, 2016, 31(3):577-583.
- [8] Weinberger K Q, Saul L K. Distance metric learning for large margin nearest neighbor classification [J]. Journal of Machine Learning Research, 2009, 10(F2b): 207-244.
- [9] Zheng Weishi, Gong Shaogang, Xiang Tao. Person re-identification by probabilistic relative distance comparison[C]//Computer Vision and Pattern Recognition. New York; IEEE, 2011; 649-656.
- [10] Pedagadi S, Orwell J, Velastin S, et al. Local fisher discriminant analysis for pedestrian re-identification[C]//Computer Vision and Pattern Recognition. New York; IEEE, 2013;3318-3325.
- [11] Li Wei, Wang Xiaogang. Locally aligned feature transforms across views[C]//Computer Vision and Pattern Recognition. New York; IEEE, 2013;3594-3601.
- [12] Mignon A, Jurie F. PCCA: A new approach for distance learning from sparse pairwise constraints[C]//Computer Vision and Pattern Recognition. New York; IEEE, 2012; 2666-2672.
- [13] Meina K, Shan Shiguang, Zhang Haihong. Multi-view discriminant analysis[M]. Berlin; Springer Berlin Heidelberg, 2012; 808-821.
- [14] Diethe T, Hardoon D R, Shawe-Taylor J. Multi-view fisher discriminant analysis[C]//NIPS Workshop on Learning from Multiple Sources. Vancouver: [s. n.], 2008;976-983.
- [15] Hou Chenping, Zhang Changshui, Wu Yi. Multiple view semi-supervised dimensionality reduction [J]. Pattern Recognition, 2010, 43(3):720-730.
- [16] Cheng Xiaohong, Chen Songcan, Xue Hui, et al. A unified dimensionality reduction framework for semi-paired and semi-supervised multi-view data [J]. Pattern Recognition, 2012, 45(5):2005-2018.
- [17] Le an, Kafai M, Yang Songfan, et al. Reference-based person re-identification[C]//Proceeding of the IEEE Conference on Advanced Video and Signal Based Surveillance. Clearwater Beach; IEEE Computer Society, 2013;244-249.
- [18] Le an, Yang Songfan, Bhanu B. Person re-identification by robust canonical correlation analysis [J]. IEEE Signal Processing Letters, 2015, 22(8):1103-1107.
- [19] Le an, Kafai M, Yang Songfan, et al. Person re-identification with reference descriptor [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(4):776-787.
- [20] Wan Jianwu, Wang Hongyuan, Yang Ming. Cost sensitive semi-supervised canonical correlation analysis for multi-view dimensionality reduction [J]. Neural Processing Letters, 2017, 45(2):411-430.
- [21] Lisanti G, Masi I, Bagdanov A D, et al. Person re-identification by iterative re-weighted sparse ranking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(8):1629-42.
- [22] Karanam S, Li Y, Radke R J. Sparse re-id: Block sparsity for person re-identification[C]//Computer Vision and Pattern Recognition. New York; IEEE, 2015;33-40.
- [23] Mirmahboub B, Kiani H, Bhuiyan A, et al. Person re-identification using sparse representation with manifold constraints [C]//Proceeding of the IEEE Conference on Image Processing. New York; IEEE, 2016;774-778.
- [24] Liu Jun, Ye Jieping. Efficient Euclidean projections in linear time [J]. Journal of Machine Learning Research, 2009, 12(7): 3371-3412.
- [25] Chen Yingcong, Zheng Weishi, Lai Jianhuang. Mirror representation for modeling view-specific transform in person re-identification[C]//Proceeding of the International Joint Conference Artificial Intelligence. Palo Alto; AAAI Press, 2015; 3402-3408.
- [26] 丘宇辉. 基于度量学习和稀疏表示的行人重识别技术研究[D]. 广州:华南理工大学, 2015.

Qiu Yuhui. Research on pedestrian recognition based on metric learning and sparse representation[D]. Guangzhou: South China University of Technology, 2015.

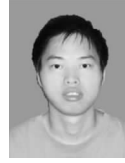
作者简介:



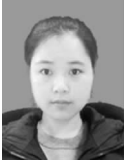
张文文(1994-),女,硕士研究生,研究方向:智能图像处理与模式识别,E-mail:564331135@qq.com。



王洪元(1960-),男,教授,研究方向:智能图像处理与模式识别等。



万建武(1986-),男,讲师,研究方向:机器学习与模式识别。



孙金玉(1991-),女,硕士研究生,研究方向:智能图像处理与模式识别。



丁宗元(1991-),男,硕士研究生,研究方向:智能图像处理与模式识别。

(编辑:张彤)