

基于无监督的非对称度量学习优化行人再识别

江雨燕¹, 吕 魏¹, 李 平², 邵 金¹

(1. 安徽工业大学 管理科学与工程学院, 安徽 马鞍山 243032;

2. 南京邮电大学 计算机学院, 江苏 南京 210023)

摘要:行人再识别是对同一行人不同视图进行相似性匹配的技术, 针对匹配过程中出现的不同相机、相机角度以及行人姿势、有无光照和障碍物等因素结合度量学习进行相似性度量。随着外界条件的变化, 由于忽略特定视图特征易导致部分信息丢失, 视图有可能展现出不同的分布。结合非对称的无监督度量学习方法 UAML (Unsupervised Asymmetric Metric Learning) 把视图分布分为共享视图和特定视图, 共享视图提取共享特征, 特定视图提取与视图相关联的特征投影到公共子空间中, 再通过 K-means 聚类, KL (Kullback - Leibler) 散度对其进行优化, 这样得到的子空间样本具有相同分布及全面的表示形式。该方法在 VIPeR 数据集、CUHK01 数据集、Market-1501 数据集上进行了实验测试, 使用累计匹配曲线 (Cumulative Match Characteristic, CMC)、Rank-1 精度和平均精度曲线等指标来衡量算法的性能。其中, 秩-1 精度在不同数据上分别达到了 40.25%、56.6%/58.09%、61.67%。该方法主要通过无监督非对称度量学习方法结合 KL 散度来进一步对行人再识别方法进行优化, 通过实验进行比较来验证该方法的有效性, 结果表明在行人再识别应用中具有更优的识别精度。

关键词:无监督学习; 非对称; 度量学习; KL 散度; 行人再识别

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2022)09-0126-08

doi:10.3969/j.issn.1673-629X.2022.09.020

Unsupervised Asymmetric Metric Learning for Person Re-identification Optimization

JIANG Yu-yan¹, LYU Wei¹, LI Ping², SHAO Jin¹

(1. School of Management Science and Engineering, Anhui University of Technology, Ma'anshan 243032, China;

2. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Person re-recognition is a technology of similarity matching for different views of the same pedestrian. In the process of matching, different cameras, camera angles, pedestrian posture, illumination and obstacles are combined with metric learning to measure the similarity. With the change of external conditions, some information may be lost due to the neglect of specific view features, and the view may show different distribution. Combined with the unsupervised asymmetric metric learning (UAML), the view distribution is divided into shared view and specific view. Shared view extracts shared features, specific view extracts features associated with view and projects them into common subspace. Then K-means and Kullback-Leibler (KL) divergence are used to optimize the view distribution. The subspace samples obtained in this way have the same distribution and comprehensive representation. The method was tested on VIPeR data set, CUHK01 data set and Market-1501 data set, and the performance of the algorithm was measured by cumulative match characteristic (CMC), rank-1 precision and average precision curve. The accuracy of rank-1 is 40.25%, 56.6% / 58.09%, 61.67% respectively. The unsupervised asymmetric metric learning method combined with KL divergence is used to further optimize the pedestrian re-recognition method. The effectiveness of the method is verified through the comparison of experiments. The results show that the proposed method has better recognition accuracy in person re-recognition application.

Key words: unsupervised learning; asymmetry; metric learning; Kullback-Leibler; personal re-identification

0 引言

行人再识别^[1]在于行人从一个相机消失到另一

个相机再次出现的时候对其进行再识别。目的是为了判别同一个人出现在两个视图中的相似性。通过度量

收稿日期: 2021-10-11

修回日期: 2022-02-17

基金项目: 国家自然科学基金项目(62006126); 安徽普通高校重点实验室开放基金项目(CS2019-ZD02)

作者简介: 江雨燕(1966-), 女, 教授, 研究方向为机器学习、数据挖掘; 吕 魏(1996-), 男, 硕士研究生, 研究方向为机器学习、数据挖掘。

学习的方法来计算两个视图特征的距离来判断其相似性。在行人再识别算法上也有关于度量学习^[2]和深度学习^[3]被提出,基于无监督学习的科学技术研究方法通过伪标签的形式进行特征学习可以有效深度学习中改善全局特征不容易轻松识别的问题并能够产生良好的效果。

基于无监督学习的研究方法在面对同样的行人出现在不同摄像机里会因为外在因素,例如光线、姿势以及障碍物等造成两者的相似度降低^[4];同时两者也存在衣服、肤色等外在特征相似的地方。但是需要对每个摄像机视图构建映射来提高相似性,并且难以区分共享视图和特定视图,无法准确地建立摄像机视图之间的共同性质和特异性质。因此通过跨视图非对称的方法^[2,4-6]把视图特征区分为共享视图特征和特定于视图特征,引入共享映射来探索共享特征,提取特定的视图特征投影到公共子空间中使得两视图之间的差异尽可能缩小,然后通过无监督学习进行聚类^[2-3,7]。但是,在特定于视图特征的相似度区分上还需要进一步提高,通过 Bregman 散度^[8]方法来衡量特定于视图之间特征的差异,减少对共享视图特征区分的重复性,提高聚类效果,增强视图之间的相似度,从而提高视图之间相似的精确度。

对不断更新增长的数据进行标签是一项艰难而又复杂的工作, Martin Köstinger 等人^[9]提出了一种通过等价约束来指定标签,并通过马氏距离度量的可扩展性和需要的监督程度来进行学习,明显优于现有的结果。Husheng Dong 等人^[10]提出了一种新的基于重叠条纹的描述子与从密集块中提取的局部最大发生率 (LOMO) 相融合的增强局部最大发生率的方法,把精细的细节和表面的粗糙有效结合起来,充分利用其互补性,结合马氏距离和双线性相似性学习广义相似性。Giuseppe Lisanti 等人^[11]提出了基于正则化相关分析的跨视图匹配的学习技术,在不相交的视图中提取描述子空间的特征投影到公共子空间进行学习,在重新识别的技术上有明显的效果。Yachuang Feng 等人^[3]基于度量学习将行人的特征投影到公共子空间中,在公共子空间中将行人的特征提取出来分为共同特征和特定视图特征,并通过无监督跨视图的度量学习方法来计算相似性,通过最大均值差异 (MMD) 来计算不同样本的均值,常作为最后损失函数的优化,在迁移学习中使用频率相对较高且计算强度低,通常用来衡量两个分布之间的距离,属于一种核学习方法。为缩减样本之间的差异性,该文采用 Bregman 散度中的 KL 散度, KL 散度在视图中满足非对称特性,增加一个样本的协方差,这样能够在一定程度上提高样本之间的精确度。

1 相关工作

1.1 跨视图非对称度量学习

传统的度量学习是学习一个通用的变换矩阵,并将样本投影到一个公共子空间,一对样本 x_i 和 x_j 的距离为:

$$D(x_i, x_j) = \sqrt{(x_i - x_j)^T \mathbf{M} (x_i - x_j)} = \|\mathbf{W}x_i - \mathbf{W}x_j\|_2 \quad (1)$$

其中, $\mathbf{M} = \mathbf{W}^T \mathbf{W}$, \mathbf{W} 是投影矩阵, x_i 和 x_j 是相机视图中获取的不同样本。由于所有视图中的样本都在(1)中处理相同的 \mathbf{W} , 只从不同的视图中提取共享特征。根据跨视图非对称度量学习, 为了处理摄像机视图之间的不相似性, 其形式如下:

$$D(x_i, x_j) = \sqrt{(x_i - x_j)^T \mathbf{M} (x_i - x_j)} = \|\mathbf{W}^p x_i^p - \mathbf{W}^q x_j^q\|_2 \quad (2)$$

其中, \mathbf{W}^p 和 \mathbf{W}^q 分别是第 p 个和第 q 个视图的投影矩阵, x_i^p 和 x_j^q 分别是第 p 个和第 q 个视图获取的样本。虽然在一定程度上缓解了相机视图之间的不一致性, 但是在共享视图和特定视图的区分上还是难以取得很好的效果, 因此这里引入了共享映射和特定于视图的映射来对相机视图的共性和特异性进行建模, 通过对第 V 个视图的投影矩阵进行表示: 令 $\mathbf{W}^v = \mathbf{U}^0 + \mathbf{U}^v$, 则样本 x_i^p 和 x_j^q 的距离为:

$$d(x_i^p, x_j^q) = \|(\mathbf{U}^0 + \mathbf{U}^p)^T x_i^p - (\mathbf{U}^0 + \mathbf{U}^q)^T x_j^q\|_2 \quad (3)$$

其中, \mathbf{U}^0 表示共享视图投影, 用于提取共同特征, \mathbf{U}^v 表示特定于视图的投影, 把 \mathbf{U}^p 和 \mathbf{U}^q 投影到子空间来消除视图之间的特异性。并让 $\mathbf{U}^v x^v$ 与 $\mathbf{U}^0 x^v$ 正交, 让提取的视图专有特征与普通特征互补。

目标模型:

$$\min_{\mathbf{U}^0, \mathbf{U}^v} F_{\text{intra}} = \sum_{k=1}^K \sum_{l \in k} \|\mathbf{W}^{lT} x_i^l - c_k\|^2 \quad (4)$$

$$\text{s. t. } \mathbf{U}^{vT} \sum^v \mathbf{U}^v = \mathbf{I} \quad (V = 1, 2, \dots, v)$$

这里 $\mathbf{W}^v = \mathbf{U}^0 + \mathbf{U}^v$, 重写目标函数:

$$\min_{\mathbf{U}^0, \mathbf{U}^v} F_{\text{intra}} = \sum_{k=1}^K \sum_{l \in k} \|(\mathbf{U}^0 + \mathbf{U}^l)^T x_i^l - c_k\|^2 \quad (5)$$

$$\text{s. t. } \mathbf{U}^{vT} \sum^v \mathbf{U}^v = \mathbf{I} \quad (V = 1, 2, \dots, v)$$

其中, $\sum^v = \mathbf{X}^v \mathbf{X}^{vT} / N_v$, \mathbf{I} 表示单位矩阵。

共享信息为 \mathbf{U}^0 , 特定于视图信息为 \mathbf{U}^v 。 \mathbf{U}^v 被分为 \mathbf{U}^p 和 \mathbf{U}^q , 为了保证 \mathbf{U}^p 和 \mathbf{U}^q 学到不同的信息, 部分信息不被划分到共享信息里面, 这里保持 \mathbf{U}^v 和 \mathbf{U}^0 尽量正交, 又因为两者属于不同的样本, 直接正交没有意义, 故使得 $\mathbf{U}^v x^v$ 与 $\mathbf{U}^0 x^v$ 正交, 这里引入 Bregman divergence, 通过 Bregman divergence 来衡量 \mathbf{U}^p 和 \mathbf{U}^q 之间的差异。

1.2 Bregman 散度

Bregman 散度是损失函数或者失真函数,假设视图 p 是视图 q 的近似样本, p 是增加了障碍物或者光线的影响而形成的 q , 所以 Bregman 散度就是用来衡量 p 和 q 的之间的差异性。定义 F 是在凸集 Ω 上可微的严格凸函数, 在函数 F 生成的 Bregman 散度的形式是^[7]:

$$D_F(p \| q) = F(p) - [F(q) + \langle \nabla F(q), p - q \rangle] \quad (6)$$

其中, $D_F(p, q)$ 表示样本 p 与 q 之间的距离, 用来衡量 p 与 q 之间的差异, $\nabla F(q)$ 表示在 q 的梯度, $\langle \nabla F(q), p - q \rangle$ 表示 $\nabla F(q)$ 与 $p - q$ 的内积。 $L(p, q) = F(q) + \langle \nabla F(q), p - q \rangle$ 表示函数 F 在 q 处附近的线性部分, Bregman 散度就是一个函数与另一个函数线性部分之间的差^[7], 选择不同的函数 F , 可以得到不同的 Bregman 散度。

假设来自不同视图的样本是从类似场景中捕获, 因此它们的概率分布被视为相同。但在原始特征空间中的分布通常是不同的, 因为它们是从不同的相机、不同的场景中捕获的, 直接计算它们的距离可能是不准确的。为此, 尽可能地提高它们在变换后的子空间中的概率分布的相似性, 选择不同的函数 F , 采用 KL 散度来衡量样本分布之间的差异性。KL 散度能够有效解决样本之间非对称的问题, 并且在衡量样本之间差异性的时候, 能够有效衡量样本的近似分布与真实分布之间的匹配程度, 主要通过计算两个样本之间的高斯分布来表示 KL 散度。这里计算 U^p 和 U^q 之间的距离形式为:

$$D_F(U^p \| U^q) = F(U^p) - [F(U^q) + \langle \nabla F(U^q), U^p - U^q \rangle] \quad (7)$$

2 模型与优化

KL 散度的形式为 $\text{KL}(p(h^p) \| p(h^q))$, 其中 $p(h^v) \sim N(\mu^v, H^v)$, $p(U^v)$ 是服从均值为 μ^v , 协方差为 H^v 的多元高斯分布, U^p 和 U^q 的多元高斯分布 $p(h^v)$ 表示为:

$$p(h^v | \mu^v, H^v) = \frac{1}{(2\pi)^{n/2} |H^v|^{1/2}} \exp\left[-\frac{1}{2} (h^v - \mu^v)^T (H^v)^{-1} (h^v - \mu^v)\right]$$

其中, 均值 μ^v 的形式为: $\mu^v = \frac{1}{N_v} \sum U^v x^v$, 协方差 H^v 的

形式为 $\text{var}(H^v) = \frac{\sum (U^v x^v - \mu^v) (U^v x^v - \mu^v)^T}{N_v - 1}$ 所以,

只要对下面的形式展开优化:

$$\min O(U^0, \dots, U^v) =$$

$$\sum_{k=1}^K \sum_{I \in k} \| (U^0 + U^v)^T x_i^v - c_k \|^2 + \lambda \text{KL}(p(h^p) \| p(h^q)) + \beta \sum_{k=1}^K \| U^v \|^2$$

s. t. $\langle U^{0T} x_i^v, U^{vT} x_j^v \rangle = 0,$

$$(U^{vT} \sum^v U^v = I), V = 1, 2, \dots, v \quad (8)$$

其中, $\sum_{v=1}^v \| U^v \|^2$ 和 $\text{KL}(p(h^p) \| p(h^q))$ 表示正则化项, 用来限制特定视图的信息使得共享信息突出。其中 $C = [c_1, c_2, \dots, c_k]$ 表示聚类中心的个数, λ 和 β 是正则化参数, 由于共享视图占的比例要比特定视图高, 所以 $\Sigma^v U^v = X^v X^{vT} / N_v$ 。

为了方便, 令 $\tilde{W} = [W^{1T}, W^{2T}, \dots, W^{vT}] \in v \times 1, \tilde{U} = [U^{0T}, U^{1T}, \dots, U^{vT}] \in (v+1) \times 1$ 使得 $\tilde{W} = \tilde{I}_1 \tilde{U}$, 这里 \tilde{I}_1 是由单位矩阵 I 组成的:

$$\tilde{I}_1 = \begin{bmatrix} I & I & 0 & \dots & 0 \\ I & 0 & I & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I & 0 & 0 & \dots & I \end{bmatrix}, \text{ 因此, 这里可以令}$$

$$\frac{1}{N} \sum_{k=1}^K \sum_{v=1}^v \sum_{x_i^v} \| W^{vT} x_i^v - c_k \|^2 = \| \tilde{W} \tilde{X} - CD \|^2. \text{ 这}$$

里 \tilde{X} 和 D 分别表示为:

$$\tilde{X} = \begin{bmatrix} x_1^1 & \dots & x_{N_1}^1 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & x_1^v & \dots & x_{N_v}^v \end{bmatrix}$$

$$D_{ki} = \begin{cases} 1, x_i^v \in C_k \\ 0, x_i^v \notin C_k \end{cases}$$

算法: 跨视图非对称度量学习的行人重识别

输入: $\tilde{X}, K, \lambda, \beta$

输出: \tilde{W}

1. 通过 K 均值使用原始训练数据 X 来初始化 D
2. 固定 D 优化 \tilde{U} 式(13)使用梯度下降求解
3. 重复:
4. 固定 \tilde{U} 优化 D : 通过 K 均值方法初始化类别中心
5. 固定 D 更新 \tilde{U} : 求解等式中的广义特征值
6. 直到收敛或迭代结束
7. 返回: $\tilde{W} = \tilde{I}_1 \tilde{U}$

聚类中心是每一个聚类样本的平均值, 即 $C = \tilde{W}^T \tilde{X} \tilde{Q}^T$, 其中 \tilde{Q}^T 的计算过程为 $Q_{ki} = \frac{D_{ki}}{n_k}$, 并且 n_k 等于 C_k 的大小, 故得到:

$$\frac{1}{N} \sum_{k=1}^K \sum_{v=1}^V \sum_{x_i^v} \| W^{vT} x_i^v - c_k \|^2 =$$

$$\text{tr}(\tilde{W}^T \tilde{X} \tilde{X}^T \tilde{W} - \tilde{W}^T \tilde{X} D^T Q \tilde{X}^T \tilde{W}) \quad (9)$$

对两个高斯分布 $p(h^p) \sim N(\mu^p, H^p)$, $p(h^q) \sim N(\mu^q, H^q)$ 计算 KL 散度 $\text{KL}(p(h^p) \| p(h^q))$ 为:

$$\text{KL}(p(h^p) \| p(h^q)) = E_{p^i}(\ln p(h^p) - \ln p(h^q)) \text{ 即:}$$

$$\text{KL}(p(h^p) \| p(h^q)) = \frac{1}{2} \ln | H^q H^{p^{-1}} | - \frac{1}{2} n +$$

$$\frac{1}{2} \text{tr}(H^{q^{-1}} H^p) + \frac{1}{2} (\mu^q - \mu^p)^T H^{q^{-1}} (\mu^q - \mu^p) \quad (10)$$

Bregman 散度中的 KL 散度计算样本均值,还计算了样本的协方差,因此与最大均值误差(MMD)方法进行比较,这两种方法均可以对不同的分布进行判断,并且满足不同分布的距离不对称性。

对于 $\beta \sum_{v=1}^V \| U^v \|^2$, 这里令其等于 $\beta \| \tilde{I}_2 U^v \|^2$, 其

$$\text{中 } \tilde{I}_2 = \begin{bmatrix} 0 & \mathbf{I} & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{I} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

放宽约束条件:

$$\sum_v W^{vT} \Sigma^v W^v = \tilde{W}^T \tilde{\Sigma} \tilde{W} = \tilde{U}^T \tilde{I}_1 \tilde{\Sigma} \tilde{I}_1 \tilde{U} = \mathbf{VI} \quad (11)$$

$$\sum_v \frac{1}{N_v} \sum_i \langle U^{0T} x_i^v, U^{vT} x_i^v \rangle = \text{tr}[\tilde{U}^T (\tilde{I}_1 - \tilde{I}_2)^T \tilde{\Sigma} \tilde{I}_2 \tilde{U}] = 0 \quad (12)$$

其中, $\tilde{\Sigma}$ 是对角矩阵。所以 $\min O(U^0, \dots, U^V)$ 优化最终可以写成如下形式:

$$\min O(U^0, \dots, U^V) =$$

$$\frac{1}{N} \text{tr}(\tilde{U}^T \tilde{I}_1 \tilde{X} (\mathbf{I} - D^T Q) \tilde{X}^T \tilde{I}_1 \tilde{U}) +$$

$$\frac{\lambda}{2} \ln | H^q H^{p^{-1}} | - \frac{1}{2} n + \frac{1}{2} \text{tr}(H^{q^{-1}} H^p) +$$

$$\frac{1}{2} (\mu^q - \mu^p)^T H^{q^{-1}} (\mu^q - \mu^p) + \beta \| \tilde{I}_2 U^v \|^2 \quad (13)$$

目标函数是非凸的和非线性的,因此对 \tilde{U} 和 D 进行优化:

固定 \tilde{U} 和 D 进行优化:指标矩阵 D 和 C 有关,通过变换后的训练数据集 $\tilde{W}^T \tilde{X} = (\tilde{I}_1 \tilde{U})^T \tilde{X}$ 。

对 C 进行 k-means 聚类优化,然后再计算 D 。

固定 D 对 \tilde{U} 进行优化:这里通过拉式数乘法对 \tilde{U} 进行求导,其中 $\tilde{U} = \tilde{I} U$, Ψ 、 φ 是拉式乘数,对一阶求导结果采取梯度下降法进行处理。构造拉式如下:

$$F = \frac{1}{N} \text{tr}(\tilde{U}^T \tilde{I}_1 \tilde{X} (\mathbf{I} - D^T Q) \tilde{X}^T \tilde{I}_1 \tilde{U}) +$$

$$\frac{\lambda}{2} [\ln | \frac{H^q}{H^p} | - n + \text{tr}(H^{q^{-1}} H^p) +$$

$$(\mu^q - \mu^p)^T H^{q^{-1}} (\mu^q - \mu^p)] + \beta \| \tilde{I}_2 U^v \|^2 +$$

$$\Psi(\tilde{U}^T \tilde{I}_1 \tilde{\Sigma} \tilde{I}_1 \tilde{U} - \mathbf{VI}) +$$

$$\varphi \text{tr}[\tilde{U}^T (\tilde{I}_1 - \tilde{I}_2)^T \tilde{\Sigma} \tilde{I}_2 \tilde{U}] \quad (14)$$

分别对 \tilde{U} 和乘子 Ψ 、 φ 进行求偏导并令其为零得到式(15)、(17)、(18):

$$\frac{\partial F}{\partial \tilde{U}} = \frac{1}{N} (\text{tr}(\tilde{I}_1 \tilde{X} (\mathbf{I} - D^T Q) \tilde{X}^T \tilde{I}_1 \tilde{U}) + \tilde{U}^T \tilde{I}_1 \tilde{X} (\mathbf{I} -$$

$$D^T Q) \tilde{X}^T \tilde{I}_1) + \frac{\lambda}{2} K \tilde{I} + 2\beta \tilde{I}_2 \tilde{U} + \Psi(\tilde{I}_1 \tilde{\Sigma} \tilde{U})$$

$$\varphi [(\tilde{I}_1 - \tilde{I}_2)^T \tilde{\Sigma} \tilde{I}_2 \tilde{U} +$$

$$\tilde{U}^T (\tilde{I}_1 - \tilde{I}_2)^T \tilde{\Sigma} \tilde{I}_2] \tilde{I} = 0 \quad (15)$$

其中 K 的形式为:

$$K = \text{tr}(\frac{H^p H^{q^{-1}}}{N_{q^{-1}}} A H^{p^{-1}} - \frac{1}{N_{p^{-1}}} B H^{p^{-1}}) \tilde{I} +$$

$$H^{q^{-1}} (\frac{1}{N_{p^{-1}}} B - \frac{1}{N_{q^{-1}}} A H^{q^{-1}} H^p) \tilde{I} + L \tilde{I} \quad (16)$$

这里 A 、 B 、 L 分别表示如下:

$$A = [\sum_{j=1}^{N_q} x_j^q x_j^{qT} U^{qT} + U^q \sum_{j=1}^{N_q} x_j^q x_j^{qT} -$$

$$\frac{1}{N_q} U^q \sum_{j=1}^{N_q} x_j^q \sum_{j=1}^{N_q} x_j^q - \frac{1}{N_q} \sum_{j=1}^{N_q} x_j^q \sum_{j=1}^{N_q} x_j^q U^{qT} -$$

$$(1 - \frac{1}{N_q}) \sum_{j=1}^{N_q} x_j^q \sum_{j=1}^{N_q} x_j^{qT} \mu^{qT} -$$

$$(1 - \frac{1}{N_q}) \mu^q \sum_{j=1}^{N_q} x_j^{qT}] \quad (17)$$

$$B = [\sum_{i=1}^{N_p} x_i^p x_i^{pT} U^{pT} + U^p \sum_{i=1}^{N_p} x_i^p x_i^{pT} -$$

$$\frac{1}{N_p} U^p \sum_{i=1}^{N_p} x_i^p \sum_{i=1}^{N_p} x_i^p - \frac{1}{N_p} \sum_{i=1}^{N_p} x_i^p \sum_{i=1}^{N_p} x_i^p U^{pT} -$$

$$(1 - \frac{1}{N_p}) \sum_{i=1}^{N_p} x_i^p \sum_{i=1}^{N_p} x_i^{pT} \mu^{pT} -$$

$$(1 - \frac{1}{N_p}) \mu^p \sum_{i=1}^{N_p} x_i^{pT}] \quad (18)$$

$$L = (\mu^{qT} + \mu^{pT}) H^{q^{-1}} A H^{q^{-1}} \mu^p + (\mu^{pT} -$$

$$\mu^{qT}) H^{q^{-1}} A H^{q^{-1}} \mu^q - \frac{1}{N_q} \sum_{j=1}^{N_q} x_j^{qT} H^{q^{-1}} \mu^q -$$

$$\frac{1}{N_p} \sum_{i=1}^{N_p} x_i^{pT} H^{q^{-1}} \mu^p - \mu^{pT} H^{q^{-1}} \frac{1}{N_p} \sum_{i=1}^{N_p} x_i^p -$$

$$\mu^{qT} H^{q^{-1}} \frac{1}{N_q} \sum_{j=1}^{N_q} x_j^q \quad (19)$$

$$\frac{\partial F}{\partial \Psi} = \tilde{U}^T \tilde{I}_1 \tilde{\Sigma} \tilde{I}_1 \tilde{U} - \mathbf{VI} = 0 \quad (20)$$

$$\frac{\partial F}{\partial \varphi} = \text{tr}[\tilde{U}^T (\tilde{I}_1 - \tilde{I}_2)^T \tilde{\Sigma} \tilde{I}_2 \tilde{U}] = 0 \quad (21)$$

取所有等式右侧函数的平方和为目标函数,计算这个函数关于自变量 U^v 以和乘子 Ψ 、 φ 偏导数,再对关于自变量及乘子的偏导数进行迭代求解。

3 实验

3.1 数据集介绍

通过数据样本来训练该方法的有效性,使用 Cumulative Match Characteristic (CMC) Curve 累计匹配曲线能够综合反映分类器的性能,Rank-1 精度通过百分比的形式,计算每百张的平均精确度对文中方法和其他方法进行比较。

VIPeR^[12]数据集(见图 1)通过随机将其分为两半,316 张图像用于训练,316 张图像用于测试。重复这个过程十次,并报告每个算法的平均性能进行比较。



图 1 VIPeR 数据集部分行人样本

CUHK01^[13](见图 2)是通过两个不重叠的摄像机捕捉的 971 个行人身份视图,每视图有 4 个图像,共有 3 884 个图像。训练集随机选择 485 人,测试集为其余 486 对行人。通过十次重复来评估并同时单镜头(SS)和多镜头(MS)实验。

Market-1501^[14](见图 3)与其他的数据集不同,只能在多镜头上进行实验。选择 751 人作为训练集,共 12 936 张图像,其余 750 人用于测试。

3.2 参数设置

该方法包含 $\text{KL}(p(h^p) \| p(h^q))$ 和 $\sum_{v=1}^V \|U^v\|^2$ 两个正则化项。 $\text{KL}(p(h^p) \| p(h^q))$ 减少相机视图之间的分布差异。 $\sum_{v=1}^V \|U^v\|^2$ 用于限制特定于视图映射的影响,并防止过拟合。将限制分布差异的 λ 设置为 0.5,正则化参数 β 设置为 0.01。对于聚类分析得出的

$K=800$ 适用于大规模数据集。



图 2 CUHK01 数据集部分行人样本



图 3 Market-1501 数据集部分行人样本

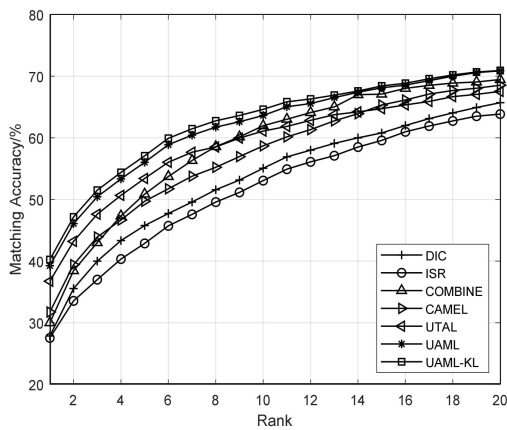
3.3 结果分析

这里主要是与稀疏字典学习模型(简称 Dic)^[15]、稀疏表示学习模型(ISR)^[16]、视跨图非对称度量学习(CAMEL)^[3]、组合度量(COMBINE)^[17]、无监督轨迹(UTAL)^[18]、无监督的跨视图度量学习(UAML)^[2]进行比较,如表 1 所示,并画出 CMC 曲线,如图 4 所示。

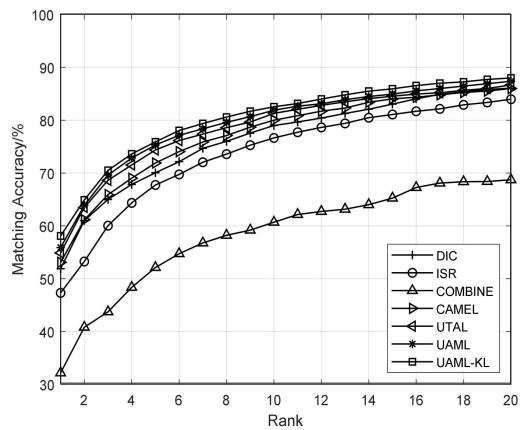
UAML - KL 通过学习每个摄像机视图的投影矩阵,不仅把投影矩阵划分成了共享投影和特定于视图的投影去学习特征,在计算样本的分布差异过程中引入了 KL 散度,通过计算两者的多元高斯分布来表示分布的 KL 散度,进一步去加强再识别的精度值,探索共同的特性并降低视图之间的不一致之外,还更加细化每个摄像机视图的特定视图的特征。

表1 用秩-1精度(%)和MAP(%)测量三个数据集的无监督方法的比较结果

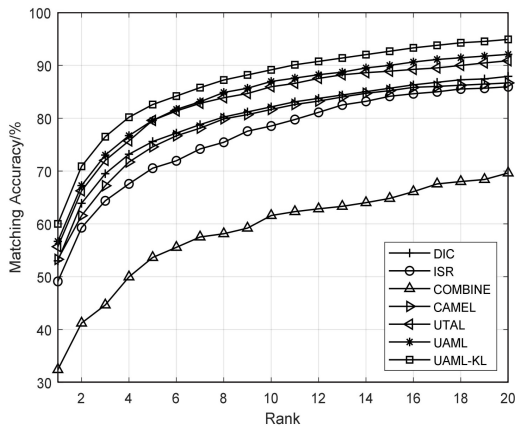
数据集\方法	VIPeR	CUHK01 (SS/MS)	Market-1501
DiC ^[15]	27.86	51.36/53.62	50.68
ISR ^[16]	27.54	47.84/51.33	38.19
COMBINE ^[17]	30.00	32.17/32.42	35.42
CAMEL ^[3]	31.96	52.51/54.37	57.61
UTAL ^[18]	35.73	53.89/55.69	61.94
UAML ^[2]	39.26	53.00/55.89	59.07
UAML-KL	40.25	56.6/58.09	62.67



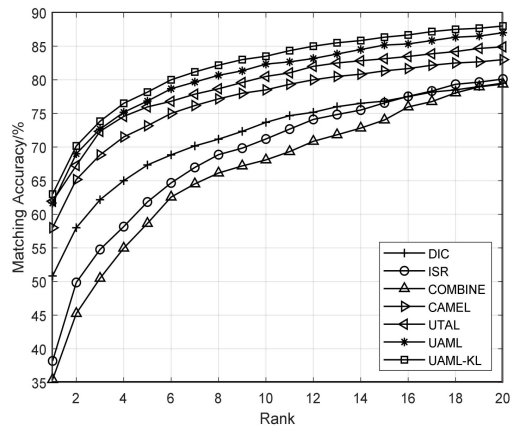
(a) VIPeR



(b) CUHK01 (SS)



(c) CUHK01 (MS)



(d) Market-1501

图4 三个数据集上的 CMC 曲线

3.3.1 共享映射和特定视图映射分析

共享映射和特定视图映射的作用主要是对摄像机视图之间的共性与不一致性进行建模来获得比较全面

的、具有区分性的行人特征。为了使得该方法具有一定的性能,通过控制共享特征和特定于视图的特征,只保留 U^0 或 U^V 来进行对比验证,如表2所示。

表2 非对称度量学习验证

数据集\方法	VIPeR	CUHK01 (SS/MS)	Market-1501
UAML(U^0)	31.6	47/50.1	55.8
UAML(U^V)	30.7	45.5/48.1	46.5
UAML- KL(U^0)	33.21	52.37/55.24	60.33
UAML- KL(U^V)	32.5	51.00/55.16	59.14
UAML - KL	40.25	56.6/58.09	62.67

KL 散度的非对称性质能够很好地与非对称度量学习相契合,通过均值和协方差,在衡量样本分布的时候,能够与真实样本分布进行匹配,匹配程度越高,则样本分布愈接近真实的样本分布,进一步改善相机视图之间的分布差异。除了特定于视图的映射可以为共享的特性提供互补的信息之外,还构建了全面的和有区别的表示。

表 3 不同聚类中心数量结果

K	200	400	600	800	1 000	1 200	1 400	1 800
VIPeR	39.8	39.93	39.98	41.25	40.03	40.01	40.00	40.00
CUHK01	51.10	52.23	53.25	56.60	55.22	53.16	50.41	49.11
Market	61.30	61.14	61.53	62.67	61.27	61.23	61.14	61.13

3.3.3 特征分析

为了适应不同特征的能力,为证明不仅在采用基于深度学习的 JSTL 特性时也很有效,还采用了 LOMO 特征,先进行 PCA 进行降维,得到 512 维

3.3.2 聚类分析

在三个数据集上验证 K 对行人再识别性能的影响实验。结果如表 3 所示, K 在 200 到 1 600 之间变化,时间间隔为 200。CUHK01 的性能变化较大,这是因为 CUHK01 与 Market 相对较小的集群。过多或过少的集群可能会阻碍对行人的样本之间关系的准确探索。

LOMO 特征,如表 4 所示。在所有模型中, Dic 和 ISR 的结果最具可比性(Dic 和 ISR 位居第二)。所以为了清晰起见,只把该方法和其他特征提取方法比较, LOMO 特征作为基线。

表 4 使用 LOMO 特征的比较结果

数据集\方法	VIPeR	CUHK01 (SS/MS)	Market-1501
Dic ^[15]	17.10	16.78/17.51	28.54
ISR ^[16]	20.54	15.96/18.32	18.10
COMBINE ^[17]	21.33	19.27/20.61	18.97
UTAL ^[18]	24.45	24.62/26.33	25.41
LOMO ^[19]	12.21	16.57/17.3	25.35
CAMEL ^[3]	22.62	22.36/24.22	30.22
UAML ^[2]	23.50	25.00/28.23	30.27
UAML - KL	26.52	27.88/30.78	32.66

4 结束语

无监督的行人再识别度量学习方法的基本思想是每个摄像机视图中的行人样本分别是两个分布中提取的:一个提取的是摄像机视图之间的共同特征,另一个提取的是特定于视图的特征。引入一个共享映射来探索共享特征,并构造特定于视图的映射与视图相关的特征提取投影到一个公共子空间中。这样不仅降低相机视图之间的一致性,还更加细化了特定视图的特征,能够更加精确地对行人进行再识别。此外,为了实现精确的相似性测量,还减少变换空间中摄像机视图之间的分布差异,在投影空间中对样本进行聚类,通过无监督的方式进行优化。实验结果表明,与其他算法相比,该方法具有较好的性能。

参考文献:

[1] 孙 锐,夏苗苗,陆伟明,等. 联合增强局部最大发生特征和 k -KISSME 度量学习的行人再识别[J]. 图学学报, 2020, 41(3): 362-371.

- [2] FENG Y, YUAN Y, LU X. Person re-identification via unsupervised cross-view metric learning [J]. IEEE Transactions on Cybernetics, 2019, 51(4): 1849-1859.
- [3] YU H, WU A, ZHENG W. Cross-view asymmetric metric learning for unsupervised person re-identification[C]//2017 IEEE international conference on computer vision (ICCV). Venice: IEEE, 2017: 994-1002.
- [4] CHEN Y, ZHENG W, LAI J, et al. An asymmetric distance model for cross-view feature mapping in person reidentification [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(8): 1661-1675.
- [5] CHEN Y, ZHU X, ZHENG W, et al. Person re-identification by camera correlation aware feature augmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(2): 392-408.
- [6] GAO X, TAO D, WANG X, et al. Transfer latent variable model based on divergence analysis [J]. Pattern Recognition, 2011, 44(10-11): 2358-2366.
- [7] 张宝华,朱思雨,吕晓琪,等. 软多标签和深度特征融合的无监督行人重识别[J]. 光电工程, 2020, 47(12): 13-22.
- [8] DAVIS J V, KULIS B, JAIN P, et al. Information-theoretic

- metric learning [C]//Proceedings of the 24th international conference on machine learning. Corvallis; Association for Computing Machinery, 2007; 209–216.
- [9] 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; IEEE, 2012; 2288–2295.
- [10] DONG Husheng, LU Ping, ZHONG Shan, et al. Person re-identification by enhanced local maximal occurrence representation and generalized similarity metric learning [J]. Neurocomputing, 2018, 307; 25–37.
- [11] LISANTI G, MASI I, BIMBO A D. Matching people across camera views using kernel canonical correlation analysis [C]//Proceedings of the international conference on distributed smart cameras. New York; Association for Computing Machinery, 2014; 1–6.
- [12] GRAY D. Evaluating appearance models for recognition, re-acquisition, and tracking [C]//IEEE international workshop on performance evaluation for tracking and surveillance. Rio de Janeiro; IEEE, 2007; 1–7.
- [13] LI WEI, ZHAO RUI, WANG XIAOGANG. Human re-identification with transferred metric learning [C]//Computer vision – ACCV 2012. Berlin; Springer, 2013; 31–44.
- [14] ZHENG L, SHEN L, TIAN L, et al. Scalable person re-identification; a benchmark [C]//2015 IEEE international conference on computer vision (ICCV). Santiago; IEEE, 2015; 1116–1124.
- [15] KODIROV E, XIANG T, GONG S. Dictionary learning with iterative Laplacian regularisation for unsupervised person re-identification [C]//Proceedings of the British machine vision conference (BMVC). Britain; The British Machine Vision Association (BMVA) Press, 2015; 1–12.
- [16] 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); 1629–1642.
- [17] ZHAO H, HAN Z, LI Z, et al. Unsupervised person re-identification via re-ranking enhanced sample-specific metric learning [C]//2017 IEEE international conference on image processing (ICIP). Beijing; IEEE, 2017; 3405–3409.
- [18] LI M, ZHU X, GONG S. Unsupervised tracklet person re-identification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(7); 1770–1782.
- [19] LIAO S, HU Y, ZHU X, et al. Person re-identification by local maximal occurrence representation and metric learning [C]//2015 IEEE conference on computer vision and pattern recognition (CVPR). Boston; IEEE, 2015; 2197–2206.
-
- (上接第 125 页)
- [11] 刘冬, 靳蓓蓓, 阙向红. 基于一种遗传算法的最小测试用例集自动生成 [J]. 计算机技术与发展, 2016, 26(4); 86–89.
- [12] 刘渊, 杨永辉, 张春瑞, 等. 一种基于遗传算法的 Fuzzing 测试用例生成新方法 [J]. 电子学报, 2017, 45(3); 552–556.
- [13] 包晓安, 熊子健, 张唯, 等. 一种基于改进遗传算法的路径测试用例生成方法 [J]. 计算机科学, 2018, 45(8); 174–178.
- [14] 桂预风, 苏鹏. 基于改进遗传算法的自动阈值图像分割方法 [J]. 数学的实践与认识, 2015, 45(19); 193–197.
- [15] 张峰峰, 张欣, 陈龙, 等. 采用改进遗传算法优化神经网络的双目相机标定 [J]. 中国机械工程, 2021, 32(12); 1423–1431.
- [16] 谷晓琳. 基于改进遗传算法的柔性作业车间调度问题的应用研究 [D]. 大连; 大连交通大学, 2020.
- [17] 罗雄, 钱谦, 伏云发. 遗传算法解柔性作业车间调度问题应用综述 [J]. 计算机工程与应用, 2019, 55(23); 15–21.
- [18] 孙晓华, 邵珊, 孙笑竹. 演化经济学研究综述 [J]. 中国地质大学学报: 社会科学版, 2010, 10(5); 119–124.
- [19] 曹道友, 程家兴. 基于改进的选择算子和交叉算子的遗传算法 [J]. 计算机技术与发展, 2010, 20(2); 44–47.
- [20] 谢燕丽, 许青林, 姜文超. 一种基于交叉和变异算子改进的遗传算法研究 [J]. 计算机技术与发展, 2014, 24(4); 80–83.