



计算机工程与应用
Computer Engineering and Applications
ISSN 1002-8331, CN 11-2127/TP

《计算机工程与应用》网络首发论文

题目: 迁移度量学习行人再识别算法
作者: 宋丽丽
网络首发日期: 2018-12-14
引用格式: 宋丽丽. 迁移度量学习行人再识别算法[J/OL]. 计算机工程与应用.
<http://kns.cnki.net/kcms/detail/11.2127.TP.20181212.1740.033.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

迁移度量学习行人再识别算法

宋丽丽

SONG Lili

成都理工大学 工程技术学院, 四川 乐山 614000

The Engineering & Technical College of Chengdu University of Technology, Leshan, Sichuan 614000, China

SONG Lili. Transfer Metric Learning for Person Re-identification. Computer Engineering and Applications

Abstract : Pedestrian re-recognition is a challenging task in the field of computer vision. This task focuses on the appearance change pattern of individuals. Due to the drastic variation of appearance feature, there is small sample problem in metric learning for person re-identification. In this paper, a transfer metric learning based method is proposed. By minimizing the difference between the distribution of source data and target data, the proposed method achieves the transform of metric model from source dataset to target dataset. The proposed method not only enhances the diversity of training samples which improves the discrimination of metric model, but also improves its generalization. Finally, the effectiveness and accuracy of the proposed method are verified on the VIPeR and CUHK01 datasets by the pre-training on iLIDS dataset.

Key words: person re-identification; metric learning; transfer learning

摘 要: 行人再识别技术是计算机视觉领域中一个具有挑战性的任务。该任务针对个体的外观变化模式展开研究, 特征变化剧烈, 存在小样本问题, 而通过提出的一种基于迁移学习的度量学习模型, 可约束不同数据集样本分布的差异, 实现度量模型在不同数据集上的迁移。该算法不仅增强了度量模型训练样本的多样性, 提高了分辨能力, 同时提升了样本的适应性。最后, 通过在 iLIDS 数据集进行度量模型的预训练, 并在 VIPeR 和 CUHK01 两个数据集上进行的迁移学习, 验证了算法的有效性和准确性。

关键词: 行人再识别; 度量学习; 迁移学习

文献标志码: A 中图分类号: TP391.41 doi: 10.3778/j.issn.1002-8331.1808-0145

1 引言

行人再识别系统的研究目的是对整个监控网络中特定行人目标进行图像检测与识别^[1-3]。该技术有助于在整个监控网络中完善行人的运动轨迹、定位目标的位置, 并在行为分析和公安系统嫌疑人追踪等方面有着重要的应用价值。行人再识别问题首先将目标行人的图像和查找范围内所有检测得到的行人图像, 通过特征表达模型转化为特征向量, 然后通过计算机视觉的相关算法对样本进行相似性度量, 得到与目标人物距离最接近的图像, 从而

完成对目标的定位。从行人再识别任务的内容可以看出, 该技术主要包含两部分内容, 即为如何将图像转化为特征描述和相似性度量模型。

在特征表达方面, 研究工作者主要致力于如何设计一种主观的特征表达模型, 提取图像的色彩和纹理特征, 使得特征不仅具有强健的鲁棒性, 同时具有强健的分辨能力。早期的行人再识别算法研究也主要基于已有的度量模型, 进行特征表达模型的设计。如图 1 中所示给出了行人再识别的样本图像, 这些图像由监控图像中检测得到, 图像分辨率较低, 使得传统的基于生物特征的人脸识别、姿态识别等

技术无法应用于行人再识别任务。因此, 图像的色彩和纹理特征成为图像特征表达的主要内容。传统的特征描述主要是基于色彩和纹理特征, 比如色彩统计直方图、LBP、Gabor 算子等^[4-6]。



图 1 行人再识别样本

此外, 行人图像在跨域场景时, 受到光照变化、步态、拍摄角度和背景等因素的干扰, 尤其是行人身体结构的对不准问题, 导致图像的外观特征变化剧烈, 行人再识别的精度处于较低的水平。如图 1 中给出了正常的正样本对和复杂的正样本对。虽然两种样本对都是正样本对, 但是 1.1(b) 中的样本对存在复杂的光照变化。因此, 人们希望建立一种对光照、步态、拍摄角度等内容鲁棒的特征模型, 同时模型需要具有强健的表达能力。Yang^[7]等人基于 color name 的外观描述方法, 提出了一种 SCNCD 特征表述方法, 通过 color names 的色彩表达方式使得特征表达对光照的变化更加鲁棒。廖胜才^[8]等人针对行人身体结构的对不准问题提出了 LOMO 特征, 该模型在局部图像内提取图像的色彩直方图和颜色直方图, 通过灰度级压缩和多尺度模型, 设计得到一种具有强分辨能力和鲁棒性的高维特征向量。袁立^[9]等人设计了一种特征融合的特征表达模型, 该模型融合了 HSV 空间下色彩直方图以及梯度直方图等特征。针对步态变化, Ziming Zhang^[10]等人提出了一种 codebook 的外观特征模型, 通过视觉词汇对图像的特征进行表达。

虽然, 上述的特征模型显著的提升了特征的分辨能力和鲁棒性, 但是由于行人图像特征变化过于剧烈, 导致行人再识别的精度仍处于较低的水平。基于度量学习的行人再识别方法是该领域中研究最广泛也是效果最好的方法, 因此研究工作者提出

了大量的新算法对行人再识别进行研究, 例如 ITML^[11]、RDC^[12]、PRDC^[13]、KISSME^[14]、XQDA^[8]和 NFST^[15]等算法。RDC 是由郑伟诗^[12,13]等人提出的算法, 该算法通过定义每个对象正负样本对的相对距离, 建立相对距离优化准则来学习度量投影空间。PRDC 则是 RDC 算法的概率模型。但是基于相对距离的方法, 计算复杂, 计算量大, 并且需要大量的训练数据学习投影矩阵。Köstinger^[14]等人基于高斯分布假设, 建立了行人图像样本对的概率推断模型 KISSME, 然而行人图像特征变化过于复杂, 测试样本与训练样本分布存在较大差异, 导致训练得到的模型适应性较差。廖胜才^[8]等人针对行人再识别问题, 基于 FDA 算法思想, 结合 KISSME 算法建立度量空间学习算法, 该算法学习了一个度量子空间, 并提出了 XQDA 算法。Zhang^[15]等人则基于 Foley-sammon 变换(FST)对行人样本进行度量, 并通过核化处理得到了更好的分辨能力。同样的, 由于行人样本的剧烈变化以及行人再识别的小样本问题, XQDA 算法和 NFST 算法通过严格约束正样本对散度与最大化负样本总体散度来学习度量空间, 使得度量模型在训练样本上过拟合, 导致识别精度较低。Li^[16]等人提出了 FPNN 深度模型对行人图像进行识别, 并达到了较好的精度。Ahmed^[17]等人对 FPNN 深度网络模型进行了改进, 引入一个对称的特征提取层, 来实现更加强健的特征表达。由于行人图像跨越场景的对不准问题, 卷积神经网络的特征提取方法无法获取稳定的特征, 并且深度学习网络需要大量训练样本。此外, 行人再识别存在小样本问题, 使得深度学习算法在行人再识别问题的研究中效果较差。

由于行人再识别样本特征变化的复杂性, 获取包含所有行人外观变化的样本总体分布是不可能完成的任务。现有的行人再识别数据库仅仅是样本总体中很小的一部分, 甚至在训练样本与测试样本的分布之间存在非常大的差异, 只有少量的测试样本可以在训练样本中找到与之近似的个体。这导致

了行人再识别任务中的小样本问题，现有的行人再识别度量模型通常在训练样本上过拟合。因此，如何利用现有数据库，提升模型的适应性是本文的研究重点。针对上述问题，本文中提出了一种迁移度量学习算法，实现交叉数据库的行人再识别。本文算法学习一个具有判别性的度量投影空间，并通过约束源域与目标域样本分布的相似性，将判别信息从有标签的源域迁移到无标签的目标域，且在学习到的度量空间中，约束正样本对的散度，同时最大化负样本对的散度。最后，通过 VIPeR 和 CUHK01 两个数据库上的仿真试验，验证了算法的有效性和精度。

2 行人再识度量模型

行人再识别任务的研究目的是对监控网络不同摄像头中，捕捉到的行人目标进行相似性度量，从而实现对特定行人目标的定位和跟踪。因此，行人再识别任务的关键是在于建立有效的距离或者相似性度量函数。

给定数据集 $X = \{x_i^p, x_j^s\}$, $x_i^p, x_j^s \in \mathbb{R}^{2 \times d \times N}$, x_i^p, x_j^s 分别表示在两个不同的摄像头下捕捉到的行人图像。令 l_{ij} 表示样本对 (x_i^p, x_j^s) 中两个样本之间关系的标签信息。若 $l_{ij}=1$ ，则 (x_i^p, x_j^s) 表示正样本对图像，即两个图像对应同一个行人；若 $l_{ij}=0$ ，则 (x_i^p, x_j^s) 表示负样本对图像，即两个图像对应不同的行人。行人再识别任务的本质是对于摄像头 A 场景中的每个待识别的 x_i^p 目标，与摄像头 B 场景中捕捉到的所有行人样本集 $\{x_j^s\}$ 中进行一一匹配，计算每对图像之间的距离，找到与之距离最接近的样本，实现对目标的定位跟踪。

在行人再识别的过程中，首先将每个图像通过特征表达模型转化为特征向量，然后基于该特征计算待匹配图像对之间的距离。令 $\Delta_{ij} = x_i^p - x_j^s$ 表示一个样本对中两个样本特征向量的差向量。基于该差值向量对样本对任意 (x_i^s, x_j^p) 的距离进行定义。如公

式(1)，通过最简单的欧式距离进行度量，则样本距离的数学表达式如下所示：

$$d(x_i^s, x_j^p) = (x_i^s - x_j^p)^T (x_i^s - x_j^p) \quad (1)$$

然而，基于欧氏距离的样本相似性度量方法过于简单，行人图像外观特征在跨越监控场景的过程中变化剧烈，识别精度十分不理想。因此，简单的欧氏距离度量方法远远不能满足行人再识别的需求。对于度量学习方法的研究工作主要是学习一个马氏距离，因此，样本对的距离度量函数表示为：

$$d(x_i^s, x_j^p) = (x_i^s - x_j^p)^T M (x_i^s - x_j^p) \quad (2)$$

其中，度量矩阵 M 为半正定矩阵。此外，研究者通过学习一个投影子空间，在度量空间中对样本之间的距离进行度量，提高模型的分别能力，如下式(3)所示：

$$d(x_i^s, x_j^p) = (x_i^s - x_j^p)^T W W^T (x_i^s - x_j^p) \quad (3)$$

3 本文算法

为了充分利用已有的公开数据集，本文引入迁移学习的方法，建立迁移度量学习模型，实现度量模型在不同数据集之间的迁移过程，这样既增加了模型训练样本的多样性，有助于提取有效的分辨信息，又增强了模型的适应性。

3.1 交叉二次判别分析行人再识别方法

交叉二次判别分析(XQDA)算法是一种引入 Fisher 判别分析(FDA)算法理论到行人再识别任务中，建立的一种度量学习模型。

给定数据集 $O = \{x_i^p, x_j^s\}$ ，其中 O 表示带标签的训练集数据。由于典型的行人再识别图像外观特征是高维的特征向量，因此直接通过距离学习的方法求解。因此，令 $M = P A P^T$ ，其中 A 为对角阵，矩阵 P 的列元素为正交特征向量。定义 $W = P A^{\frac{1}{2}}$ ，则上述对于度量矩阵的分解形式可以表示为：

$$M = W W^T \quad (4)$$

因此，距离度量模型改写为如下形式：

$$d(x_i^s, x_j^p) = (x_i^s - x_j^p)^T W W^T (x_i^s - x_j^p) \quad (5)$$

其中 W 可以看作作为差值特征向量的子空间投

影矩阵。于是行人再识别度量学习问题转化为一个投影子空间的学习问题。对于该子空间的学习，投影矩阵 W 的目标是使得在投影空间中，正负样本对总体之间的差异尽量大，类内散度尽量小，从而使得正负样本两个类别尽可能的分离。

令 $\mathbf{o}_{ij} = \mathbf{x}_i^p - \mathbf{x}_j^g$ 表示待识别样本对差向量， $\mathbf{O}_+ = \{\mathbf{o}_{ij} | l_{ij} = 1\}$ 表示正样本对集合， $\mathbf{O}_- = \{\mathbf{o}_{ij} | l_{ij} = 0\}$ 表示负样本对集合。本文将行人再识别问题中的样本之间的距离度量问题转化为样本对的二分类问题，即样本对为正样本对或者负样本对。定义正负样本对总体散度为如下形式：

$$S_I = \frac{1}{n_1} \sum_{\mathbf{o}_{ij} \in \mathbf{O}_+} (\mathbf{o}_{ij} - \mathbf{u}_1)(\mathbf{o}_{ij} - \mathbf{u}_1)^T \quad (6)$$

$$S_E = \frac{1}{n_2} \sum_{\mathbf{o}_{ij} \in \mathbf{O}_-} (\mathbf{o}_{ij} - \mathbf{u}_2)(\mathbf{o}_{ij} - \mathbf{u}_2)^T \quad (7)$$

其中 S_I 表示正样本对总体散度， S_E 表示负样本对总体散度； \mathbf{u}_1 ， \mathbf{u}_2 为正负样本对差向量总体的均值。假设正负样本的差向量总体分别服从不同协方差的零均值高斯分布，因此：

$$S_I = \frac{1}{n_1} \sum_{\mathbf{o}_{ij} \in \mathbf{O}_+} \mathbf{o}_{ij} \mathbf{o}_{ij}^T \quad (8)$$

$$S_E = \frac{1}{n_2} \sum_{\mathbf{o}_{ij} \in \mathbf{O}_-} \mathbf{o}_{ij} \mathbf{o}_{ij}^T \quad (9)$$

投影子空间的学习通过解决最大化式(10)所示的 Fisher 判别函数的优化问题来实现。

$$\max_{\mathbf{w}} J(\mathbf{w}) = \frac{\mathbf{w}^T S_E \mathbf{w}}{\mathbf{w}^T S_I \mathbf{w}} \quad (10)$$

式(10)所示的 Fisher 判别准则等价于如下形式：

$$\max_{\mathbf{w}} \mathbf{w}^T S_E \mathbf{w}, \quad \mathbf{w}^T S_I \mathbf{w} = 1 \quad (11)$$

式(11)所示的问题可以通过拉格朗日方法来求解。

于是公式(11)优化问题转化为广义的特征根问题：

$$\lambda \mathbf{w} = S_I^{-1} S_E \mathbf{w} \quad (12)$$

其中矩阵 $S_I^{-1} S_E$ 的第一个特征向量 \mathbf{w}_1 是公式(10)所示优化问题最大特征根对应的最大特征向量，使得目标函数的值最优。同时矩阵 $S_I^{-1} S_E$ 的第二个特征向量 \mathbf{w}_2 是其次优的最大特征根对应的特征

向量。因此，投影子空间可以通过上述优化问题的前 m 个特征向量来组成 $W = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m)$ 。

此外，基于正负样本对的总体分布假设，在投影空间内学习一个马氏距离，来提高样本相似性的度量效果。从统计推断的角度来看，待检测样本对 $(\mathbf{W}^T \mathbf{x}_i^p, \mathbf{W}^T \mathbf{x}_j^g)$ 的相似性优化决策方案可以通过样本对分布的概率模型进行求解，定义正负样本的概率模型如下：

$$p_0(\mathbf{W}^T \mathbf{x}_i^p, \mathbf{W}^T \mathbf{x}_j^g) = \frac{1}{\sqrt{2\pi} |\sum_{l_{ij}=0}^{-1}|} \exp(-1/2 \mathbf{o}_{ij}^T \sum_{l_{ij}=0}^{-1} \mathbf{o}_{ij}) \quad (13)$$

$$p_1(\mathbf{W}^T \mathbf{x}_i^p, \mathbf{W}^T \mathbf{x}_j^g) = \frac{1}{\sqrt{2\pi} |\sum_{l_{ij}=1}^{-1}|} \exp(-1/2 \mathbf{o}_{ij}^T \sum_{l_{ij}=1}^{-1} \mathbf{o}_{ij}) \quad (14)$$

其中 $\sum_{l_{ij}=1}^{-1}$ ， $\sum_{l_{ij}=0}^{-1}$ 分别表示正负样本对差向量在投影空间中的协方差矩阵。其中 $p_0(\mathbf{W}^T \mathbf{x}_i^p, \mathbf{W}^T \mathbf{x}_j^g)$ 表示样本 $(\mathbf{x}_i^p, \mathbf{x}_j^g)$ 对属于负样本的概率，概率值越大则属于负样本对的可能性也就越大； $p_1(\mathbf{W}^T \mathbf{x}_i^p, \mathbf{W}^T \mathbf{x}_j^g)$ 表示样本 $(\mathbf{x}_i^p, \mathbf{x}_j^g)$ 对属于正样本的概率，概率值越大则属于正样本对的可能性也就越大。而对于样本对的距离，数值越大表示差异越大，属于正样本的概率越低，因此我们通过如下公式(15)定义样本距离函数：

$$\delta(\mathbf{x}_i^p, \mathbf{x}_j^g) = \gamma p_0(\mathbf{W}^T \mathbf{x}_i^p, \mathbf{W}^T \mathbf{x}_j^g) - p_1(\mathbf{W}^T \mathbf{x}_i^p, \mathbf{W}^T \mathbf{x}_j^g) \quad (15)$$

其中 γ 为平衡系数，通过平衡系数可以有效解决由于正负样本对数量不均衡问题导致的模型分辨能力差的问题。对距离模型(15)进行简化，首先对公式(15)右边取对数，得到公式(16)：

$$\begin{aligned} \delta(\mathbf{x}_i^p, \mathbf{x}_j^g) = & \frac{1}{2} \mathbf{o}_{ij}^T \sum_{l_{ij}=1}^{-1} \mathbf{o}_{ij} + \ln \left(\sqrt{2\pi} |\sum_{l_{ij}=1}^{-1}| \right) \\ & - \frac{\gamma}{2} \mathbf{o}_{ij}^T \sum_{l_{ij}=0}^{-1} \mathbf{o}_{ij} - \ln \left(\sqrt{2\pi} |\sum_{l_{ij}=0}^{-1}| \right) \end{aligned} \quad (16)$$

上式中的常数项使得样本间相似性在整体上发生一定的偏移，但并不影响样本对相似性的对比结果，因此移除上式中的常数项，得到如下公式(17)：

$$d(\mathbf{x}_i^p, \mathbf{x}_j^s) = \mathbf{o}_{ij}^T \left(\sum_{l_{ij}=1}^{-1} - \gamma \sum_{l_{ij}=0}^{-1} \right) \mathbf{o}_{ij} \quad (17)$$

最终，得到马氏距离的度量模型如下：

$$d(\mathbf{x}_i^p, \mathbf{x}_j^s) = (\mathbf{x}_i^p - \mathbf{x}_j^s)^T \mathbf{W} \mathbf{M} \mathbf{W}^T (\mathbf{x}_i^p - \mathbf{x}_j^s) \quad (18)$$

其中 $\mathbf{M} = \left(\sum_{l_{ij}=1}^{-1} - \gamma \sum_{l_{ij}=0}^{-1} \right)$ 。 $\sum_{l_{ij}=1}^{-1}$ ， $\sum_{l_{ij}=0}^{-1}$ 的数学表达式如下：

$$\sum_{l_{ij}=1}^{-1} = \frac{1}{n_1} \sum_{l_{ij}=1} \mathbf{W}^T (\mathbf{x}_i^p - \mathbf{x}_j^s) (\mathbf{x}_i^p - \mathbf{x}_j^s)^T \mathbf{W} \quad (19)$$

$$\sum_{l_{ij}=0}^{-1} = \frac{1}{n_2} \sum_{l_{ij}=0} \mathbf{W}^T (\mathbf{x}_i^p - \mathbf{x}_j^s) (\mathbf{x}_i^p - \mathbf{x}_j^s)^T \mathbf{W}$$

3.2 迁移度量学习模型

现有的度量学习算法通过建立距离度量函数的优化模型，约束正样本对相对于负样本对的距离学习一个投影度量空间或马氏距离，从而获取具有强分辨能力的度量模型。基于度量学习的样本距离(相似性)函数构造方法的假设前提是训练数据和测试数据的分布与总体分布相同。然而由于行人再识别样本特征变化的复杂性，包含有所有行人外观变化的样本总体分布是一个体量庞大的集合。已有的公开数据集仅仅是样本总体中很小的一部分。这导致了行人再识别任务中的小样本问题，现有的行人再识别度量模型通常在训练样本上过拟合。对此，本文中基于 VIPeR 和 iLIDS 数据集进行了分析。首先，通过 3.1 中的度量学习算法，在 VIPeR 数据集上训练得到度量模型；其次，基于训练得到的度量模型对 VIPeR 和 iLIDS 数据集中样本分别进行相似性度量，统计两个数据集的正负样本在度量空间中的距离分布的均值，结果如下表 1 所示。表 1 中给出了两个数据集中正负样本总体在度量空间中的相似距离的均值。由统计结果可以看出，基于现有的度量学习模型，目标域和源域数据在投影空间中的分布存在着巨大的差异。导致差异的原因是由于行人图像跨越场景时外观特征变化剧烈，现有模型训练数据过少，训练数据与测试数据存在较大差异。

表 1 样本相似距离统计结果

	VIPeR(均值)	iLIDS(均值)
正样本	8.2	259.4
负样本	1406.3	367.9

虽然不同数据集的拍摄环境有一定的差异性，但由于行人样本图像的结构具有共性，故基于不同数据集训练得到的度量模型对于行人再识别样本度量函数的学习仍具有很大的价值。因此，如何将已有的度量模型迁移到当前数据集的行人再识别模型中是本文研究要解决的问题。针对该问题，本文中引入迁移学习方法对现有度量学习算法进行改进，增强模型的适应性。迁移学习方法的目的是解决源域与目标域训练数据的分布不相同，机器学习算法的适应性问题。

给定原始目标域样本集 $T = \{\mathbf{t}_i^p, \mathbf{t}_j^s\}$ 和源域样本集 $S = \{\mathbf{s}_i^p, \mathbf{s}_j^s\}$ 。基于公式(12)和(18)的度量学习模型在目标域样本集上计算得到的度量空间 $(\mathbf{W}_0, \mathbf{H}_0)$ 。由于目标域和源域数据集的采集场景、外部环境以及设备条件通常具有很大差异，因此两个数据集的概率分布存在着很大的不同。为了减小两个数据集概率分布上的差异，可以通过在度量空间中约束源域数据的概率分布，以尽可能的接近目标域来实现。为了实现对两个样本集概率分布相似性的约束，本文在交叉二次判别分析的基础上，增加对样本在度量空间中概率分布的相似性约束。对于数据集概率分布相似性，定义约束项如下：

$$D_{TS}(T, S) = \left\| \frac{1}{N_T} \sum_{i=1}^{N_T} \mathbf{w}^T \mathbf{x}_{ij} - \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{w}^T \mathbf{y}_{ij} \right\|_2^2 \quad (20)$$

通过引入公式(20)的约束项，度量投影空间的优化目标函数改写为：

$$\max_{\mathbf{w}} J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_E \mathbf{w}}{\mathbf{w}^T \mathbf{S}_I \mathbf{w} + \alpha D_{TS}(T, S)} \quad (21)$$

公式(20)通过约束源域样本与目标域样本数据，在投影空间内总体分布中心之间的距离，来实现迁移度量学习。基于机器学习方法进行图像识别的假设，前提是训练数据与测试数据服从相同的分布，对于测试数据，均能在训练数据中找到与之相

似的样本，从而实现有效的识别。在行人再识别问题中，虽然训练样本与测试样本来自相同的样本总体，然而由于行人图像特征变化的复杂性以及小样本问题，基于训练数据得到的度量模型参数是有偏。训练数据与测试数据在投影空间中的分布存在较大差异。因此，我们将公式(20)源域与目标域数据分布中心的差异约束，引入到度量投影空间的学习模型中，得到公式(21)。从而使得源域数据与目标域数据在度量投影空间中的分布尽可能相似。

同样的，公式(21)所示的优化函数可以改写为等价形式如下：

$$\max_w w^T S_E w, \quad w^T (S_I + \alpha D_{TS}(T, S)) w = 1 \quad (22)$$

其中 α 表示平衡系数，上述公式(22)中的优化模型通过约束负样本对总体散度和不同数据集在投影空间中总体分布的差异性，最大化负样本散度来实现模型的迁移。通过求解公式(22)，得到目标域的度量模型。通过拉格朗日乘数法，将公式(22)中的优化模型转化为如下特征根问题：

$$\lambda w = (S_I + \alpha D_{TS}(T, S))^{-1} S_E w \quad (23)$$

通过求解公式(22)中的特征根问题，求解得到度量量子空间的投影矩阵 w 。基于投影矩阵，将样本数据投影到度量量子空间，并计算正负样本对的协方差，并基于公式(19)得到距离度量矩阵 M 。

3.3 计算方法

针对行人再识别任务中的小样本问题和过拟合现象，本文算法引入迁移学习的方法，对现有的度量学习算法进行了改进，通过约束不同数据集之间样本分布的差异，提高度量模型的分辨能力和适应能力。该方法一方面增加了度量模型训练样本的多样性，另一方面通过通过迁移度量学习增加了度量模型的适应性，本文算法具体的计算流程如图 2 所示。

如图 2 所示为基于迁移度量学习的行人再识别方法。本文算法首先在源域数据集上基于交叉二次判别分析，根据公式(12)和(19)利用标签信息学习得

到初始度量模型 (w_0, H_0) ；然后带入目标域数据，并基于公式(23)和(19)计算得到迁移度量学习模型；再带入测试样本到迁移度量模型训练得到的度量函数中，计算样本的识别结果；之后根据样本对之间距离的度量结果，按照距离由小到大的顺序进行排序；最后统计识别精度指标数据，输出识别结果。

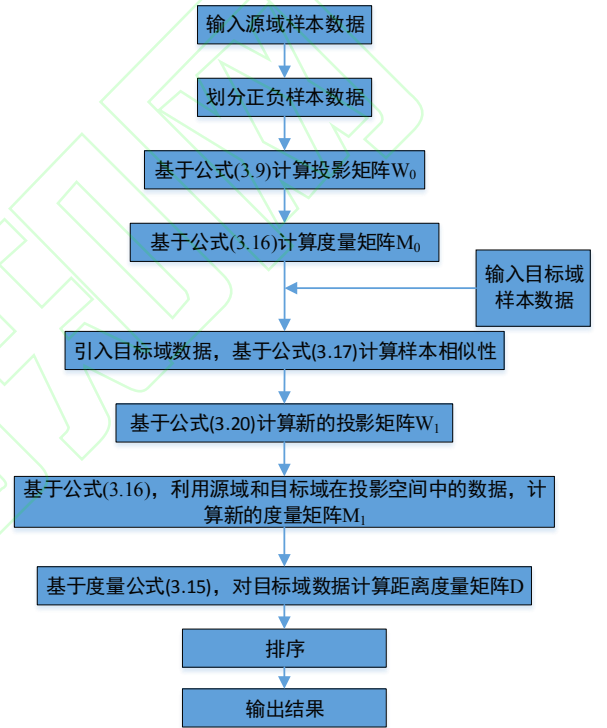


图 2 本文算法的计算流程图

4 仿真实验

4.1 数据库与仿真设定

本文研究中选取了 VIPeR、iLIDS、CUHK01、CUHK03 和 GRID 等五个数据集对基于迁移度量学习的行人再识别算法进行仿真实验，验证算法的有效性和识别精度。VIPeR^[3]数据集是一个被广泛应用的行人再识别算法测试数据集，该数据集共有 632 个行人的 1264 张图像，其中对应每个行人均有两张图像，并且这两张图像是由两个不同的摄像头在不同的时间不同的场景中捕捉得到的。数据集中所有的图像被归一化为相同的尺寸 128×48。图像的拍摄场景为户外场景，存在复杂的光照条件、拍摄角度以及背景的变化。与 VIPeR 数据集不同，

iLIDS^[6]是 multi-shot 的行人再识别数据集，即对应数据集中的每个行人，在每个摄像场景中包含有多张图像，并且每个人拥有的图像数量是不固定的。该数据集是来自中央电视台 CCTV 视频的基准库，用于对潜在威胁进行事件检测，该数据集的场景接近于政府要求下的实际操作条件。该数据集共有 119 个行人的 476 张图像，平均每个行人拥有 4 张图像。

本文的迁移度量学习，首先基于 iLIDS 数据库进行有监督的模型训练，然后基于迁移度量学习模型对 VIPeR 数据库中的样本进行测试。此外，设定算法参数 $\alpha=0.3$ 。

4.2 评价指标

本文中选取了累计精度曲线 (CMC curve) 作为算法的识别精度的评价指标，计算方法如下：

$$CMC(l) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(\text{rank}(P_i) < l) \quad (24)$$

其中 l 表示累积精度的排名数值，也就是计算排名小于等于 l 的识别结果中，存在正样本的概率。 N 表示用于匹配的样本的个数； $\mathbb{I}(\cdot)$ 为符号函数，若满足条件则取值 1，否则为 0。 $\text{rank}(\cdot)$ 表示样本距离排名计算， P_i 为第 i 个 gallery 样本的正样本距离， $\text{rank}(P_i)$ 表示其正样本的排序。

4.3 仿真实验

首先，基于 iLIDS 数据集样本，训练度量模型。在模型训练的过程中，一半个体作为训练样本，另一半作为测试样本。然后将基于 iLIDS 数据库训练得到的度量模型，通过迁移度量学习模型将度量模型迁移到 VIPeR 数据集。基于 VIPeR 数据库的测试结果如下表 2 和图 3 所示，图中给出了本文算法以及多种对比算法(SLD²L^[18]、RDC^[12]、ITML^[11]、KISSME^[14]、XQDA^[8]、和 MLAPG^[19])识别精度的 CMC 曲线。为确保对比实验结果的公平性，本文采用常见的测试方案，进行 10 次重复独立实验并计

算识别精度的平均值作为最终的识别精度评价指标。在行人再识别任务中，与其他分类和识别任务不同，我们关心的是不仅仅是距离最近的样本，对于排名比较靠前的结果我们都十分关心。因此，我们通过 CMC 曲线中选取了 rank-1、rank-5、rank-10 和 rank-20 共 4 个指标作为评价算法识别精度的依据，分别表示排名前 1、5、10、20 的正确识别结果所占全部识别结果的比率。从表中数据我们可以清楚的看出，本文算法在所有指标上均取得了最优的识别精度，尤其在 rank-1 的识别精度上，相较于 MLAPG 算法精度提升了 5.79%，达到了 46.14%。同时，本文算法在 rank-5、rank-10 和 rank-20 等指标上的识别精度也达到了最好的水平。同时，对于参数 α ，通过仿真试验，取 $\alpha=0.2$ 时模型的识别精度表现最好，因此本文中取 $\alpha=0.2$ 。

表 2 算法识别精度统计表

方法	p=316			
	r=1	r=5	r=10	r=20
SLD ² L ^[18]	16.86	41.22	58.06	95.57
RDC ^[12]	11.71	25.32	35.44	45.57
ITML ^[11]	15.46	38.32	53.06	69.68
KISSME ^[14]	19.35	47.89	62.77	77.21
MVSLD ² L ^[18]	20.79	45.08	61.24	81.36
XQDA ^[8]	37.44	64.81	77.63	89.40
MLAPG ^[19]	40.35	65.26	77.92	89.78
本文算法	46.14	65.48	77.68	89.12

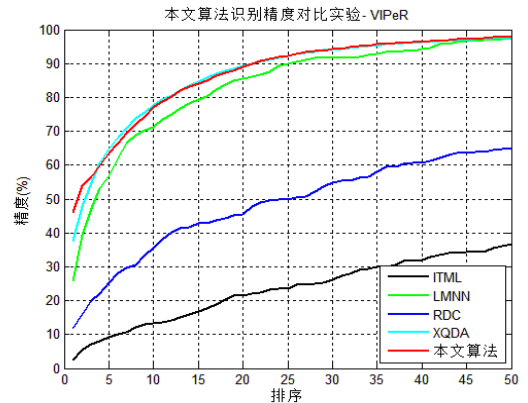


图 3 算法识别精度对比

针对度量模型中平衡系数 γ 的灵敏度测试，本

文中基于 VIPeR 数据集,在 γ 取不同值的条件下计算其识别精度,测试结果见图 4。如图 4 中所示,给出了 rank-1、rank-5、rank-10 和 rank-20 识别精度指标随 γ 数值变化的识别精度曲线。从图中可以看出,4 个识别精度指标的变化曲线具有相同的变化趋势,当 γ 的数值从 0 逐渐增长到 1.5 时,识别精度先逐渐增大,然后降低,并且当 $\gamma=1.2$ 时,模型具有最优的识别精度。

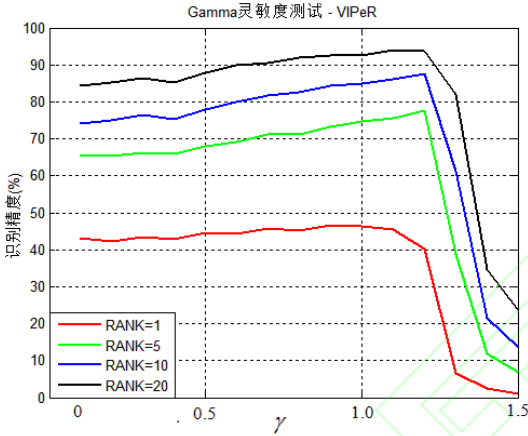


图 4 参数 γ 灵敏度测试曲线图

此外,我们在 CUHK01 数据库上对本文算法的识别精度进行了进一步的验证,结果如表 3 中所示。由表 3 中的识别结果可以明显的看出,在该数据库上,本文算法相较于其他对比算法具有最好的表现,尤其是 rank-1 的识别精度达到了 67.32%。

表 3 对比实验识别精度统计表(CUHK01)

方法	p=485			
	r=1	r=5	r=10	r=20
Improved Deep ^[16]	47.53	71.00	80.00	-
NFST ^[15]	65.0	85.0	89.9	94.4
XQDA ^[8]	63.21	83.89	90.04	94.14
MLAPG ^[19]	64.24	85.41	90.84	94.92
本文算法	67.32	86.75	91.25	95.39

为了进一步验证算法的有效性,本文在 GRID 数据集上进行了测试。GRID 数据集是现有的行人再识别数据集中最具挑战的一个数据集。与 VIPeR、CUHK01 等数据集不同,GRID 数据集包含有两部分,一部分是正常的数据样本,包含 250 个行人的 500 张图像,每个行人有两张图像,分别来自不同摄像头;另外一部分是干扰样本,共有 775 张图像,

并且图像分辨率较低。我们首先在 iLIDS 数据集上训练得到度量模型,然后迁移到 GRID 数据集对样本进行识别,识别精度如表 4 所示。由表中数据可以看出,基于迁移度量学习的行人再识别方法相较于 XQDA 算法,在识别精度上有了一定的提升,但由于 GRID 数据集样本质量较差,干扰样本较多,识别率仍然处于较低水平。

表 4 对比实验识别精度统计表(GRID)

方法	p=125		
	r=1	r=10	r=20
RDC ^[12]	9.68	32.96	44.32
NFST ^[15]	10.24	33.28	43.68
XQDA ^[8]	19.67	51.44	61.36
LSSCDL ^[18]	22.40	51.28	61.20
本文算法	21.52	51.67	61.58

将本文算法在 CUHK03 数据集上进行进一步验证。CUHK03 数据集是行人再识别数据集中样本量最大的数据集之一,包含了 1360 个行人的图像样本。本文选取该数据集中的人工标注图像进行算法验证,并且采用标准的测试方法,即选取 100 个行人样本作为测试数据,其余作为训练数据。通过表 5 中统计结果可以看出,基于迁移度量学习方法,有效的提升了识别精度,在 rank-1、rank-5、rank-10、rank-20 均取得了最好的识别精度,成功实现了度量模型在不同数据集之间的迁移。

表 5 对比实验识别精度统计表(CUHK03)

方法	Labeled			
	r=1	r=5	r=10	r=20
Improved Deep ^[16]	54.74	86.50	93.88	98.10
XQDA ^[8]	52.20	82.23	92.14	96.25
NFST ^[15]	58.90	85.60	92.45	96.30
本文算法	61.91	88.09	94.64	98.09

最后,本文对不同测试数据集在迁移度量学习和非迁移度量学习下,度量空间中样本分布进行了统计,结果如下表 6 所示,给出了不同数据集正样本在二次判别分析度量模型(通过 iLIDS 数据集训练得到)中的投影坐标均值以及在迁移度量学习投影空间中的坐标均值。由公式(11)可以看出,度量模型的学习是通过约束源域数据中正样本对的

散度，来获得具有强分辨能力的度量投影空间。通过对源域中正样本训练数据的度量，统计正样本训练数据在度量投影空间中距离的均值为 11.52。行人再识别度量学习模型中正、负样本对总体服从多元高斯分布假设，度量投影矩阵 W 是由原始空间到度量空间的线性变换，根据多元高斯分布的性质，原始数据在线性投影空间中的分布同样服从高斯分布。因此，目标域中匹配得到的样本对，在度量空间中的距离越接近源域正样本对距离的均值，其落在源域正样本对在度量空间中的分布概率越大，两个分布的相似性也就越高。根据表 6 中的统计结果可以得知，本文的迁移度量学习方法有效的提升了，目标域识别结果中样本对与源域正样本对在投影空间中分布的相似性。同时，迁移度量学习增加了模型训练样本的多样性，有效提升了识别精度。

表 6 rank-1 匹配样本对距离均值

数据集	二次判别分析	本文算法
VIPeR	267.3	156.3
CUHK01	219.6	139.8
GRID	326.8	207.4
CUHK03	248.5	167.3

5 结束语

本文针对行人再识别中的小样本问题和过拟合现象，提出了一种迁移度量学习算法，该算法在交叉二次判别分析度量模型的基础上，通过约束源域和目标域数据样本总体概率分布之间的差异，实现度量模型由源域到目标域的迁移。首先通过交叉二次判别分析方法，利用标签信息在源域上学习一个度量模型；然后，带入目标域数据信息，利用迁移度量学习模型，学习得到最终的距离度量模型。本文算法通过引入迁移学习方法，有效增加了度量模型学习的样本多样性，提升了度量模型的适应能力，改善了模型的过拟合问题。最后在多个数据集上对算法进行了测试，充分验证了本文方法的有效性。

参考文献:

- [1] Liu X., Tao D., Song M., and Zhang L., Bu, J. and Chen C. Learning to Track Multiple Targets[J]. IEEE Transactions on Neural Networks & Learning Systems, 2015, 26(5):1060.
- [2] Gray D and Tao Hai. Viewpoint invariant pedestrian recognition with an ensemble of localized features[C]// European Conference on Computer Vision, Marseille, 2008: 262-275.
- [3] Du Y., Ai H., and Lao S. Evaluation of color spaces for person re-identification[C]// IEEE Conference on Pattern Recognition, Tsukuba, 2012: 1371-1374.
- [4] Chen C L, Liu C, Gong S. Person re-identification by manifold ranking[C]// IEEE International Conference on Image Processing. IEEE, 2014:3567-3571.
- [5] M. Hirzer, P. M. Roth, M. Köstinger, and H. Bischof. Relaxed pairwise learned metric for person re-identification[C]// European Conference on Computer Vision, 2012:780-793.
- [6] Ojala T, Pietikäinen M, and Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 27(7):971-987.
- [7] Yang Y, Yang J, and Yan J et.al. Salientcolor names for person re-identification[C]// European Conference on Computer Vision, 2014: 536-551.
- [8] Liao S, Hu Y, Zhu X, and Li S Z. Person re-identification by local maximal occurrence representation and metric learning[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2015: 2197-2206.
- [9] 袁立, 田子茹. 基于融合特征的行人再识别方法[J]. 模式识别与人工智能, 2017, 30(3):269-278.
- [10] Zhang Z, Chen Y, Saligrama V. A Novel Visual Word Co-occurrence Model for Person Re-identification[C]// Workshop at the European Conference on Computer Vision. Springer International Publishing, 2014:122-133.
- [11] Davis J V, Kulis B, Jain P, et al. Information-theoretic metric learning [C]// in International Conference on Machine Learning, 2007:209-216.
- [12] ZhengWeishi, Gong Shaogang, Xiang Tao. Reidentification by relative distance comparison [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2013, 35(3): 653-668.
- [13] Zheng W S, Gong S, Xiang T. Person re-identification by probabilistic relative distance comparison[C]// Computer Vision and Pattern Recognition, 2011:649-656.
- [14] M. Köstinger, M. Hirzer, P. Wohlhart, P. Roth, and H. Bischof. Large scale metric learning from equivalence constraints[C]// IEEE Conference on Computer Vision and Pattern Recognition, pp. 2288-2295, 2012.
- [15] Zhang Li, Xiang Tao, Gong Shaogang. Learning a Discriminative Null Space for Person Re-identification[C]// Computer Vision and Pattern Recognition, 2016:1239-1248.
- [16] Li W, Zhao R, Xiao T, et al. DeepReID: Deep Filter Pairing Neural Network for Person Re-identification[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014:152-159.
- [17] Ahmed E, Jones M, Marks T K. An improved deep learning architecture for person re-identification[C]// Computer Vision and Pattern Recognition. IEEE, 2015: 3908-3916.
- [18] Chen F, Chai J, Ren D, et al. Semi-supervised distance metric learning for person re-identification[C]// Interna-

tional Conference on Computer and Information Science, 2017:733-738.

- [19] Liao Shengcai, Li S Z. Efficient PSD Constrained Asymmetric Metric Learning for Person Re-Identification

[C]// IEEE International Conference on Computer Vision. IEEE, 2015:3685-3693.

