

# 基于度量学习的步态识别比较研究

刘 东<sup>1</sup>, 胡峻林<sup>2\*</sup>

(1. 北京化工大学 信息科学与技术学院, 北京 100029;

2. 北京航空航天大学 软件学院, 北京 100191)

**摘 要:**步态识别是一项根据人在行走中的动作特征进行身份识别的技术,与其他生物特征识别方法相比,其可以实现远距离、非接触识别,在监控、安防等领域有着广泛的应用。度量学习在模式识别任务中起着非常重要的作用,其从训练样本中学习出合适的距离函数来度量样本间的相似性,以提高识别率。为此,从度量学习视角对步态识别问题进行研究,比较分析几种经典的度量学习方法在步态识别中的性能表现。首先从步态序列中提取步态能量图作为行人的步态特征,然后使用度量学习方法学习距离度量,使得在该度量下同一类样本间的距离最小化和不同类样本间的距离最大化,以提升步态识别的正确识别率。在广泛使用的CASIA-B与CASIA-C步态数据集上进行了一系列的对比实验,实验结果展示了几种度量学习方法的识别性能,为今后的步态识别研究提供了一些基准结果。

**关键词:**步态识别;度量学习;特征提取;人体生物特征;模式识别

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2022)09-0089-06

doi:10.3969/j.issn.1673-629X.2022.09.014

## Comparative Study of Gait Recognition Based on Metric Learning

LIU Dong<sup>1</sup>, HU Jun-lin<sup>2\*</sup>

(1. School of Information Science and Technology, Beijing University of Chemical Technology,  
Beijing 100029, China;

2. School of Software, Beihang University, Beijing 100191, China)

**Abstract:**Gait recognition is a kind of technology for human identification based on the characteristics of a person's movement during walking. Compared with other biometrics methods, gait recognition can realize non-contactless recognition at a distance, which has been widely used in the fields of surveillance, security and so on. Metric learning plays a quite important role in pattern recognition tasks, which aims to learn appropriate distance functions from training data itself to measure the similarity between samples, thereby improving recognition accuracy. We investigate the gait recognition problem from the perspective of metric learning, and compare and analyze the performance of several classic metric learning methods for gait recognition. The gait energy image (GEI) of each gait sequence is extracted to obtain the robust gait feature representation. Then the metric learning methods are utilized to learn the distance metrics, under which the distance between the samples from the same class is minimized and that of samples from different classes is maximized simultaneously, to improve the correct recognition accuracy of gait recognition. Extensive experiments are carried out on the widely used CASIA-B and CASIA-C gait datasets, which show the recognition performance of several metric learning methods and provide benchmark results for gait recognition.

**Key words:**gait recognition; metric learning; feature extraction; biometrics; pattern recognition

## 0 引言

随着计算机视觉和模式识别的快速发展,传统的生物识别技术已经取得了极大的突破,如指纹识别、面部识别、虹膜识别等,但是这些方法通常需要待识别者一定的配合,对采集的图像的质量要求也比较高,并且

需要在近距离内完成识别。与传统的识别技术相比,步态识别是从行人的行走姿态中提取特征进行检测和比较,因而可以实现远距离、无接触的人体身份识别。近年来,步态识别已经广泛应用于安防、监控、医疗<sup>[1-3]</sup>等领域。

收稿日期:2021-10-18

修回日期:2022-02-23

基金项目:国家自然科学基金项目(62006013);北京市自然科学基金项目(4204108)

作者简介:刘 东(1996-),男,硕士研究生,研究方向为计算机视觉与模式识别;通讯作者:胡峻林(1986-),男,博士,副教授,CCF会员(96236M),研究方向为计算机视觉与模式识别。

近二十年来,不断有新的步态识别方法提出,但是基于视觉的步态识别仍然面临着一些挑战:首先是拍摄角度问题<sup>[4]</sup>,在实际应用中,行人通常以不同的角度出现在画面中,不同视角下行人的轮廓会存在明显的差别;其次,行人的状态如着装的变化、是否携带物体如背包等也是影响识别的重要因素;此外,一些环境因素如光线、遮挡、低分辨率等同样也会影响算法模型的性能。这些问题往往会改变行人的外观轮廓,进而干扰识别结果。因此,步态识别技术仍然需要进一步研究发展。

该文从度量学习视角来研究步态识别问题,并比较分析了几种经典的度量学习方法在步态识别上的性能表现,这些方法包括余弦相似度度量学习(Cosine Similarity Metric Learning, CSML)<sup>[5]</sup>、信息论度量学习(Information Theoretic Metric Learning, ITML)<sup>[6]</sup>、几何平均度量学习(Geometric Mean Metric Learning, GMML)<sup>[7]</sup>、KISS 度量学习(KISS Metric Learning)<sup>[8]</sup>、边信息线性判别分析(Side-Information based Linear Discriminant Analysis, SILD)<sup>[9]</sup>。

## 1 相关工作

### 1.1 步态识别

现有的步态识别方法主要分为两大类:基于模板序列的和基于视频的<sup>[10]</sup>。基于模板序列的步态识别方法主要将步态序列转化为单张图像进行识别,这类方法首先从视频序列的每一帧中得到人体轮廓,然后将所有人体轮廓对齐并归一化成大小相同的步态模板。得到步态模板之后,可以通过机器学习方法或者深度学习方法从步态模板中提取步态特征,最后使用某种距离度量如欧几里得距离、余弦距离等度量图像之间的相似度并完成识别。步态模板的作用是将一个行走序列的信息压缩到一张图像中,典型的方法是步态能量图(Gait Energy Image, GEI)<sup>[11]</sup>,由于行人在行走的时候动作是有周期性的,因此可以用单个二维步态模板代替视频序列来表示行走的特征信息。将一个行走周期内的人体轮廓图像累加后再求平均就得到一张 GEI。使用 GEI 作为特征代替视频序列能够节省样本的存储空间并缩短识别的计算时间。基于视频的方法则直接将轮廓序列作为输入,从轮廓序列中提取特征,代表性的方法包括 LSTM(Long-Short-Term Memory)<sup>[12]</sup>和 3D-CNN<sup>[13]</sup>等基于深度学习的方法。这类方法虽然能够保留更多的时间和空间信息,但是计算量巨大,并且训练深度模型具有一定的困难。

近年来,不断有新的步态识别方法提出。为了解决在煤矿井下人脸、指纹等生物识别效果不理想的问题,刘晓阳等<sup>[14]</sup>提出一种基于双流神经网络的步态识

别方法。该方法利用残差神经网络和栈式卷积自动编码器从步态中提取特征,并使用特征融合方法实现动态和静态特征的融合。受 CNN 在图像识别方面应用的启发,Shiraga 等<sup>[12]</sup>提出 GEINet,将 GEI 作为 CNN 网络的输入,经过两个连续的卷积层、池化层、归一化层,最后经过两个全连接层完成识别。Wu 等<sup>[15]</sup>提出一种基于深度卷积神经网络的步态识别方法,该方法中设计了 3 个不同的以 GEI 为输入的网络结构,最后将这些网络融合得到识别的结果。相较于已有的方法,该方法在视角变化较大时具有明显的优势。Yu 等<sup>[16]</sup>提出一种基于生成性对抗网络(Generative Adversarial Networks, GAN)的步态识别方法,称为 GaitGANv2。在 GaitGANv2 中,生成器用于生成待识别者的标准步态图像,即穿着正常衣服且没有携带任何物品的侧视图,这样可以消除角度、着装等因素的干扰。与传统 GAN 不同的是,GaitGANv2 中包含两个判别器,第一个为真假判别器,用于判别生成的步态图像的真假,第二个为身份判别器,用于保证生成步态图像的同时不丢失行人的身份信息。同时,GaitGANv2 还采用多损耗策略优化网络,在增大类间距离的同时减小类内距离。Chao 等<sup>[17]</sup>提出一种端到端的深度模型 GaitSet。GaitSet 中,步态序列不再保持特定的顺序,而是被视为一组相互独立的帧。该方法不受帧排列的影响,并且可以自然地整合不同场景下拍摄的不同视频的帧,例如不同的视角、不同的衣服等,在复杂的场景下有较好的性能。

### 1.2 度量学习

度量学习的思想起源于分类问题中的最近邻方法,度量样本数据间的相似性并将一个样本划分为距离其最近的类别。作为一种最常见的距离,欧氏距离在早期度量学习中有着广泛的应用,然而单一的距离度量无法应对更多复杂的问题。因此,度量学习的基本目的是学习一种合适的距离度量,使得分类问题中,同一类内样本间的距离最小化,不同类样本间的距离最大化,从而完成分类。

作为欧氏距离的一种推广,马氏距离越来越多地被应用于度量学习算法中,典型的有信息论度量学习(ITML)<sup>[6]</sup>、大边界成分分析(Large Margin Component Analysis, LMCA)<sup>[18]</sup>等。然而,单一的马氏距离只能学习到线性变换,一些非线性的问题就难以得到解决。因此,近年来一些深度度量学习方法不断被提出。Hadshell 等<sup>[19]</sup>最早将深度学习引入度量学习中,他们提出一种对比损失函数用于实现最大化类间距离并最小化类内距离。在此基础上,后续不断有研究者做出改进,提出了三元组损失<sup>[20]</sup>、N 元组损失<sup>[21]</sup>、中心损失<sup>[22]</sup>、代理损失<sup>[23]</sup>等方法。近年来,度量学习被广泛

运用于人脸识别<sup>[24]</sup>、行人重识别<sup>[25]</sup>等领域。

鉴于度量学习在模式识别中的优势,该文从度量学习视角来研究步态识别问题,并比较分析了几种经典的度量学习方法在步态识别上的性能。

## 2 度量学习方法

在本节中,主要介绍几种用于步态识别的度量学习方法,这些方法包括余弦相似度度量学习(CSML)<sup>[5]</sup>、信息论度量学习(ITML)<sup>[6]</sup>、几何平均度量学习(GMML)<sup>[7]</sup>、KISS度量学习(KISSME)<sup>[8]</sup>和边信息线性判别分析(SILD)<sup>[9]</sup>。

### 2.1 CSML

余弦相似度度量学习的目的是学习一个线性变换  $W \in R^{d \times m}$ ,  $m \leq d$ , 其中  $d$  是变换前的空间维度,  $m$  是变换后的空间维度,来计算一对样本  $x_i$  和  $x_j$  在这个变换子空间中的余弦相似度:

$$CS_W(x_i, x_j) = \frac{x_i^T W W^T x_j}{\sqrt{x_i^T W W^T x_i} \sqrt{x_j^T W W^T x_j}} \quad (1)$$

为了学习出最优的变换  $W$ , CSML 通过将目标函数表示为以下形式来最小化交叉验证误差:

$$\max_W F(W) = \sum_{l_{ij}=1} CS_W(x_i, x_j) - \alpha \sum_{l_{ij}=-1} CS_W(x_i, x_j) - \beta \|W - W_0\|^2 \quad (2)$$

其中,  $l_{ij}=1$  表示  $x_i$  和  $x_j$  为一对正样本对(即来自同一类别);  $l_{ij}=-1$  表示  $x_i$  和  $x_j$  为一对负样本对(即来自不同的类别);  $W_0$  是一个预定义的矩阵,如随机矩阵;参数  $\alpha$  用于平衡正负样本对对目标函数的贡献;  $\beta$  用于控制正则化项  $\|W - W_0\|^2$  的权重。最后采用基于梯度的策略求解出合适的  $W$ 。

### 2.2 ITML

ITML 方法采用广泛使用的马氏距离度量学习框架。马氏距离度量学习的目的是寻找一个半正定矩阵  $M \in R^{d \times d}$  来计算两个样本  $x_i$  和  $x_j$  之间的马氏距离平方:

$$d_M^2(x_i, x_j) = (x_i - x_j)^T M (x_i - x_j) \quad (3)$$

ITML 的核心思想就是,在正样本对之间的马氏距离小于一个阈值  $\tau_p$  和负样本对之间的马氏距离大于另一个阈值  $\tau_n$  的约束条件下,  $\tau_n > \tau_p > 0$ , 通过最小化两个矩阵之间的对数离散度来寻找一个接近先验度量  $M_0$  的距离度量  $M$ 。具体可以表示为以下的优化问题:

$$\begin{aligned} \min_M D_{ld}(M, M_0) &= \text{tr}(M M_0^{-1}) - \log \det(M M_0^{-1}) - d \\ \text{s. t. } d_M^2(x_i, x_j) &\leq \tau_p, \forall l_{ij} = 1 \\ d_M^2(x_i, x_j) &\geq \tau_n, \forall l_{ij} = -1 \end{aligned} \quad (4)$$

其中,  $\text{tr}$  表示矩阵的迹运算,  $\det$  表示矩阵的行列式运算。

通过迭代计算 Bregman 投影,可以将式(4)简化为以下形式:

$$M_{t+1} = M_t + \beta M_t (x_i - x_j) (x_i - x_j)^T M_t \quad (5)$$

其中,投影参数  $\beta$  由样本对标签和学习率决定;  $t$  表示迭代次数。依照式(5)迭代直到算法满足一定收敛条件,便可获得需要的  $M$ 。在测试阶段,当两个样本之间的距离小于一个给定的阈值时,即认为它们是同类的样本,当样本间的距离大于一个给定的阈值时,即认为它们是来自不同类的样本。

### 2.3 GMML

几何平均度量学习的目的是寻找一个矩阵  $M$ , 用于减小正样本对之间的距离,但是与传统方法不同的是,对于负样本对,GMML 使用  $M^{-1}$  来度量它们之间的距离,而不是使用非对称的方法。GMML 是通过使用  $M$  最小化所有正样本对之间的距离,同时用  $M^{-1}$  最小化所有负样本对之间的距离来学习一个矩阵  $M$ 。GMML 的目标函数可以表示为如下的优化问题:

$$\min J(M) = \text{tr}(MS) + \text{tr}(M^{-1}D) \quad (6)$$

其中,  $S$  和  $D$  分别表示:

$$S = \frac{1}{N_s} \sum_{l_{ij}=1} (x_i - x_j) (x_i - x_j)^T \quad (7)$$

$$D = \frac{1}{N_d} \sum_{l_{ij}=-1} (x_i - x_j) (x_i - x_j)^T$$

这里,  $N_s$  和  $N_d$  分别表示正样本对和负样本对的总数目。

根据文献[7]中的理论3可知,GMML 的目标函数是严格凸的。因此,式(6)中目标函数的全局最小值可以由下面过程求得闭式解:

$$\frac{\partial J(M)}{\partial M} = S - M^{-1} D M^{-1} = 0 \quad (8)$$

$$M = S^{-1/2} (S^{1/2} D S^{1/2})^{-1/2} S^{-1/2} \quad (9)$$

### 2.4 KISSME

KISSME 方法将一对样本是否为同一类(正样本对或负样本对)考虑为两个独立的过程。从统计学的角度来看,可以通过似然比检验得到一对样本对是否为同一类的最优统计决策。检验假设  $H_0$ : 样本对不是同类样本;  $H_1$ : 样本是同一类样本,则似然比检验为:

$$\delta(x_i, x_j) = \log \left( \frac{p(x_i, x_j | H_0)}{p(x_i, x_j | H_1)} \right) \quad (10)$$

当式(10)的值大于一个阈值时,接受  $H_0$ ; 当式(10)的值小于阈值时,接受  $H_1$ , 即认为两个样本为同一类。KISSME 方法假定样本对之间的差为单高斯分布,通过取对数并剥离常数项式(10)简化为:

$$\delta(x_i, x_j) = x_{ij}^T (S^{-1} - D^{-1}) x_{ij} \quad (11)$$

这里,  $x_{ij} = x_i - x_j$  表示样本对之间的差向量; 方差矩阵  $S$  和  $D$  通过式(7)计算求得。



最终可以通过将  $\hat{M} = S^{-1} - D^{-1}$  投影到半正定矩阵  $M$  的锥上,得到反映对数似然比检验性质的马氏距离度量  $M$ 。

## 2.5 SILD

与传统方法 Fisher 线性判别分析 (FLDA) 相似, SILD 通过最大化类间散度矩阵  $S_B$  的行列式与类内散度矩阵  $S_W$  的行列式的比率来找到一组最具判别力的线性投影,投影矩阵可以通过求解以下的优化问题获得:

$$W_{\text{opt}} = \underset{W}{\operatorname{argmax}} \frac{|W^T S_B W|}{|W^T S_W W|} \quad (12)$$

然而在 FLDA 中,所有样本的标签信息都必须已知,在某些类标签信息未知的情况下,  $S_B$  和  $S_W$  将无法求解。为了解决这一问题, SILD 使用边信息计算散度矩阵。具体来说,通过式 (7) 计算正样本对的  $S$  作为类内散度矩阵,即  $S_W = S$ ; 计算负样本对的  $D$  作为类间散度矩阵,即  $S_B = D$ 。

后续的求解方式与 FLDA 类似,由于正负样本对集合的大小显然会影响  $S_B$  和  $S_W$  的稳定性,导致类内散度矩阵会有大量较小的特征值,因此 SILD 中将  $S_W$  对角化,只使用最大特征值对应的特征向量:

$$S_W = H \Lambda H^T \quad (13)$$

在 SILD 中,定义一个由  $\Lambda$  的部分列组成的矩阵  $\Lambda'$ ,  $\Lambda'$  保留了  $\Lambda$  中较大的部分特征值,使用  $\Lambda'$  代替  $\Lambda$  进行后续运算可以有效地解决不稳定性问题。

当求得变换矩阵  $W = W_{\text{opt}}$  之后,一对样本之间的距离可以表示为:

$$d_w^2(x_i, x_j) = \|W^T x_i - W^T x_j\|_2^2 = (x_i - x_j)^T M (x_i - x_j) \quad (14)$$

其中,  $M = W W^T$ 。

## 3 实验与结果

本小节在两个大规模的步态数据集上进行了一系列实验,评估了几种度量学习方法的性能。

### 3.1 数据集

选择在 CASIA-C<sup>[26]</sup> 与 CASIA-B<sup>[27]</sup> 步态数据集上进行实验。CASIA-C 是一个使用红外摄像机在夜间拍摄的大规模数据集,包含 153 个样本,每个样本共有 10 组步态序列,分别为 4 组正常行走的序列,2 组快步行走的序列,2 组慢步行走的序列以及 2 组背包行走的序列。在该数据集中不包含视角的变化,因此每一组序列中行人都以 90° 的视角出现。在 CASIA-C 中,使用前 76 个样本的所有数据作为训练集,测试集中,使用每个样本的前 2 个正常行走序列作为 Gallery 集 (fn00–fn01), Probe 集则划分了 4 组进行实验,分别是

后 2 个正常行走序列 (fn02–fn03), 2 个快步行走的序列 (fq00–fq01), 2 个慢步行走的序列 (fs00–fs01) 以及 2 个背包行走的序列 (fb00–fb01)。CASIA-B 是一个大规模、多视角的步态数据集,共包含 124 个样本,每个样本有 10 种步态序列,分别为 6 个正常行走的序列,2 个穿着长外套的行走序列和 2 个背包的行走序列。与 CASIA-C 相比, CASIA-B 更关注视角的变化,因此每种行走序列下,又分为 11 个不同的角度 (0°, 18°, ..., 180°), 所以,每个样本共包含 110 个步态序列。实验的设置上,选择目前比较主流的方案,使用前 62 个样本的所有数据作为训练集。测试集中,使用每个样本的前 4 个正常行走序列作为 Gallery 集 (NM#1–4), 为了研究模型在轮廓变化时的性能表现,划分了 3 个 Probe 集,分别为正常行走的后 2 个序列 (NM#5–6), 2 个穿着长外套的序列 (CL#1–2) 和 2 个背包的序列 (BG#1–2)。考虑到角度的影响,在实验时,对 Probe 集中每一个角度都单独测试了识别率。

在实验中,使用步态能量图 (GEI) 来表示行人的步态特征,最终每一个步态序列生成的 GEI 大小为 64×32, 如图 1 和图 2 所示。

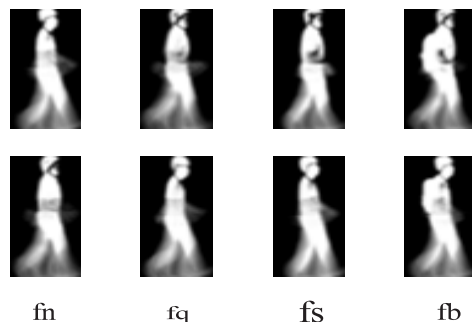


图 1 CASIA-C 数据集生成的 GEI

(fn、fq、fs 和 fb 分别表示正常行走情形下、快步行走情形下、慢步行走情形下和背包行走的步态序列)

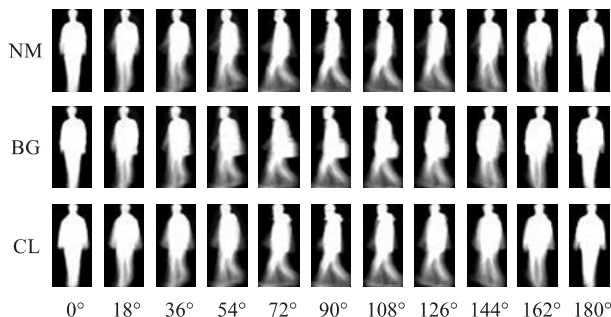


图 2 CASIA-B 数据集生成的 GEI

(NM、BG 和 CL 分别表示正常行走情形下、背包情形下和穿外套情形下的步态序列)

### 3.2 实验结果和分析

在本节中,分别评估了几种度量学习方法在 CASIA-C 和 CASIA-B 步态数据集上的识别率,来比较几个度量学习模型在步态识别中的性能表现。首先

将  $64 \times 32$  的 GEI 转成  $64 \times 32 = 2\,048$  维的特征向量,然后通过主成分分析(PCA)方法将训练样本的特征降到 500 维,以加快计算速度和去除部分噪声的干扰。然后分别使用 CSML、ITML、GMML、SILD、KISSME 以及欧氏距离(L2)度量样本之间的距离,并对结果进行了详细的对比。

表 1 显示了几种方法在 CASIA-C 数据集上的识别准确率,其中,KISSME 方法在 4 组实验中均取得了最好的识别精度。快步行走和慢步行走的情况下识别精度相较于正常行走略有下降,KISSME 方法在这两组实验中都达到了 98% 以上的精度。背包的情形下,由于行人的轮廓发生较大的变化,识别精度相较于前 3 组实验的精度有所下降,KISSME 方法在此情形下的识别精度为 60.39%。

表 1 CASIA-C 数据集的识别准确率 %

Method	fn02-03	fq00-01	fs00-01	fb00-01
CSML	100.00	89.61	92.21	42.86
ITML	99.35	90.91	89.61	33.12
GMML	100.00	97.40	97.40	57.79
L2	100.00	93.51	95.45	49.35
KISSME	100.00	98.70	98.05	60.39
SILD	92.86	68.83	79.87	29.87

表 2 至表 4 显示了几种方法在 CASIA-B 数据集上的表现。表 2 显示了几种方法在正常行走情形下的识别准确率,其中,SILD 方法的平均识别准确率达到 99.78%,明显优于其他几种方法。CSML,GMML,L2 三种方法的平均识别准确率也达到了 97% 以上。在正常行走的场景下,平均识别率普遍较高。表 3 给出了在背包情况下的识别准确率,SILD 方法的平均识别准确率达到 64.26%,在  $90^\circ$  及  $108^\circ$  视角下的识别率都在 50% 以上,而其他方法的识别率在该视角下都低于 40%,SILD 方法明显优于其他方法。CSML,GMML,L2 三种方法的平均识别率都在 50% 以上,而 ITML 的平均识别率最低,只有 34.44%。同时可以观察到,背包的情形中, $90^\circ$  视角下的识别率普遍较低,而  $0^\circ$  和  $180^\circ$  视角下的识别率则相对较高,这是因为背包对行人侧面视角的轮廓影响较大,对  $0^\circ$  和  $180^\circ$  视角的影响相对较小。表 4 列出了穿外套情形下的识别准确率,SILD 方法在各个角度下的识别准确率都优于其他方法,但平均准确率仅有 26.54%。在穿外套情形中,各个方法的识别率都较低,而且不同角度下的识别率没有明显的区别,这是因为穿外套对各个角度下的行人轮廓都有较大的影响。

表 2 CASIA-B 数据集正常行走(NM#5-6)下的识别准确率 %

Method	$0^\circ$	$18^\circ$	$36^\circ$	$54^\circ$	$72^\circ$	$90^\circ$	$108^\circ$	$126^\circ$	$144^\circ$	$162^\circ$	$180^\circ$	Mean
CSML	100	100	99.19	95.16	95.97	97.58	95.97	96.77	98.39	97.58	100	97.87
ITML	96.77	97.58	97.58	95.97	91.94	93.55	97.58	94.35	94.35	95.16	95.16	95.45
GMML	100	100	97.58	95.97	98.39	99.19	97.58	96.77	98.39	97.58	100	98.31
L2	100	100	97.58	95.16	97.58	97.58	95.97	95.97	97.58	97.58	100	97.73
KISSME	97.58	24.19	95.97	96.77	91.94	95.16	96.77	94.35	93.55	92.74	95.16	88.56
SILD	100	100	100	100	99.19	99.19	100	100	100	99.19	100	99.78

表 3 CASIA-B 数据集背包状态下(BG#1-2)的识别准确率 %

Method	$0^\circ$	$18^\circ$	$36^\circ$	$54^\circ$	$72^\circ$	$90^\circ$	$108^\circ$	$126^\circ$	$144^\circ$	$162^\circ$	$180^\circ$	Mean
CSML	70.97	67.74	59.68	55.28	39.52	28.23	33.06	47.58	48.39	60.98	67.74	52.65
ITML	46.77	46.77	39.52	29.27	25.00	19.35	14.52	32.26	33.06	43.90	48.39	34.44
GMML	75.00	72.58	68.55	60.16	39.52	33.06	38.71	56.45	57.26	66.67	71.77	58.16
L2	73.39	70.16	62.9	55.28	37.90	33.06	38.71	52.42	54.03	63.41	71.77	55.73
KISSME	58.87	16.13	58.06	44.72	20.97	30.65	28.23	45.97	44.35	47.15	48.39	40.32
SILD	72.58	73.39	73.39	65.04	64.52	51.61	55.65	56.45	61.29	65.04	67.74	64.26

表 4 CASIA-B 数据集穿外套状态下(CL#1-2)的识别准确率 %

Method	$0^\circ$	$18^\circ$	$36^\circ$	$54^\circ$	$72^\circ$	$90^\circ$	$108^\circ$	$126^\circ$	$144^\circ$	$162^\circ$	$180^\circ$	Mean
CSML	9.68	16.13	15.32	22.58	15.32	16.13	13.71	12.9	12.9	9.68	10.48	14.08
ITML	5.65	5.65	5.65	4.84	4.84	4.84	4.03	7.26	8.06	5.65	8.87	5.94
GMML	10.48	16.94	16.94	20.16	12.90	12.10	11.29	12.10	11.29	11.29	11.29	13.34
L2	9.68	15.32	17.74	17.74	12.10	13.71	12.90	13.71	12.10	12.10	11.29	13.49

续表 4

Method	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	Mean
KISSME	12.90	7.26	18.55	19.35	14.52	14.52	23.39	18.55	16.13	16.13	12.10	15.76
SILD	21.77	25.81	37.90	33.06	27.42	26.61	25.81	25.00	28.23	19.35	20.97	26.54

对比几种度量学习方法在两个数据集上的表现,可以发现 KISSME 在应对单一视角的情况下有较好的性能表现,而在视角有较大变化的情况下,SILD 方法则有绝对的优势。因此,在步态识别任务中,应对行人以单一视角出现的场景,KISSME 是一种推荐的度量学习方法,而当行人以多视角出现时,SILD 拥有更优异的性能表现。同时,对比几组实验的结果可以观察到,背包情况下的识别精度相对于正常行走的识别精度有所下降,在此情形下最高的识别精度为 64.26%,而穿外套情形下的识别精度最低,SILD 方法在此情形下的识别精度也仅为 26.54%。与传统的生物识别方式相比,步态识别有着巨大的应用潜力,但是现阶段的步态识别仍然面临一些挑战,结合实验结果来看,目前轮廓的变化仍然是影响步态识别算法性能的一大障碍。今后的步态识别研究中,可以重点关注算法在轮廓发生变化时的性能表现,优化算法在背包情形和穿外套情形下的识别精度,进一步提高算法应对复杂多变场景的能力,让实时高效的步态识别应用成为可能。

#### 4 结束语

从度量学习角度研究了步态识别问题,并评估了几种度量学习方法在步态识别中的性能表现。首先使用步态能量图 GEI 作为行人步态特征,然后采用主成分分析方法将目标特征进行降维,最后使用几种度量学习方法在训练集上学习出数据之间的距离度量,并应用距离度量计算测试集样本之间的相似性。在 CASIA-C 与 CASIA-B 步态数据集上的实验结果展示了几种方法的识别性能,也表明了跨视角步态识别问题的挑战性,为今后的研究提供了一些基准结果。

#### 参考文献:

- [1] VERLEKART T, CORREIA P L, SOARES L D. Using transfer learning for classification of gait pathologies [C]//2018 IEEE international conference on bioinformatics and biomedicine. Madrid:IEEE,2018:2376-2381.
- [2] 高毅. 基于步态识别的跨场景多目标跟踪算法[J]. 控制工程,2021,28(7):1375-1381.
- [3] 尔雅莉. 基于计算机人工智能下识别技术的应用[J]. 电子技术与软件工程,2021(5):124-125.
- [4] 韩东岳,桑海峰. 基于双路网络的步态识别方法[C]//第十八届沈阳科学学术年会论文集. 沈阳:沈阳市科学技术协会,2021:8.
- [5] NGUYEN H V, BAI L. Cosine similarity metric learning for face verification [C]//10th Asian conference on computer vision. Queenstown:Springer-Verlag,2010:709-720.
- [6] DAVIS J V, KULIS B, JAIN P, et al. Information-theoretic metric learning [C]//International conference on machine learning. New York:ACM,2007:209-216.
- [7] ZADEH P, HOSSEINI R, SRA S. Geometric mean metric learning [C]//International conference on machine learning. New York:IMLS,2016:2464-2471.
- [8] MARTIN K, 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.
- [9] KAN M, SHAN S, XU D, et al. Side-information based linear discriminant analysis for face recognition [C]//Proceedings of the British machine vision conference. Dundee:BMVA Press,2011.
- [10] 朱应钊,李嫚. 步态识别现状及发展趋势[J]. 电信科学,2020,36(8):130-138.
- [11] HAN J, BHANU B. Individual recognition using gait energy image [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2006,28(2):316-322.
- [12] SHIRAGA K, MAKIHARA Y, MURAMATSU D, et al. GEI-Net:view-invariant gait recognition using a convolutional neural network [C]//2016 international conference on biometrics. Halmstad:IEEE,2016:1-8.
- [13] WOLF T, BABAEI M, RIGOLL G. Multi-view gait recognition using 3D convolutional neural networks [C]//2016 IEEE international conference on image processing. Phoenix:IEEE,2016:4165-4169.
- [14] 刘晓阳,刘金强,郑昊琳. 基于双流神经网络的煤矿井下人员步态识别方法[J]. 矿业科学学报,2021,6(2):218-227.
- [15] WU Z, HUANG Y, WANG L, et al. A comprehensive study on cross-view gait based human identification with deep CNNs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(2):209-226.
- [16] YU S, LIAO R, AN W, et al. GaitGANv2:invariant gait feature extraction using generative adversarial networks [J]. Pattern Recognition,2019,87:179-189.
- [17] CHAO H, HE Y, ZHANG J, et al. GaitSet:regarding gait as a set for cross-view gait recognition [C]//Proceedings of the AAAI conference on artificial intelligence. Hawaii:American Association for Artificial Intelligence,2019:8126-8133.
- [18] TORRESANI L, KC L. Large margin component analysis [C]//Advances in neural information processing systems.

(下转第 113 页)