

多视角步态识别综述

王科俊¹ 丁欣楠¹ 邢向磊¹ 刘美辰¹

摘要 步态识别作为生物特征识别中的一种, 具有远距离、非接触和难以模仿等优点. 其中视角或行走方向的变化使提取的人体轮廓产生巨大差异, 是影响步态识别系统性能的最主要因素之一. 本文首先介绍了现有的多角度步态数据库, 然后根据特征提取方式的不同, 将当前已提出的方法分为三维模型法、视角不变性特征法、映射投影法和深度神经网络法四类, 并详细阐述了每一类的原理、特点以及优缺点. 最后, 结合实际应用指出当前研究的局限性与发展趋势.

关键词 步态识别, 多视角, 行为特征识别, 生物特征识别

引用格式 王科俊, 丁欣楠, 邢向磊, 刘美辰. 多视角步态识别综述. 自动化学报, 2019, 45(5): 841–852

DOI 10.16383/j.aas.2018.c170559

A Survey of Multi-view Gait Recognition

WANG Ke-Jun¹ DING Xin-Nan¹ XING Xiang-Lei¹ LIU Mei-Chen¹

Abstract As one of biometric identification methods, gait recognition has many advantages. It is suitable for human identification at a long distance, requiring no contact and hard to imitate. One of the most important factors that affect the performance of gait recognition is that the change of view or the direction of the walk, which makes the human silhouette have a large variation. We first introduce the existing multi-view gait database. Then according to the characteristics of different extraction methods, the current approaches are divided into four kinds: three-dimensional model, view-invariant features, subspace projection as well as deep neural network, and analyses are conducted on their principle, advantages and disadvantages. Finally the limitations of current research and development trend are pointed out.

Key words Gait recognition, multi-view, behavioral characteristics recognition, biometrics

Citation Wang Ke-Jun, Ding Xin-Nan, Xing Xiang-Lei, Liu Mei-Chen. A survey of multi-view gait recognition. *Acta Automatica Sinica*, 2019, 45(5): 841–852

步态识别是指通过人走路的姿态或足迹对身份进行认证或识别, 被认为是远距离身份识别中最具潜力的方法之一^[1]. 优势主要包括无需接触、非侵犯、识别过程不需要配合、难于隐藏和伪装等. 因此步态识别在安全监控、人机交互、医疗诊断和门禁系统等领域具有广泛的应用前景和经济价值.

步态识别在实际应用中面临许多难点, 主要表现在行人在行走过程中会受到外在环境和自身因素的影响^[2–3] (例如不同行走路面、不同时间、不同视角、不同服饰、不同携带物等因素), 导致提取到的步态特征呈现很强的类内变化. 其中视角因素是影响系统识别性能最主要的因素之一. 当行人行走方向

发生变化, 或由一个摄像监控区域转入另一个具有不同设置的摄像监控区域时都会发生视角变化. 图 1 为同一个人在不同视角下的步态图像, 可以观察到不同视角下的步态图像均具有较大差异. 研究^[4–6]普遍认为侧面视角的步态轮廓包含了更有价值的信息, 特征提取绝大多数也都是基于侧面轮廓的, 而传统的单视角步态识别技术在视角变化时, 识别性能也随之明显下降^[7–8].



图 1 不同视角下的步态图像 (CASIA-B)

Fig. 1 Gait images from different views (CASIA-B)

自 1994 年 Niyogi 等^[9] 最早利用步态信息作为特征进行身份认证后, 步态识别得到发展快速, 并涌现出大量的步态识别算法, 其中不乏相关的综述文

收稿日期 2017-09-29 录用日期 2018-04-04

Manuscript received September 29, 2017; accepted April 4, 2018

国家自然科学基金项目 (61573114), 中央高校基本科研基金项目 (HEUCF160415), 黑龙江省自然科学基金项目 (F2015033) 资助

Supported by National Natural Science Foundation of China (61573114), Fundamental Research Funds for the Central Universities (HEUCF160415), and Provincial Natural Science Foundation of Heilongjiang (F2015033)

本文责任编辑 王亮

Recommended by Associate Editor WANG Liang

1. 哈尔滨工程大学自动化学院 哈尔滨 150001

1. College of automation, Harbin Engineering University, Harbin 150001

章^[10-12], 但多是基于对步态识别的整体概述(相同视角下的步态周期检测与识别), 无针对解决视角这一主要难点对现有研究成果进行归纳总结. 为了弥补这个不足, 有必要对现阶段多视角步态识别研究情况进行总结分析, 以期对本领域研究人员有所裨益.

数据库对于学习角度因素对步态识别的影响、评估和性能比较是至关重要的. 本文首先总结可用于多视角识别的步态数据库. 然后对现有文献的研究方法进行综述. 根据特征提取的方式不同, 将当前已提出的多视角步态识别方法分为四类, 分别是 3D 模型法、视角不变性特征法、映射投影法和深度神经网络法. 最后指出当前研究的局限性和发展方向.

1 现有多视角步态数据库

进行身份识别算法研究、系统开发和评估必须要有共同的数据. 因此, 大型步态数据库是必不可少的. 为此, 步态数据库应该包含一个大的类别数以及各种协变量条件. 要消除角度因素对步态识别的影响, 需数据库提供针对各种复杂角度

设定的步态序列. 为了便于研究, 目前国际上已经建立了多个用步态识别研究的数据库, 其中具有角度变量, 可用于的多视角识别研究的数据库主要有 USF^[3]、CASIA-A^[13]、CASIA-B^[14]、HID-UMD1 和 2^[15]、CMU MoBo^[16]、OU-ISIR Treadmill^[17]、OU-ISIR LP^[18] 和 SZU RBG-D^[19] 等, 表 1 详细地总结了上述多角度步态数据集.

此外, 还有 UMST^[20]、KY4Ddata^[21]、AVAMVG^[22] 和 TUM-IITKGP^[23] 等 3D 步态数据库, 用于 3D 步态建模.

而现存的多视角步态数据库虽已系统地对各个行人不同视角的下的数据进行注册, 但仍存在以下不足:

- 1) 高维的步态特征和小样本问题: 除 OU-ISIR LP 数据库外, 步态数据库的注册样本都不足 200 人. 但通常情况下, 步态识别技术的特征维数很高, 而数据集中的训练样本数目很少, 一般的识别算法可能会造成数据过拟合.
- 2) 视角问题: 需应用步态识别的视频监控场景的摄像头通常安装是有一定俯角的, 然而现存数

表 1 多视角步态库
Table 1 Databases for multiview gait

库名	建立机构	样本容量	具体角度	示例
USF ^[3]	南佛罗里达大学	122 人, 1870 序列 (地面、鞋子、负重和时间)	在摄像机前 (左右两个) 绕椭圆路线行走	
CASIA-A ^[13]	中国科学院自动化研究所	20 人 × 3 视角 × 4 序列 = 240	3 个 (侧面的 3 个点)	
CASIA-B ^[14]	中国科学院自动化研究所	124 人 × 11 视角 × (8 正常 + 2 背包 + 2 外套) = 13 640	11 个 (180° 每间隔 18° 度一个视角)	
HID-UMD ^[15]	马里兰大学	1: 25 人 × 4 个视角 = 100	正面 (走向、走出)、侧面 (向左、向右)	
		2: 55 人 × 4 (2 视角) = 220	T 形路径 (正面、侧面)	
CMU MoBo ^[16]	卡耐基梅隆大学	25 人 × 6 视角 × (3 速度 + 1 抱球 + 1 上坡) = 600	6 个 (360° 圆周每 60° 一个视角)	
OU-ISIR Treadmill ^[17]	大阪大学	168 人	25 个 (12 个方位角 × 2 个倾斜角度 + 1 个俯视)	
OU-ISIR LP ^[18]	大阪大学	4 007 人	8 个 (2 摄像机 × 4 个侧面角度)	
SZU RBG-D ^[19]	深圳大学	99 人 × 2 视角 × 4 序列 = 792	90° 和 30° ~ 60° 之间的一个角度	

据库大多是人的行走方向与摄像机镜头主轴方向垂直,视角变量仅限制在行走平面的 360 度内,无立体视角变量的大型步态库。

3) 遮挡、服饰或携带物和夜间识别问题:实际的步态识别中,很可能存在障碍物和行人之间的相互遮挡,现存的步态数据库中,无论室内室外都是在空旷环境中且画面只有一个行人情况下采集的;同时现有的步态数据库服饰和携带物变量较少,通常不超过 5 种,且所有行人都采用同样的服饰和携带物进行注册;夜间红外摄像头采集的视频信息通常光线较暗,步态轮廓与背景较为接近,视角因素与夜间因素相互影响大大增加了识别难度,而目前仍没有夜间的多视角步态数据库。

2 研究动态

步态分析最早是医学^[24]、心理学方面的课题^[25],如果考虑步态运动的所有信息,每个人的步态都是唯一的。随着计算机运算能力的增强和生物特征识别技术的兴起,步态分析在计算机视觉领域的发展逐渐引起了关注。美国国防部高级研究项目署在 2000 年资助了远距离身份识别的重大研究项目^[1],研究远距离步态和动态人脸以及其他因素对身份识别的影响,最终开发远距离下具有高可靠性、鲁棒性的大规模身份识别系统,对步态识别的研究产生了深远的影响,在一定程度上促进了步态识别的发展。国内外许多知名的研究机构都开展了步态识别方面的研究,国外比较著名的有麻省理工学院、马里兰大学、南安普顿大学、南佛罗里达大学等。国内的研究机构主要有中国科学院自动化研究所、山东大学、哈尔滨工程大学、复旦大学、深圳大学等。但目前步态识别研究均是理论性的,尚没有成熟的步态识别系统出现。多数的步态识别算法是在理想环境下(背景简单,无遮挡、携带物和服饰变换且画面中只有一行人)对目标侧影图像的分析,与实际应用环境差别较大。因此,虽然现有的方法大多都已取得超过 80% 识别率^[26],但把步态用于个人身份识别还没有达到在实际复杂环境中应用的要求。

本文将现有的多视角步态识别研究方法分为基于三维建模和基于二维图像或视频序列特征两大类。其中,三维建模主要采用多台校准的不同视角的摄像机构建 3D 人体步态模型;而基于二维图像或视频序列特征的方法,构造能够有效整合步态视频轮廓与时域信息的步态特征模板直接影响步态识别的精度,故本文先介绍各常用类能量图的构造方法,再根据特征提取方式的不同,又分为提取视角不变性特征法,学习不同视角下映射投影关系法和深度神经网络法三类。

2.1 基于建立 3D 步态模型的方法

由于步态信息本质是三维的,而二维步态图像序列只提供单一视角信息,限制了任意视角的步态识别。通过多摄像机对人体结构或人体运动进行三维建模,能够更准确地表达人体各个部位的物理空间,充分利用关节的角度约束和人体各个部位的运动特性。且在 3D 空间中,步态识别的人体检测、人体跟踪等预处理工作变得比较简单,能够降低遮挡等因素的负面影响。

Shakhnarovich 等^[27]提出基于图像的可视外壳(Image-based visual hull, IBVH)以绘制用于步态识别的视角。IBVH 从多个校准摄像机的一系列单视角中计算得出。该方法首先估计规范视觉相机的位置,然后使用从这些视角获得的绘制图像来做视角规范化。Bodor 等^[28]应用基于图像的绘制技术于 3D 可视外壳模型,得以在任意所需视角下重构步态特征。该方法可以将不同视角下的多个摄像机所获取的步态信息综合起来,但需要在校准单个摄像机的基础上进行交叉校准使得它们具有相同的参考帧。Zhang 等^[29]提出基于 3D 角线性模型和贝叶斯规则的视角对立步态识别方法。该方法在傅立叶表示的样本中使用主成分分析来构建 3D 线性模型。通过最大后验概率估计将不同视角下的 2D 步态序列投影到一个 3D 模型中,由此得出一系列系数用来描述步态特征。Tang 等^[30]利用先进的 3D 成像设备进行 3D 重建和目标跟踪,但很难只用距离数据精确地分割出人体轮廓。为了解决这个问题,Tang 等^[31]提出建立 3D 人体模型,通过二维轮廓和姿态建立拉普拉斯形变能量函数将模型产生相应角度和姿势变形,再将局部投影至二维空间构建部分步态能量图再识别的方法。Zhao 等^[32]从多个摄像机捕获的视频序列中重构 3D 步态模型。该方法使用从 3D 模型中提取的下肢的运动轨迹作为动态特征,同时利用线性时间规整化进行匹配和识别。López-Fernández^[22]等同样利用多摄像机构成的多视点视频序列重构三维步态序列,提出了一种基于 3D 角度分析的旋转不变的特征,运用子空间分量与判别分析和支持向量机(Support vector machine, SVM)进行分类识别。Deng 等^[33-34]提出了一种基于多视点融合和确定性学习的方法,利用不同视角合成轮廓图像。Iwashita 等^[35]利用 4D 步态数据库合成虚拟图像估计行走方向,提取仿射不变矩为特征进行识别。

一般 3D 分析至少需要两台摄像机,通常因为存在遮挡,为了进行充分的 3D 步态分析,至少需要来自 4 台摄像机的步态信息。然而,由于复杂的摄像机平衡视角和建模计算,这一系列的方法一般只适

用于完全可控的多摄像机协作环境,且难以在实际的实时应用中使用。

2.2 基于二维图像或视频序列特征的方法

2.2.1 步态特征表征模板 (类能量图)

与基于三维建模的方法不同,基于二维图像的步态识别通常需构造步态表征模板集中整合步态的静态、动态和时序信息,一般是对视频图像序列按照一定规则的叠加,即构造步态类能量图。学习不同视角下映射投影关系方法的研究绝大多数都是在类能量图基础上进行二次特征提取完成的;也可从能量图中直接提取视角不变特征或将其送入深度神经网络。因此类能量图的构造与选择直接影响了步态识别精度。

类能量图法将一个人周期性的、连续的时空运动序列生成一幅或几幅图像,对步态图像质量要求不高,且无需考虑人体模型结构和计算人体各部分的精确参数,能够节省存储空间和计算代价,对图像噪声有较好的鲁棒性。同时基于序列的步态表征方法更有效地利用了步态序列的时空连续性,包含更多运动特征。Lv 等^[12]根据类能量图的生成方式将其分成步态信息累计类、步态信息引入类、步态信息融合类能量图三类,并全面地对已提出的各种类能量图的特性进行了理论分析和对比实验研究。其中较为常用的类能量图主要有运动历史图像 (Motion history image, MHI)^[36]、步态能量图 (Gait energy image, GEI)^[37]、运动轮廓图 (Gait energy image, MSI)^[38]、步态历史图 (Gait history image, GHI)^[39]、主动能量图 (Active energy image, AEI)^[40]、步态熵图 (Gait entropy image, GENI)^[41]、帧差能量图 (Frame difference energy image, FDEI)^[42]和基于时间保持的步态能量图 (Chrono-gait image, CGI)^[43]等。其构造方法与性能分析如表 2 所示。

2.2.2 提取视角不变性特征

直接提取不同视角下视频序列或图像中不随着角度或行走方向而改变的步态特征进行身份识别,从而避免因角度变化而引起的人体轮廓的巨大差异对识别的影响。直接提取视角不变性步态特征的方法各式,通常思路直观,计算较为简单,而提取的特征也较为多样。由于基于局部特征和聚类图像的方法较多,本文大体将其分为局部特征法、聚类图像法和其他三类。

1) 局部特征

Jean 等^[44]提出一种计算视频序列中人体部位轨迹的视角规范化方法。该方法选取轮廓序列中规范化脚和头的 2D 轨迹将行走轨迹分割成分段线形

的部分。然而,该技术仅对有限范围内的视角有效。为了解决 Jean 等的方法由于遮挡等原因,提取到的人体轮廓中可能丢失头部或脚的问题,Ng 等^[45]提出从自动检测到的人体关节 (臀、膝和脚踝) 中计算关节的角轨迹,并应用透视校正来提取视角不变的步态特征。彭彰等^[46]提出了一种基于肢体长度参数的方法,利用脚间距计算方法和动态身体分割方法,拟合出场景的转换参数,并以此估计出人运动情况下的 5 个肢体长度参数用于识别。这种方法受限于图像分割的准确度且只适用于视角与行人水平的情况,立体视角的变化对肢体长度参数影响较大。Goffredo 等^[47]提出基于模型的步态特征自标定视角不变步态识别。下肢姿态由无标记运动来估计,然后这些姿态在腿关节运动近似平坦的假设下,使用视点矫正在矢状面进行重构。但该方法对于下肢姿态的估计缺乏鲁棒性且无法应用于正面视角,两视角差异较小时性能较差。

2) 聚类图像法

Lu 等^[48]将不同的视角分成几个聚类,提出聚类的平均步态图像 (Average gait image, AGI) 作为特征表达,利用稀疏重构的度量学习 (Sparse reconstruction metric learning, SRML) 进行身份分类。Darwish^[49]采用聚类的空间域能量偏差图像 (Energy deviation image, EDI)^[50]作为步态特征,再利用区间二型模糊 K 近邻 (Interval type-2 fuzzy K-nearest neighbor, IT2FKNN) 进行步态识别。聚类图像法大多先利用聚类进行视角估计解决角度问题,需要大量的注册样本,且如果在步态序列中没有相似的视角,识别率将下降。

3) 其他

Han 等^[37]从 GEI 中提取视角不变特征。该方法仅选取部分在视角间相互重叠的步态序列来构建交叉视角的步态匹配表示。Kale 等^[51]提出一种从任意视角生成侧视图的方法。该方法采用透视投影模型和基于运动方程的光流结构。这种方法要求行人距离相机的距离足够远,当像平面和矢状面之间的夹角变大时,该方法受到自遮挡的影响性能显著下降。

直接提取视角不变特征的方法适用于视角变化范围有限或较小的情形,并且该类方法提取步态特征的过程易受到遮挡因素或服饰变化的破坏。

2.2.3 学习不同视角下的映射或投影关系

在步态相似性度量之前,训练好的映射关系模型可以将不同视角下的步态特征规范化到相同视角的特征空间中,利用多个角度的训练数据学习视角判别子空间,步态特征被投影到子空间 (通常在较低的维度) 中获得视角不变特征,以解决多角度

表 2 类能量图构造方法与性能分析

Table 2 The construction methods and performances analysis of class energy image

名称	公式	含义及性能	图像
运动历史图 像 MHI ^[36]	$E_{MHI}(x, y, n) = \begin{cases} \tau, & D(x, y, n) = 1 \\ \max(0, E_{MHI}(x, y, n-1) - 1), & \text{其他} \end{cases}$	为在每一像素与该位置的持续运动时间成正比. 描述时序信息, 静态信息较少. ($D(x, y, n)$ 表示运动区域的二值差分图像, 当图像像素点处有运动发生时为 1. τ 为序列帧数.)	
步态能量图 GEI ^[37]	$E_{GEI}(x, y) = \frac{1}{N} \sum_{n=1}^N B(x, y, n)$	人体运动归一化的一个周期图像能量的叠加. 像素的浓度代表在该像素位置频率, 即为能量. 包含动态信息和静态信息, 对于单一帧的偶尔错误不敏感, 但丢失了时序信息. ($B(x, y, n)$ 表示第 n 帧的二值化图像, N 为一个步态周期的帧数.)	
运动轮廓图 MSI ^[38]	$E_{MSI}(x, y) = \begin{cases} 255, & B(x, y, n) = 1 \\ \max(0, E_{MSI}(x, y, n-1) - 1), & \text{其他} \end{cases}$	受 MHI 启发而来, 像素的浓度表征着在该像素点上的运动历史. 相比于 MHI, 计算较为简单, 存储空间更少.	
步态历史图 GHI ^[39]	$E_{GHI}(x, y) = \begin{cases} P, & \bigcup_{n=1}^p B(x, y, n) = 1 \\ \sum_{n=1}^p D(x, y, n) \cdot (N-1), & \text{其他} \end{cases}$	p 是 1/4 步态周期的帧数, GHI 以 1/4 周期为处理对象, 记录了随时间变化的区域, 用不同的灰度值代表其变化的时序. 包含了动态和静态信息, 同时有 MHI 能够反映时序变化的特点, 但可能会丢失有用信息.	
主动能量图 AEI ^[40]	$E_{AEI} = \frac{1}{N} \sum_{n=1}^N D(x, y, n)$	由帧差图像叠加构成, 很好地反应运动部分的时序特征, 减轻图像质量低, 衣着和负重的影响, 静态信息较少.	
步态熵图 GEnI ^[41]	$E_{GEnI}(x, y) = -E_{GEI}(x, y) \log_2 E_{GEI}(x, y) - (1 - E_{GEI}(x, y)) \log_2 (1 - E_{GEI}(x, y))$	计算 GEI 的每个像素点香农熵区分动态和静态区域, 香农熵度量轮廓的不同位置的密度的不确定性. 有效解决对背包和衣着静态信息较敏感的问题.	
帧差能量图 FDEI ^[42]	$E_{FDEI}(x, y) = F(x, y, z) + E_{DEI}(x, y)$ $E_{DEI}(x, y) = \begin{cases} \frac{1}{N} \sum_{n \in N} B(x, y, n), & \frac{1}{N} \sum_{n \in N} B(x, y, n) \geq T \\ 0, & \text{其他} \end{cases}$ $F(x, y, n) = \begin{cases} 0, & B(x, y, n) \geq B(x, y, n-1) \\ B(x, y, n-1) - B(x, y, n), & \text{其他} \end{cases}$	由帧差图和去噪能量图相加得到, 包含了大部分形状和帧间变化信息. 对有残缺的步态图像序列有较好的鲁棒性.	
基于时间保持的步态能量图 CGI ^[43]	$r_t = (W_t - W_{\min}) / (W_{\max} - W_{\min})$ $B(r_t) = \begin{cases} (1 - 2r_t)I, & r_t \leq 1/2 \\ 0, & r_t > 1/2 \end{cases}$ $G(r_t) = \begin{cases} 2r_tI, & r_t \leq 1/2 \\ (2 - 2r_t)I, & r_t > 1/2 \end{cases}$ $R(r_t) = \begin{cases} 0, & r_t \leq 1/2 \\ (2r_t - 1)I, & r_t > 1/2 \end{cases}$ $CGI(x, y) = \frac{1}{p} \sum_{i=1}^p PGI_i(x, y)$	为了避免 GEI 的时序信息损失, 采用一个线性插值函数对 1/2 步态周期的时间信息编码, 并计算权重将每帧外轮廓投影至 RGB 空间累加得到 PGI. 一个序列的所有 PGI 图像叠加平均得到 CGI. (W_t 为第 t 帧的腿部平均宽度, $W_{\max}$ 和 $W_{\min}$ 为 1/2 周期内腿部平均宽度的最大值与最小值.)	

问题. 基于投影映射的方法是一种很好的实时应用解决方案, 已有的研究成果多具有较高的识别精度. 其中已采用的映射投影方法很多, 包括典范相关分析 (Canonical correlation analysis, CCA)、视角转换模型 (View transformation model, VTM)、线性判别分析 (Linear discriminant analysis, LDA)、多线性主成分分析 (Multilinear principal component analysis, MPCA)、核主成分分析 (Kernel principal component analysis, KPCA) 和耦合学习等. 由于 LDA、MPCA、耦合学习与核思想多交叉融合构造子空间, 将其归为一类, 故本文将从 CCA、VTM 和其他三类分别说明.

1) 典范相关分析 (CCA)

典范相关分析 (CCA)^[52] 是一种著名的多元分析方法, 其目的是寻找和量化两个多维变量之间的相关性. CCA 利用两种相同模式的视图, 并将它们投射到一个使其相关性最大的低维空间中. 应用 CCA 解决跨视角步态识别问题, 将不同视角的步态特征投影到一个统一的特征空间, 并在该公共空间中进行相似性度量.

Bashir 等^[53] 采用高斯过程分类框架进行视角估计, 再利 CCA 进行不同角度建模. Hu^[54] 将判别典范相关分析 (Discriminant canonical correlation analysis, DCCA) 进行了高阶张量扩展, 应用多重线性分析, 利用张量到向量投影直接从张量数据中提取不相关判别特征, 增进了 DCCA 方法在多视角步态识别中的性能. Xing 等^[55] 针对传统 CCA 方法在处理两个高维数据集时, 存在的广义特征分析的奇异矩阵问题以及解的不稳健和不完备性, 提出完备典范相关分析方法 (Complete canonical correlation analysis, C3A), 并将其应用于多视角步态识别, 提升了传统 CCA 在步态识别中的性能. Wang 等^[56] 对通过改进优化目标函数和类关系矩阵对原有的核判别典型相关分析进行优化, 在 CCA 中引入了类信息并减少对应元素的相关性, 在跨视角识别取得了较好的识别效果. Luo 等^[57] 将 GEI 分割成 5 个子部分进行 CCA, 并在不同携带物和行走条件下优化训练每组子 GEI, 减少携带物和行走条件等变量因素的影响.

然而, 基于 CCA 类的方法仅能利用两个视角间的互补信息, 处理 N 个视角时要重复 N 次来学习 N 对特征映射, 计算负担沉重.

2) 视角转换模型 (VTM)

视角转化模型可以将不同视角下的步态特征转化到相同的视角下, 解决多角度的步态识别问题. Makihara 等^[58] 提出视角转化模型的概念, 由采用奇异值分解的矩阵分解过程创建. 训练数据集中

的步态矩阵每行包含来自相同视角不同对象的步态信息, 每列包含来自相同对象, 不同视角的步态信息, 应用奇异值分解 (Singular value decomposition, SVD) 将步态矩阵分解成视角独立的矩阵和对象独立的矩阵, 对象独立矩阵用于构建 VTM. 该方法基于傅立叶变换获得的频域步态特征来创建 VTM. 为了增进性能, Kusakunniran 等^[59-60] 基于线性判别分析所获取的最优 GEI 特征来创建 VTM, 应用截断奇异值分解 (Truncated singular value decomposition, TSVD) 来缓解训练 VTM 时的过拟合现象. 进一步将 VTM 的构建重新表述成回归问题. 使用回归概念来揭示不同视角间步态的运动相关性. Zheng 等^[61] 通过首先对步态矩阵进行低秩分解再应用 SVD 分解构造 VTM 的方法来实现鲁棒的 VTM 模型. Hu 等^[62] 应用高阶奇异值分解将 VTM 模型扩展为在四阶张量空间的多重线性投影模型, 然后提取视角独立、站姿独立的单位矢量, 以对多视角、不完整步态周期的步态序列进行识别. Muramatsu 等^[63-65] 在 VTM 中引入一种质量评价措施. 由于一般是通过 VTM 视角转换至规定视角的步态图像与目标视角图像进行相似性度量评判是否为同一身份, 步态特征组引入了相似度的不均匀偏差. 通过引入转化质量 (源视角内在特征质量) 和不均匀边缘质量 (目标视角内在特征质量), 利用这两个质量度量来计算真实配对的后验概率和原始特征的相似度量结合作为最终的匹配结果以提高识别精度.

VTM 类方法虽然可以将一个视角下的步态特征转化为另一视角下的步态特征, 从而解决不同视角之间的相似性度量问题, 但无法有效地同时利用多个视角之间的互补信息. 且基于 VTM 的方法都有在进行模型构建和视角转化时容易造成噪声传播, 致使识别性能退化的问题.

3) 其他 (LDA、耦合学习、MPCA 与核扩展等)

LDA 是一种监督学习的降维技术, 投影后类内方差最小, 类间方差最大, 提取需要的判别信息并减少维度. 通常步态特征空间的维数非常高, 且有许多在相似度测量中冗余的零值的像素, 在经过 LDA 投影后减小同一行人不同视角下的类内方差, 提高识别精度. Choudhury 等^[66] 提出一种首先分割得到腿部步态能量图用于估计视角, 再利用随机子空间学习进行身份分类的方法. Liu 等^[67] 在每个视角下的训练数据集中提取步态特征在 LDA 子空间中的判别信息. 在测试阶段, 每个步态特征分别投影到每个子空间中, 然后最终的步态距离由每个子空间匹配结果的加权和组成. Liu 等^[68] 提出一种联合子空间学习的方法 (Joint subspace learning, JSL), 构

造含有不同视角原型的 JSL, 不同视角的注册和待测样本分别被表示为这些原型在相应视角中的线性组合, 并提取特征表示的系数利用最近邻识别分类。

核方法^[69]是对 SVM 中主要思想核映射的应用的扩展, 很多线性子空间算法都可以运用核函数扩展为非线性子空间, 如核主成分分析^[70]和核判别分析^[71]等。Connie 等^[72]以双核主成分分析进行系数膨胀建立非线性子空间, 形成 Grassman 流形描述多角度的步态特征, 非线性子空间的结构能够更恰当地在视角变化中保留步态特征。

耦合度量学习 (Coupled metric learning, CML) 受启发于局部保留投影, 旨在通过寻找保留局部信息的低维嵌入获得一个可以保持数据内部流形结构的子空间。与局部保留投影不同, 耦合距离度量学习寻找一对线性变换矩阵将不同的样本映射到共同的子空间, 在这个子空间中不同视角的步态数据差异被削减。Xu 等^[73]提出一种耦合局部保留投影的方法, 学习耦合投影矩阵, 在保证基本流形结构的同时, 将交叉视角特征投影到统一的子空间中。Ben 等^[74]基于耦合距离度量学习思想, 通过广义特征值分解将不同视角下的步态特征联系起来, 提高了跨域生物特征的识别率, 在跨视角步态识别中也取得了良好的效果。在此基础上, Ben 等^[75]引入了核的思想, 通过核耦合距离度量学习 (Kernel coupled distance metric learning, KCDML) 使不同类样本线性可分, 并保持样本局部结构数据的几何特征。Wang 等^[76]在传统 CML 中引入可分离标准, 将标签信息加入其中, 应用这种新的方法将步态特征从不同视角转换为一个耦合特征空间。但基于耦合度量学习的方法同 CCA 方法有相同缺点, 仅能利用两个视角间的互补信息。

MPCA 是一种无监督多线性子空间学习方法, 实现了由张量至张量的投影技术, 将高阶张量对象投影到下维张量中, 即直接在张量维度进行降维。Al-Tayyan 等^[77]提出了一种基于累计预测图像的方法, 同时定义累积流量图像和边缘掩蔽活性能量图像两种新的步态表达方式, 采用 MPCA 与 LDA 结合通过 K 近邻进行身份分类, 以最大化类间散射矩阵与类内散射矩阵之间的比例, 提高分类的准确性。

子空间学习的方法通常计算较为复杂, 步态图像转换成向量后维数常常高达上万维, 计算量很大。此外, 在视角的变化较大时, 这一类方法性能下降较大。

2.2.4 基于深度神经网络

深度学习^[78]是含多隐层的多层感知器的深度神经网络结构, 通过组合低层特征形成更加抽象的

高层表示属性类别或特征, 以发现数据的分布式特征表示。因其在语音识别、图像目标分类等实际应用中的出色性能近年来备受重视。但将深度学习应用于解决多视角步态识别问题的研究相对较少。

卷积神经网络 (Convolutional neural network, CNN)^[79]近年发展迅速, 并引起广泛重视。其避免了对图像的复杂前期预处理, 可以直接输入原始图像, 对于大型图像处理有出色表现。且具有非常强的自主学习能力和高度非线性映射的能力, 能够学习非线性度量函数以解决跨视角步态识别问题。Yan 等^[80]提出将步态能量图送入 CNN 来提取高级步态特征, 并引入多任务学习模型, 在步态识别的同时预测状态, 联合每个任务的损失函数进行反向传播, 获得比单独步态识别更好的性能。Wu^[81]提出一种对一组图片集进行特征提取的方法, 有效抑制过拟合问题。将步态轮廓集送入相同的 CNN 网络中, 积累这些特征以获得集合的全局表示, 能够较好地应对步态中的角度变换。Zhang 等^[82]提出将 GEI 送入具有两个卷积子网络的对称结构的孪生深度卷积神经网络。Tan 等^[83]提出将 GEI 送入使用共享权重的双通道卷积神经网络训练匹配模型, 从而匹配步态识别人的身份, 对跨较大视角的步态变化有很强的鲁棒性。Wu 等^[84]提出深度 CNN 网络用于步态识别, 网络分为局部特征匹配的底层网络、中级特征匹配和全局特征匹配的上层网络, 在跨视角和多状态识别中都有较好表现。Wolf 等^[85]提出了一种用于步态识别的 3D CNN 方法, 网络的输入由灰度步态图像和光流组成, 能够在多个角度下提取步态的时空特征。Li 等^[86]提出一种基于深度学习 VGG 网络^[87]的识别方法。步态序列经过周期检测后直接送入 VGG 网络进行特征提取, 最后利用联合贝叶斯进行步态识别。

自动编码器 (AutoEncoder)^[78]是近年来流行的网络模型, 它可以用来提取紧凑的特征。Yu 等^[88]提出基于 GEI 进行的层叠式逐步自动编码 (Stacked progressive auto-encoders, SPAE) 的完成视角转化, 每层转化 18° 的视角。即利用自动编码器搭建 VTM 模型, 来解决跨视角步态识别问题。

深度学习本质是数据驱动的, 需要大量的不同行人的不同角度的步态信息, 而现有的数据量相对较少。且现有的方法大多基于步态能量图和卷积神经网络的。然而步态能量图在轮廓序列的周期叠加后会丢失时序信息。卷积神经网络本身无法直接处理时间序列信号, 也缺乏对时间序列信号的记忆功能。而基于自动编码的视角转换也有在转化时容易造成噪声传播使识别性能退化的问题。

3 对比与总结

3.1 实验结果对比

为了对现有方法的性能进行直观比较,表3选取了在CASIA-B数据集上验证过的若干方法的实验结果作为对比.由于跨视角情况下的步态识别暂没有统一的性能评价标准,研究中通常进行多组实验,选择每个行人不同的一个或若干个角度为注册样本,验证其他视角下的识别率以全面评估性能(即选取不同的Gallery和Probe set进行验证).表3中仅仅展示以90°为待检测样本(Probe)的情况下,不同方法分别在54~126°为已知参考样本(Gallery)的识别准确率.

实验中都只采用正常行走状态下的行人步态样本(即无携带物和着装变化的视频样本).而在训练与测试集的划分上,不少工作采用在库中124行人中选择部分行人的样本作为训练,剩余行人作为测试集的验证方式,也有部分工作选择了其他的划分方式.3D模型局部能量图投影和KCDML的验证方式分别为选取数据库中所有行人的6个正常步态视频中的3个和2个作为训练,剩余视频作为测试集;其余方法均使用上述主要验证方式,其中GEI+CCA训练和测试集行人样本数量划分为74/50,GEI+SPAE为62/62,剩余方法为24/100.

3.2 现有方法总结

早期的多视角步态识别方法主要是提取视角不

变的步态特征^[37,44–51].随着三维建模与多摄像机协作技术的发展,建立3D步态模型^[22,27–35]也很好解决了识别中的视角问题;同时基于度量学习的方法也被用于多视角的步态识别中^[53–77],通过投影到子空间获取视角不变特征,取得了相对较高的识别精度.而深度学习利用深层神经网络学习出高层抽象的步态特征,在步态识别的视角变化中也取得了良好的识别效果^[80–88].表4比较了现有的多视角步态识别算法.

表3 CASIA-B数据集上现有步态识别方法的准确率对比
Table 3 Recognition accuracy of existing approaches on CASIA-B datasets

方法	识别率 (%)				
	54°	72°	90°	108°	126°
3D模型局部能量图投影 ^[31]	86.1	98.1	–	92.3	90.4
下肢姿态重构 ^[47]	74.4	72.6	86.5	69.2	67.4
聚类AGI+SRML ^[48]	68.3	92.7	–	93.4	70.3
聚类EDI+IT2FKNN ^[49]	72.2	94.8	–	95.3	74.9
GEI+CCA ^[53]	52.0	94.6	–	93.8	78.3
C3A ^[55]	86.6	98.7	–	95.4	86.0
GEI+TSVD ^[59]	53.7	81.8	–	86.3	45.7
低秩GEI+SVD ^[61]	42.5	86.2	–	88.3	49.8
JSL ^[68]	71.4	91.7	–	90.5	73.6
KCDML ^[75]	76.7	94.3	–	88.1	88.4
GEI+SPAE ^[88]	82.3	94.4	96.0	96.0	86.3

表4 现有多视角步态识别方法
Table 4 Existing approaches for multiview gait recognition

模板	方法	实现途径	优点	缺点
三维模板	建立三维步态模型	利用多摄像机对人体结构或人体运动进行3D建模 ^[22,27–35]	更准确地表达人体各个部位的物理空间;能够降低遮挡等因素的负面影响.	需要全可控的多摄像机协作的识别环境;摄像机平衡视角和建模计算复杂.
	提取视角不变性特征	提取局部特征 ^[44–47] ;聚类图像估计视角 ^[48,49] ;其他 ^[38,51] .	直接提取不随着角度而改变的步态特征进行身份识别,避免轮廓差异;思路直观,计算简单,易于理解和实现.	仅仅适用于视角变化有限的情况下;易受到遮挡因素或服饰变化的破坏.
二维图像或视频序列特征	学习不同视角下的映射或投影关系	典范相关分析(CCA) ^[53–57] ;视角转换模型(VTM) ^[58–65] ;其他(LDA、CML、MPCA与核扩展等) ^[66–77] .	投影到子空间中获得步态的视角不变特征,减小同一行人不同视角下的类内方差;具有相对较高的识别精度.	步态图像转换成向量后维数常高达上万维,计算量很大;视角变化较大时效果不理想;CCA和CML类的方法仅能利用两个视角间的互补信息,处理N个视角时要重复N次来学习N对特征映射;VTM方法在进行模型构建和视角转化时容易造成噪声传播,致使识别性能退化.
	基于深度神经网络	CNN ^[80–86] 与AutoEncoder ^[88]	无图像的预处理过程;有效提取步态特征;具有相对较高的识别精度.	需要大量数据用于训练;卷积神经网络缺乏对时间序列信号的记忆功能;基于自动编码的VTM也有在转化时容易造成噪声传播使识别性能退化的问题.

4 研究难点与发展趋势

4.1 研究难点

当前的研究难点主要集中以下在三个方面:

1) 与指纹、人脸识别等相比, 步态数据库的样本量过小, 且通常的实际生活中摄像机的安装位置为俯视视角, 但当前研究主要集中在行人行走平面的多视角识别, 即缺乏模拟真实环境下立体视角的具有大量样本的大型步态数据库. 同时现有的数据库都是在行人已知的情况下采集步态信息的, 注册行人在面对复杂的采集环境时可能产生不自觉的姿态变化.

2) 实际行走过程中会受到行走路面、不同时间、不同视角、不同服饰和不同携带物等的多种因素综合影响, 目前的研究多着重于解决视角问题, 但在视角与其他影响因素结合的复杂真实环境中的识别率仍然较低.

3) 当前的基于图像或视频序列的多视角步态识别多是利用步态序列叠加合成图像构造类能量图模板. 而图像的合成过程中可能会丢失信息, 且因涉及周期检测等问题, 此过程中可能已经引入了误差, 影响识别率.

4.2 发展趋势

1) 现有的数据库与真实环境中步态识别差距较大且样本数量较少. 一致的评估和性能比较需要构造一个大型的、涵盖各种变量、环境因素并适合实际应用的数据库. 同时构造隐藏摄像机的步态采集环境, 获取在行人未知状态下的步态信息也是未来研究中数据库建设的一个发展方向.

2) 深度学习用更多的数据或是更好的算法来提高学习算法的结果. 对某些应用而言, 深度学习在大数据集上的表现比其他机器学习方法都要好. 其中长短期记忆网络 (Long short term memory networks, LSTM)^[89] 对时序信息的处理能力为实现步态视频帧序列直接输入提供了可能性; 深度学习的实现需要大数据集作为依托, 而实际中很难采集同一行人遍历所有不同变量的大量样本, 生成对抗网络 (Generative adversarial network, GAN)^[90] 可以通过生成的方式生成多角度、多状态的大量不同步态样本用于深度学习的训练中.

3) 深度学习可以依靠深层的网络结果自动提取特征, 但过深的结构不利于参数训练, 同时导致信息不断稀释. 在深度学习中引入传统的步态特征提取方法, 并将两者结合起来或许能取得更好的步态识别效果.

4) 人体骨骼关键点检测技术^[91] 能够实时的抽象出人体的比例结构与姿态信息, 而步态识别本质

也是通过行人的轮廓和运动姿态信息进行识别. 可考虑将提取的人体骨骼步态模板用于步态识别中, 此种方法还可以避免服饰、携带物和遮挡带来的影响.

5) 研究中步态识别视频数据的形式往往是采用已分割好的个体步态视频流. 实现个体步态视频流的自动截取, 是未来实现端对端的步态识别系统的实际应用的关键步骤.

6) 步态识别仅仅利用步态信息进行身份认证, 但每种生物特征识别都有相应的适用场合, 以及各自的优缺点. 故开发多模态系统, 代替现有的使用单一特征的生物特征识别, 使其能在各种环境下都能提供有效的身份认证与识别, 有至关重要的意义.

5 结语

步态作为生物特征识别领域的一个新的研究方向, 多视角识别对其应用具有极大的实际意义, 近年来也引起了广大科研工作者的广泛关注. 针对步态识别中的视角问题, 本文首先介绍了现有的可用于多视角步态识别的数据库, 然后分别从 3D 模型法、视角不变特征法、映射或投影法和深度神经网络法 4 个方面对现有研究成果进行综述, 阐述了各种方法的原理和优缺点; 同时结合步态识别实际应用的需求, 针对现有工作中存在的不足, 提出一些有待深入研究的问题并指明未来的研究方向. 这些问题的解决将促使步态识别具有更宽泛的识别条件、更好的实时性与识别率, 从而将步态识别真正用于实际的远距离身份判断中.

References

- 1 Phillips P J. Human identification technical challenges. In: Proceedings of the 2002 International Conference on Image Processing. New York, USA: IEEE, 2002. I-49-I-52
- 2 Tariq M, Shah M A. Review of model-free gait recognition in biometric systems. In: Proceedings of the 23rd International Conference on Automation and Computing. Huddersfield, UK: IEEE, 2017. 1-7
- 3 Sarkar S, Phillips P J, Liu Z, Vega, I R, Grother P, Bowyer K W. The humanID gait challenge problem: data sets, performance, and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(2): 162-177
- 4 Liu Z, Sarkar S. Effect of silhouette quality on hard problems in gait recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2005, **35**(2): 170-183
- 5 Kale A, Sundaresan A, Rajagopalan A N, Cuntoor N P, Roy-Chowdhury A K, Kruger V, et al. Identification of humans using gait. *IEEE Transactions on Image Processing*, 2004, **13**(9): 1163-1173
- 6 Masood H, Farooq H. A proposed framework for vision based gait biometric system against spoofing attacks. In: Proceedings of the 2017 International Conference on Communication, Computing and Digital Systems. Islamabad, Pakistan: IEEE, 2017. 357-362

- 7 Arseneau S, Cooperstock J R. Real-time image segmentation for action recognition. In: Proceedings of the 1999 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing. British Columbia, Canada: IEEE, 1999. 86–89
- 8 Verlekar T T, Correia P L, Soares L D. View-invariant gait recognition system using a gait energy image decomposition method. *IET Biometrics*, 2017, **6**(4): 299–306
- 9 Niyogi S A, Adelson E H. Analyzing gait with spatiotemporal surfaces. In: Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects. Texas, USA: IEEE, 1994. 64–69
- 10 Wang Ke-Jun, Hou Ben-Bo. A survey of gait recognition. *Journal of Image and Graphics*, 2007, **12**(7): 1152–1160 (王科俊, 侯本博. 步态识别综述. 中国图象图形学报, 2007, **12**(7): 1152–1160)
- 11 Sugandhi K, Wahid F F, Raju G. Feature extraction methods for human gait recognition — a survey. In: Proceedings of the 2017 Advances in Computing and Data Sciences. Communications in Computer and Information Science, vol. 721. Singapore: Springer, 2017. 377–385
- 12 Lv Z W, Xing X L, Wang K J, Guan D H. Class energy image analysis for video sensor-based gait recognition: a review. *Sensors*, 2015, **15**(1): 932–964
- 13 Wang L, Tan T N, Ning H Z, Hu W M. Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(12): 1505–1518
- 14 Yu S Q, Tan D L, Tan T N. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: Proceedings of 18th International Conference on Pattern Recognition. Hong Kong, China: IEEE, 2006. 441–444
- 15 Chalidabhongse T, Kruger V, Chellappa R. The UMD Database for Human Identification at a Distance, Technical Report, University of Maryland, USA, 2001
- 16 Gross R, Shi J B. The CMU Motion of Body (MoBo) Database, Technical Report CMU-RI-TR-01-18, Carnegie Mellon University, USA, 2001
- 17 Makihara Y, Mannami H, Yagi Y. Gait analysis of gender and age using a large-scale multi-view gait database computer vision. In: Proceedings of Asian Conference on Computer Vision. Queenstown, New Zealand: ACM, 2010. 440–451
- 18 Iwama H, Okumura M, Makihara Y, Yagi Y. The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. *IEEE Transactions on Information Forensics and Security*, 2012, **7**(5): 1511–1521
- 19 Yu S Q, Wang Q, Huang Y Z. A large RGB-D gait dataset and the baseline algorithm. In: Biometric Recognition. Lecture Notes in Computer Science, vol. 8232. Cham: Springer, 2013. 417–424
- 20 Seely R D, Samangoeei S, Lee M, Carter J N, Nixon M S. The University of Southampton Multi-Biometric Tunnel and introducing a novel 3D gait dataset. In: Proceedings of IEEE 2nd International Conference on Biometrics: Theory, Applications and Systems. Virginia USA: IEEE, 2008. 1–6
- 21 Anguelov D, Koller D, Pang H C, Srinivasan P, Thrun S. Recovering articulated object models from 3D range data. In: Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence. Banff, Canada: ACM, 2004. 18–26
- 22 López-Fernández D, Madrid-Cuevas F J, Carmona-Poyato A, Muñoz-Salinas R, Medina-Carnicer R. A new approach for multi-view gait recognition on unconstrained paths. *Journal of Visual Communication and Image Representation*, 2016, **38**: 396–406
- 23 Hofmann M, Sural S, Rigoll G. Gait recognition in the presence of occlusion: a new dataset and baseline algorithms. In: Proceedings of the 19th International Conference in Central Europe Computer Graphics, Visualization and Computer Vision. Plzen, Czech Republic: Václav Skala UNION Agency, 2011. 99–104
- 24 Cho C W, Chao W H, Lin S H, Chen Y Y. A vision-based analysis system for gait recognition in patients with Parkinson's disease. *Expert Systems with Applications*, 2009, **36**(3): 7033–7039
- 25 Stevenage S V, Nixon M S, Vince K. Visual analysis of gait as a cue to identity. *Applied Cognitive Psychology*, 1999, **13**(6): 513–526
- 26 Lai D T H, Begg R K, Palaniswami M. Computational intelligence in gait research: a perspective on current applications and future challenges. *IEEE Transactions on Information Technology in Biomedicine*, 2009, **13**(5): 687–702
- 27 Shakhnarovich G, Lee L, Darrell T. Integrated face and gait recognition from multiple views. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE, 2001. 439–446
- 28 Bodor R, Drenner A, Fehr D, Masoud O, Papanikolopoulos N. View-independent human motion classification using image-based reconstruction. *Image and Vision Computing*, 2009, **27**(8): 1194–1206
- 29 Zhang Z H, Troje N F. View-independent person identification from human gait. *Neurocomputing*, 2005, **69**(1–3): 250–256
- 30 Tang J, Luo J, Tjahjadi T, Gao Y. 2.5D multi-view gait recognition based on point cloud registration. *Sensors*, 2014, **14**(4): 6124–6143
- 31 Tang J, Luo J, Tjahjadi T, Guo F. Robust arbitrary-view gait recognition based on 3D partial similarity matching. *IEEE Transactions on Image Processing*, 2017, **26**(1): 7–22
- 32 Zhao G Y, Liu G Y, Li H, Pietikinen M. 3D gait recognition using multiple cameras. In: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition. Southampton, UK: IEEE, 2006. 529–534
- 33 Deng M Q, Wang C, Chen Q F. Human gait recognition based on deterministic learning through multiple views fusion. *Pattern Recognition Letters*, 2016, **78**: 56–63
- 34 Deng M Q, Wang C, Cheng F J, Zeng W. Fusion of spatial-temporal and kinematic features for gait recognition with deterministic learning. *Pattern Recognition*, 2017, **67**: 186–200
- 35 Iwashita Y, Ogawara K, Kurazume R. Identification of people walking along curved trajectories. *Pattern Recognition Letters*, 2014, **48**: 60–69
- 36 Bobick A F, Davis J W. The recognition of human movement using temporal templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(3): 257–267
- 37 Han J, Bhanu B. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(2): 316–323
- 38 Lam T H W, Lee R S T. A new representation for human gait recognition: motion silhouettes image (MSI). In: Advances in Biometrics. Lecture Notes in Computer Science, vol. 3832. Berlin, Heidelberg: Springer-Verlag, 2005. 612–618

- 39 Liu J Y, Zheng N N. Gait history image: a novel temporal template for gait recognition. In: Proceedings of the 2007 IEEE International Conference on Multimedia and Expo. Beijing, China: IEEE, 2007. 663–666
- 40 Zhang E H, Zhao Y W, Xiong W. Active energy image plus 2DLPP for gait recognition. *Signal Processing*, 2010, **90**(7): 2295–2302
- 41 Bashir K, Xiang T, Gong S G. Gait recognition without subject cooperation. *Pattern Recognition Letters*, 2010, **31**(13): 2052–2060
- 42 Chen C H, Liang J M, Zhu X C. Gait recognition based on improved dynamic Bayesian networks. *Pattern Recognition*, 2011, **44**(4): 988–995
- 43 Wang C, Zhang J P, Wang L, Pu J, Yuan X R. Human identification using temporal information preserving gait template. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(11): 2164–2176
- 44 Jean F, Albu A B, Bergevin R. Towards view-invariant gait modeling: computing view-normalized body part trajectories. *Pattern Recognition*, 2009, **42**(11): 2936–2949
- 45 Ng H, Tan W H, Abdullah J, Tong H L. Development of vision based multiview gait recognition system with MMU-Gait database. *The Scientific World Journal*, 2014, **2014**: Article ID 376569
- 46 Peng Zhang, Wu Xiao-Juan, Yang Jun. A multi-view method for gait recognition based on the length of body's parts. *Acta Automatica Sinica*, 2007, **33**(2): 210–213 (彭彰, 吴晓娟, 杨军. 基于肢体长度参数的多视角步态识别算法. *自动化学报*, 2007, **33**(2): 210–213)
- 47 Goffredo M, Bouchrika I, Carter J N, Nixon M S. Self-calibrating view-invariant gait biometrics. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2010, **40**(4): 997–1008
- 48 Lu J W, Wang G, Moulin P. Human identity and gender recognition from gait sequences with arbitrary walking directions. *IEEE Transactions on Information Forensics and Security*, 2014, **9**(1): 51–61
- 49 Darwish S M. Design of adaptive biometric gait recognition algorithm with free walking directions. *IET Biometrics*, 2017, **6**(2): 53–60
- 50 Ma Q Y, Wang S K, Nie D D, Qiu J F. Gait recognition at a distance based on energy deviation image. In: Proceedings of the 1st International Conference on Bioinformatics and Biomedical Engineering. Wuhan, China: IEEE, 2007. 621–624
- 51 Kale A, Chowdhury A K R, Chellappa R. Towards a view invariant gait recognition algorithm. In: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance. Miami, Florida, USA: IEEE, 2003. 143–150
- 52 Hotelling H. Relations between two sets of variates. *Biometrika*, 1936, **28**(3–4): 321–377
- 53 Bashir K, Xiang T, Gong S G. Cross-view gait recognition using correlation strength. In: Proceedings of the British Machine Vision Conference. Aberystwyth, UK: BMVA Press, 2010. 109.1–109.11
- 54 Hu H F. Multiview gait recognition based on patch distribution features and uncorrelated multilinear sparse local discriminant canonical correlation analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, 2014, **24**(4): 617–630
- 55 Xing X L, Wang K J, Yan T, Lv Z W. Complete canonical correlation analysis with application to multi-view gait recognition. *Pattern Recognition*, 2016, **50**: 107–117
- 56 Wang K J, Yan T. An improved kernelized discriminative canonical correlation analysis and its application to gait recognition. In: Proceedings of the 10th World Congress on Intelligent Control and Automation. Beijing, China: IEEE, 2012. 4869–4874
- 57 Luo C, Xu W J, Zhu C Y. Robust gait recognition based on partitioning and canonical correlation analysis. In: Proceedings of the 2015 IEEE International Conference on Imaging Systems and Techniques. Macau, China: IEEE, 2015. 1–5
- 58 Makihara Y, Sagawa R, Mukaigawa Y, Echigo T, Yagi Y. Gait recognition using a view transformation model in the frequency domain. In: Proceedings of European Conference on Computer Vision. Graz, Austria: Springer-Verlag, 2006. 151–163
- 59 Kusakunniran W, Wu Q, Li H D, Zhang J. Multiple views gait recognition using view transformation model based on optimized gait energy image. In: Proceedings of the 12th International Conference on Computer Vision Workshops. Kyoto, Japan: IEEE, 2009. 1058–1064
- 60 Kusakunniran W, Wu Q, Zhang J, Li H D. Gait recognition under various viewing angles based on correlated motion regression. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, **22**(6): 966–980
- 61 Zheng S, Zhang J G, Huang K Q, He R, Tan T. Robust view transformation model for gait recognition. In: Proceedings of the 18th International Conference on Image Processing. Brussels, Belgium: IEEE, 2011. 2073–2076
- 62 Hu M D, Wang Y H, Zhang Z X. Cross-view gait recognition with short probe sequences: from view transformation model to view-independent stance-independent identity vector. *International Journal of Pattern Recognition and Artificial Intelligence*, 2013, **27**(6): 1350017
- 63 Muramatsu D, Shiraishi A, Makihara Y, Uddin M Z, Yagi Y. Gait-based person recognition using arbitrary view transformation model. *IEEE Transactions on Image Processing*, 2015, **24**(1): 140–54
- 64 Muramatsu D, Makihara Y, Yagi Y. Cross-view gait recognition by fusion of multiple transformation consistency measures. *IET Biometrics*, 2015, **4**(2): 62–73
- 65 Muramatsu D, Makihara Y, Yagi Y. View transformation model incorporating quality measures for cross-view gait recognition. *IEEE Transactions on Cybernetics*, 2016, **46**(7): 1602–1615
- 66 Choudhury S D, Tjahjadi T. Robust view-invariant multiscale gait recognition. *Pattern Recognition*, 2015, **48**(3): 798–811
- 67 Liu N, Tan Y P. View invariant gait recognition. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. Texas USA: IEEE, 2010. 1410–1413
- 68 Liu N N, Lu J W, Tan Y P. Joint subspace learning for view-invariant gait recognition. *IEEE Signal Processing Letters*, 2011, **18**(7): 431–434
- 69 Shawe-Taylor J, Cristianini N. Properties of kernels. *Kernel Methods for Pattern Analysis*. Cambridge: Cambridge University Press, 2004. 47–84
- 70 Alzate C, Suykens J A K. Multiway spectral clustering with out-of-sample extensions through weighted kernel PCA. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(2): 335–347
- 71 Yang J, Frangi A F, Yang J Y, Zhang D, Jin Z. KPCA Plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(2): 230–244

- 72 Connie T, Goh K O M, Teoh A B J. Multi-view gait recognition using a doubly-kernel approach on the Grassmann manifold. *Neurocomputing*, 2016, **216**: 534–542
- 73 Xu W J, Luo C, Ji A M, Zhu C Y. Coupled locality preserving projections for cross-view gait recognition. *Neurocomputing*, 2017, **224**: 37–44
- 74 Ben X Y, Meng W X, Yan R, Wang K J. An improved biometrics technique based on metric learning approach. *Neurocomputing*, 2012, **97**(1): 44–51
- 75 Ben X Y, Meng W X, Yan R, Wang K J. Kernel coupled distance metric learning for gait recognition and face recognition. *Neurocomputing*, 2013, **120**: 577–589
- 76 Wang K J, Xing X L, Yan T, Lv Z W. Couple metric learning based on separable criteria with its application in cross-view gait recognition. In: *Proceedings of the Biometric Recognition. Lecture Notes in Computer Science*, vol. 8833. Shenyang, China: Springer, 2014. 347–356
- 77 Al-Tayyan A, Assaleh K, Shanableh T. Decision-level fusion for single-view gait recognition with various carrying and clothing conditions. *Image and Vision Computing*, 2017, **61**: 54–69
- 78 Chen Wei-Hong, An Ji-Yao, Li Ren-Fa, Li Wan-Li. Review on deep-learning-based cognitive computing. *Acta Automatica Sinica*, 2017, **43**(11): 1886–1897
(陈伟宏, 安吉尧, 李仁发, 李万里. 深度学习认知计算综述. 自动化学报, 2017, **43**(11): 1886–1897)
- 79 LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision. In: *Proceedings of the 2010 IEEE International Symposium on Circuits and Systems*. Paris, France: IEEE, 2010. 253–256
- 80 Yan C, Zhang B L, Coenen F. Multi-attributes gait identification by convolutional neural networks. In: *Proceedings of the 8th International Congress on Image and Signal Processing*. Shenyang, China: IEEE, 2015. 642–647
- 81 Wu Z F, Huang Y Z, Wang L. Learning representative deep features for image set analysis. *IEEE Transactions on Multimedia*, 2015, **17**(11): 1960–1968
- 82 Zhang C, Liu W, Ma H D, Fu H Y. Siamese neural network based gait recognition for human identification. In: *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*. Shanghai, China: IEEE, 2016. 2832–2836
- 83 Tan T N, Wang L, Huang Y Z, Wu Z F. A Gait Recognition Method Based on Depth Learning, CN Patent 201410587758, June 2017
- 84 Wu Z F, Huang Y Z, Wang L, Wang X G, Tan T N. A comprehensive study on cross-view gait based human identification with deep CNNs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(2): 209–226
- 85 Wolf T, Babae M, Rigoll G. Multi-view gait recognition using 3D convolutional neural networks. In: *Proceedings of the 2016 IEEE International Conference on Image Processing*. Arizona USA: IEEE, 2016. 4165–4169
- 86 Li C, Min X, Sun S Q, Lin W Q, Tang Z C. Deep-Gait: a learning deep convolutional representation for view-invariant gait recognition using joint Bayesian. *Applied Sciences*, 2017, **7**(3): 210
- 87 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: *Proceedings of the 2015 International Conference on Pattern Recognition*. California, USA: ICPR, 2015. 212–218
- 88 Yu S Q, Chen H F, Wang Q, Shen L L, Huang Y Z. Invariant feature extraction for gait recognition using only one uniform model. *Neurocomputing*, 2017, **239**: 81–93
- 89 Gers F A, Schraudolph N N, Schmidhuber J. Learning precise timing with LSTM recurrent networks. *Journal of Machine Learning Research*, 2002, **3**: 115–143
- 90 Wang Kun-Feng, Gou Chao, Duan Yan-Jie, Lin Yi-Lun, Zheng Xin-Hu, Wang Fei-Yue. Generative adversarial networks: The state of the art and beyond. *Acta Automatica Sinica*, 2017, **43**(3): 321–332
(王坤峰, 苟超, 段艳杰, 林懿伦, 郑心湖, 王飞跃. 生成式对抗网络 GAN 的研究进展与展望. 自动化学报, 2017, **43**(3): 321–332)
- 91 Cao Z, Simon T, Wei S E, Sheikh Y. Realtime multi-person 2D pose estimation using part affinity fields. In: *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Hawaii, USA: CCPR, 2017. 1302–1310



王科俊 哈尔滨工程大学自动化学院教授. 主要研究方向为步态识别, 行人再识别, 机器学习. 本文通信作者.

E-mail: wangkejun@hrbeu.edu.cn

(WANG Ke-Jun Professor at the College of Automation, Harbin Engineering University. His research interest covers gait recognition, person re-identification, and machine learning. Corresponding author of this paper.)



丁欣楠 哈尔滨工程大学自动化学院硕士研究生. 主要研究方向为步态识别和机器学习.

E-mail: dingxinnan@hrbeu.edu.cn

(DING Xin-Nan Master student at the College of Automation, Harbin Engineering University. Her research interest covers gait recognition and machine learning.)



邢向磊 哈尔滨工程大学自动化学院讲师. 主要研究方向为模式识别和机器学习. E-mail: xingxl@hrbeu.edu.cn

(XING Xiang-Lei Lecturer at the College of Automation, Harbin Engineering University. His research interest covers pattern recognition and machine learning.)



刘美辰 哈尔滨工程大学自动化学院博士研究生. 主要研究方向为行人再识别和机器学习.

E-mail: meichen_0417@163.com

(LIU Mei-Chen Ph.D. candidate at the College of Automation, Harbin Engineering University. Her research interest covers person re-identification and machine learning.)