

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2024)07-1921-13

论文引用格式: Shen S, Zhang W H, Ding H, Zhang H, Sha C, Wang S and Chen S J. 2024. Research progress of three-dimensional gait recognition. Journal of Image and Graphics, 29(07):1921-1933(沈澍, 张文昊, 丁浩, 张浩, 沙超, 王森, 陈书军. 2024. 三维步态识别研究进展. 中国图象图形学报, 29(07):1921-1933)[DOI:10.11834/jig.230328]

三维步态识别研究进展

沈澍^{1*}, 张文昊¹, 丁浩², 张浩³, 沙超¹, 王森¹, 陈书军¹

1. 南京邮电大学计算机学院, 南京 210023; 2. 江苏警官学院刑事科学技术系, 南京 210031;
3. 南京工业大学机械与动力工程学院, 南京 211800

摘要: 步态识别在身份识别领域具有重要的研究意义。随着技术的发展, 步态识别的研究热点正从二维(2D)转向三维(3D)。与图像固有的2D信息相比, 用视觉技术还原的3D信息能更有效地预测人员的身份。在2D视觉领域中, 由于受到物体遮挡、视角变化等因素的影响, 传统的步态识别方法在实际应用中难以取得理想的识别性能。基于人体3D重建和人体3D姿态估计等3D人体技术, 近年来的研究在3D步态识别领域取得了一系列进展。本文介绍了3D步态识别方法, 探讨了基于3D步态的身份识别领域的研究现状、优势与不足; 总结了主要的3D步态数据集; 讨论了3D识别方法与2D识别方法的对比; 提出了3D身份识别领域未来潜在的研究方向, 包括3D数据集的采集和整理、2D和3D数据的多模态融合等。

关键词: 计算机视觉; 生物特征识别; 步态识别; 三维人体; 身份识别; 三维建模

Research progress of three-dimensional gait recognition

Shen Shu^{1*}, Zhang Wenhao¹, Ding Hao², Zhang Hao³, Sha Chao¹, Wang Sen¹, Chen Shujun¹

1. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;
2. Department of Forensic Science, Jiangsu Police Institute, Nanjing 210031, China;
3. School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing 211800, China

Abstract: Gait recognition is a new biometric identification technology that uses human walking posture and gait to determine a person's identity. Face recognition, which is considered traditional biometric recognition technology, is widely used, but it has the following defects: 1) recognition distance is limited; 2) it is vulnerable to occlusion and light and other factors; and 3) the results are at risk of being attacked by using face photos, video playback, and three-dimensional masks. In contrast, gait has the following advantages: 1) it can be identified from a long distance; 2) it is less affected by occlusion and illumination; and 3) it is not easy to disguise and deceive. Therefore, gait recognition is playing an increasingly important role in public safety, biometric authentication, video surveillance, and other fields. Gait recognition is mainly divided into two categories according to the dimension of input data: 1) two-dimensional (2D) gait recognition

收稿日期: 2023-06-16; 修回日期: 2023-09-20; 预印本日期: 2023-09-27

* 通信作者: 沈澍 shens@njupt.edu.cn

基金项目: 国家自然科学基金项目(52275535, 62272244); 江苏省高等学校自然科学研究重大项目(22KJA520010); 江苏省研究生科研与实践创新计划项目(SJCX22_0266, SJCX23_0273, KYCX22_0973); 痕迹检验鉴定技术公安部重点实验室开放课题项目(HJKF201915); 刑事检验四川省高校重点实验室开放课题项目(2023YB01); 南京邮电大学研究生教改课题项目(JGKT23_XJ07)

Supported by: National Natural Science Foundation of China (52275535, 62272244); Nature Science Research of Jiangsu Higher Education Institutions of China (22KJA520010); Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX22_0266, SJCX23_0273, KYCX22_0973)

based on 2D data and 2) three-dimensional (3D) gait recognition based on 3D data. At present, the research review in the field of gait recognition focuses on 2D gait recognition, usually from the perspective of traditional machine learning or deep learning. Gait recognition is moving from 2D to 3D. Compared with the inherent 2D information of the image, the 3D information restored by visual technology can more effectively predict human identity. In the field of 2D vision, traditional gait recognition methods have difficulty achieving better recognition performance because of the influence of object occlusion and view changes. On the basis of 3D human technology such as 3D human reconstruction and 3D human pose estimation, a series of progress has been made in the field of 3D gait recognition in recent years. To fully understand the existing research in the field of 3D gait recognition, this paper reviews and summarizes the research in this field. This paper discusses the research status, advantages, and disadvantages of identity recognition based on 3D gait; summarizes 3D gait recognition methods and 3D gait datasets; and provides potential research directions in the field of 3D identity recognition. This paper summarizes the different input data of existing 3D gait recognition methods and the recognition effect (recognition accuracy and speed) of these methods. These methods include multi-view-based, depth image-based, 3D skeleton-based, 3D point cloud-based, and 3D reconstruction-based recognition methods. This paper divides the 3D gait dataset into the indoor dataset and the outdoor dataset according to the acquisition environment. The 3D gait data include depth images, 3D skeleton, 3D human body grid, and 3D point cloud. In addition, this paper collates and compares the experimental results of different 3D gait recognition methods on various 3D gait datasets. Finally, this paper provides potential research directions for the field of 3D identity recognition. 1) Performance improvement and model optimization. Different from 2D gait, the performance of 3D gait is closely related to the 3D model. The 3D deep learning model needs to be optimized to improve the recognition performance of 3D gait in real scenes. For example, the training skills of vision Transformer (ViT) to improve performance are applied to 3D models such as 3D convolutional neural networks to improve the generalization and robustness of the model. The 3D model with the ViT concept is expected to learn more discriminative features from 3D gait data. 2) Collection and collation of 3D datasets. Compared with the 2D gait data set, the number of 3D gait datasets is small and the data types are not rich enough, which limits the development of 3D gait and requires further data collection by researchers. When the collected 3D gait dataset is sorted out, the training set and the test set can be divided in advance. For the test set, the registration set and the verification set are expected to be divided again, and the baseline algorithm that is easy to reproduce is used for evaluation. Rank-1 accuracy and mean average precision can be used as evaluation metrics. 3) Multi-modal fusion of 2D and 3D data. Compared with 2D data, 3D data contains more information, so the effective use of 3D data can improve the recognition performance in real scenes. In the field of gait recognition, current research mainly focuses on 2D data (human 2D skeleton, gait contour map, etc.) but has gradually shifted to 3D data (human 3D skeleton, human 3D mesh, depth image, etc.) in recent years. Future researchers can explore multidimensional gait recognition networks based on multimodal fusion to dynamically fuse 2D and 3D gait data. This fusion network combines the advantages of high 2D recognition efficiency and high 3D recognition accuracy and is expected to improve the performance of gait recognition in complex outdoor scenes. 4) Promotion and application of 3D vision technology. This paper mainly discusses the application of 3D vision technology in the emerging field of biometrics, particularly in gait recognition. Traditional biometrics, such as face recognition and fingerprint recognition, are also gradually transitioning from 2D to 3D. It is anticipated that this paper will aid researchers in understanding the latest advancements in 3D gait recognition and inspire the development of novel and advanced algorithms and models.

Key words: computer vision; biometric recognition; gait recognition; three-dimensional human body; identification recognition; three-dimensional modeling

0 引言

步态识别作为一种新型的生物特征识别技术,通过分析人体步行的姿态即步态来识别身份。传统

的生物特征识别技术——人脸识别已全面走向应用,但是存在以下缺陷:识别距离上存在局限性,容易受到遮挡和光照等因素的影响,存在被人脸照片、视频回放以及三维面具等方式攻击的风险。相比之下,步态具有以下优势:可远距离识别,受遮挡和光

照影响小,不易受到伪装和欺骗。因此,步态识别在公共安全、生物认证和视频监控等领域发挥着愈加重要的作用。

步态识别按照输入数据的维度主要分为两类:基于二维(2D)数据的2D步态识别和基于三维(3D)数据的3D步态识别,如表1所示。当前,步态识别领域的研究重点关注2D步态识别,通常从传统机器学习或深度学习的角度进行探讨。传统机器学习的综述文献包括王科俊和侯本博(2007)、贾晔焯等人(2012)的成果。王科俊和侯本博(2007)论述了早期步态识别的运动检测、周期检测、特征提取和识别算法这4个方面;贾晔焯等人(2012)综述了2012年以前发表的步态特征表达和识别方法。在深度学习方面,国内外的研究者均发表了相关的综述文献(何逸炜和张军平,2018;Sepas-Moghaddam和Etemad,2023;dos Santos等,2023)。何逸炜和张军平(2018)总结了基于深度学习的步态识别领域的进展及相关的识别算法;Sepas-Moghaddam和Etemad(2023)从方法的性能、特点、优点和缺点角度对基于深度学习的步态识别方法进行了全面的综述;dos Santos等人(2023)综述了以深度学习方法为主体的步态识别近

期工作(包括方法的优缺点、数据集的分类等)。本课题组已经在基于深度学习的步态识别方面进行了一定的研究(Shen等,2023b)。

3D步态识别是相对于2D步态识别而言的。2D步态识别在小规模室内数据集上的识别精度已趋近饱和,但在真实场景中的识别性能(如精准度及鲁棒性等)仍存在提升空间。在真实场景中,步态数据通常涵盖多个视角,这些视角往往因时间或环境因素而变化(王科俊等,2019;许文正等,2023)。然而,2D步态识别主要基于单一视角的信息,这限制了其在复杂场景下的识别性能。鉴于步态具有天然的3D特性(人在三维空间中行走),有效利用3D数据能够提升步态识别的性能。

与传统的2D步态识别相比,新型的3D步态识别往往能在非受控环境中提升识别性能,从而更好地满足实际应用需求。2D步态识别与3D步态识别从不同的角度进行了对比,如表1所示。

不同于现有的2D步态识别方面的综述,本文对基于3D步态的身份识别领域进行了综述,总结了3D步态识别方法和3D步态数据集,并比较了3D识别方法与2D识别方法在真实场景中的识别性能。

表1 2D步态识别与3D步态识别的对比

Table 1 Comparison of 2D gait recognition and 3D gait recognition

分类	2D步态识别	3D步态识别
输入数据	轮廓图、2D骨骼等2D数据	深度图像、3D骨骼、3D点云、人体3D网格等3D数据
识别方法	基于外观的识别方法(以轮廓图为输入) 基于模型的识别方法(以2D骨骼为输入)	基于深度图像的识别方法;基于3D骨骼、3D点云、人体3D重建的识别方法等
识别性能	在真实场景中偏低,远低于实验室受控场景	在非受控环境中有所提升,能较好地满足实际场景中的识别需求
识别效率	较高(得益于轻量级的2D识别模型)	偏低(3D建模比2D建模复杂度高)

1 3D步态识别方法

3D步态识别方法输入的3D数据可以从2D数据估计得到,也可以利用3D成像设备直接获取。现有的3D步态识别方法包括:基于多视角的识别方法、基于深度图像的识别方法、基于3D骨骼的识别方法、基于3D点云的识别方法和基于人体3D重建的识别方法。这些方法的输入数据不同,在识别效果(识别精度和速度)方面各有千秋,如表2所示。

1.1 基于多视角的识别方法

多视角的识别方法一般通过输入多台摄像机采集的不同视角下的数据进行识别。Zhao等人(2006)提出了一种从多台摄像机采集的视频序列中重构3D步态模型的方法。该方法从3D模型中提取下肢的运动轨迹,将其作为动态特征,并使用线性时间规整化(linear time normalization, LTN)来匹配和识别。López-Fernández等人(2016)使用多台摄像机采集多视点视频序列来重构3D步态序列,提出一种基于3D角度分析的旋转不变的特征,应用判别分析

(discriminant analysis, DA)和支持向量机(support vector machine, SVM)来分类识别。Deng 等人(2016, 2017)提出一种基于多视点融合与确定性学习的方法,通过不同视角来合成具有多种步态特征的轮廓图像。

多视角识别方法中,使用多台摄像机可以有效地缓解视角变化带来的识别难题。同时,多台摄像机的设置也有助于解决行人之间的遮挡问题,提高识别的准确性。然而,由于摄像机平衡视角和建模计算复杂,这些方法通常仅在全可控的多摄像机协作环境中使用,且难以适用于实际场景。

1.2 基于深度图像的识别方法

当与彩色图像相比较时,深度图像可以提供人体的3D结构,因为深度图像中的像素值代表了采集设备与被采集者之间的距离。深度图像的采集设备主要包括 Kinect(Zhang, 2012)等深度相机。深度相机采集深度图像的速度快,易于大规模使用,推动了基于深度图像的步态识别的发展。与用RGB相机采集彩色图像相比,使用深度相机采集深度图像成本更低、受外界因素干扰较少,提高了步态识别的性能和鲁棒性。深度图像和彩色图像进行融合识别可

以弥补单独使用彩色图像的不足。

Hofmann 等人(2012)将传统的2D步态能量图(gait energy image, GEI)(Han 和 Bhanu, 2006)与基于深度的3D模板进行比较。该3D模板包括深度数据上的步态能量图(gait energy image on depth data, Depth-GEI)、步态能量体积(gait energy volume, GEV)(Sivapalan 等, 2011)和深度梯度直方图能量图(depth gradient histogram energy image, DGHEI)(Hofmann 等, 2012)。实验结果表明,基于深度的3D模板可以获得更好的性能。Nunes 等人(2019)对基于深度图像的步态识别进行了全面的综述,并介绍了基于深度图像的数据集 GAID(gait from audio, image and depth)(Hofmann 等, 2014)、SZU RGB-D Gait(Yu 等, 2013)、SDUGait(Wang 等, 2016)等。

基于深度图像的步态识别面临的主要挑战是深度相机采集数据的范围有限(仅限室内近距离)和采集的数据往往质量不佳。深度相机主动发射的红外光在传输过程中,其强度会随着与目标物体距离的增加而显著下降,从而影响深度测量的准确性。因此,在实际应用中,需充分考虑深度相机的工作距离,以确保其能稳定、准确地获取深度信息。

表2 3D步态识别方法的识别效果(识别精度和识别速度)

Table 2 The recognition effect (recognition accuracy and speed) of 3D gait recognition method

类别	方法	输入数据	识别精度	识别速度
基于多视角的识别方法	Zhao 等人(2006)	使用多台摄像机采集的不同视角下的数据	较高,能够更精确地表达人体不同部位的物理空间	偏低,因为摄像机平衡视角和建模计算复杂
	López-Fernández 等人(2016) Deng 等人(2016) Deng 等人(2017)			
基于深度图像的识别方法	Hofmann 等人(2012) Sivapalan 等人(2011) Yu 等人(2013) Nunes 等人(2019)	通过 Kinect 等深度相机采集的深度图像	中等,深度图像比彩色图像受干扰影响小,但图像质量制约了精度的提高	中等,深度图像的采集速度取决于具体设备
	Owaidah 等人(2019) Liao 等人(2020) An 等人(2020) Zeng 等人(2021)	用3D姿态估计或人体骨骼追踪等方式提取的3D骨骼	有所提高(相比2D骨骼),能更好地应用于人体的遮挡问题	偏低(相比2D骨骼),因为3D骨骼建模难度和检测时间均较高
基于3D点云的识别方法	Ryu 和 Kamata(2011) Uy 等人(2019) Caesar 等人(2020) Shen 等人(2023a)	运用人体点云数据表示的3D人体模型	有所提高(相比2D数据),能较好地表示人体的3D特征	偏低(相比2D数据),因为3D数据的表示比2D数据的表示复杂度高
基于人体三维重建的识别方法	Loper 等人(2015) Zheng 等人(2022) Li 等人(2022) Meng 等人(2023)	从2D人物图像或视频恢复出人体的3D几何形状	较高,重建的人体3D模型适用于跨视角下的识别	偏低,3D建模的过程较为耗时,建模的计算量与特征维数较高

1.3 基于3D骨骼的识别方法

不同于仅拥有人体骨骼点平面坐标的2D骨骼, 3D骨骼提供了人体骨骼点的3D坐标, 从而能构建出性能更佳的人体3D模型。3D骨骼可通过以下方式获取: 1) 用3D姿态估计方法从2D图像中估计; 2) 直接用Kinect相机等3D体感摄像机采集。

3D姿态估计方法能预测出人体关节的3D坐标位置等信息。与2D姿态估计不同的是, 3D姿态估计在2D骨骼检测的基础上加入了深度信息, 因此更加精准。PoseGait(Liao等, 2020)使用一种3D姿态估计方法(Chen和Ramanan, 2017)从RGB图像中估计出行人3D姿态, 并将其作为特征, 结合基于先验知识的时空特征与人体姿态特征, 通过丰富的特征提高了识别率。姿态估计方法也应用到基于人体骨骼姿态的数据库 OUMVLP-Pose (OU-ISIR gait database, multi-view large population dataset with pose sequence)(An等, 2020)。

Kinect相机可通过骨骼追踪技术来较为准确地采集3D人体骨骼, 从而获取高鲁棒性的3D姿态数据进行步态识别。Owaidah等人(2019)运用Kinect的骨骼追踪技术对穿着传统长袍的沙特人进行步态识别, 面对长袍隐藏大部分关节而难以识别的挑战, 使用K最近邻(K-nearest neighbor, KNN)算法提取识别结果最佳的前3个关节, 将其用于步态识别。Kim等人(2020)提出一种基于Kinect的人体步态识别方法, 该方法结合静态和动态步态特征来提高识别的准确性, 并建立KNN分类器和长短期记忆网络(long-short term memory, LSTM)分类器来对不同的行人分类。Zeng等人(2021)采用Kinect的最新设备Azure Kinect采集了10个健康人体的运动轨迹, 并基于采集的3D人体骨骼提取步态特征, 证明了Azure Kinect能够提取有效的步态特征并进行步态分析。

3D骨骼比2D骨骼能更好地应用于人体的遮挡问题, 但建模难度和检测时间均高于2D骨骼。

1.4 基于3D点云的识别方法

深度学习作为人工智能领域的主流技术, 已成功解决了各种2D视觉问题, 但处理点云等3D数据面临着挑战(Loper等, 2015)。随着3D视觉的发展, 基于深度学习的3D点云处理技术应用于3D形状分类(Uy等, 2019)、3D目标检测与跟踪(Caesar等, 2020)和3D点云分割(王振燕等, 2023)等任务。

3D点云被引入步态识别领域来更有效地提取3D步态特征, 以提高识别率。Ryu和Kamata(2011)提出了一种基于人体点云的步态识别框架—球面空间模型(spherical space model, SSM)。该模型输入用立体相机获取的人体3D点云数据, 在正面视角下识别率优于现有的2D步态识别方法。Shen等人(2023a)研究了基于点云的3D步态识别, 提出了一种3D步态识别框架LidarGait。该框架将稀疏点云投影至深度图像中以学习物体几何信息的3D表示, 识别性能优于现有基于轮廓图等2D数据的步态识别方法。

点云数据能描述3D空间中物体的形状, 从而较好地表示人体3D特征、提升识别性能。然而, 点云数据分布稀疏, 且表示的复杂度比2D数据高, 在几何建模等领域仍然存在着较大的挑战。

1.5 基于人体3D重建的识别方法

人体3D重建能够从输入的2D人物图像或视频中恢复出人体的3D几何形状, 从而构建出人体3D模型。3D重建方法通常基于一种3D人体网格模型—蒙皮多人线性模型(skinned multi-person linear model, SMPL)(Loper等, 2015)。SMPL模型作为参数化的3D人体模型, 能够从2D图像或视频中精确估计出人体的3D网格和3D视角。3D网格应用于步态识别, 优势分为两个方面: 一是3D网格不仅能够提供人体在3D空间中的姿态, 还能够提供人体的形状, 这对于学习步态的判别性特征至关重要; 二是在跨视角匹配时, 可以探索3D视角来将人体的不同方向规范为有限个方向。

SMPL模型现已应用于基于人体网格的步态数据库 OUMVLP-Mesh(Li等, 2022)。OUMVLP-Mesh在大规模、多视角的步态数据库 OU-MVLP(OU-ISIR gait database, multi-view large population dataset)(Takemura等, 2018)(<http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitMVLP.html>)中应用了异步多视角训练框架, 获得的SMPL模型精度高且同时适用于多视角和单视角步态序列。与基于人体骨骼姿态的数据库 OUMVLP-Pose(An等, 2020)相比, OUMVLP-Mesh数据库构建的人体模型质量更高, 在相同基准条件下识别性能显著提高。

为了进一步探索人体的3D步态表示, Zheng等人(2022)设计了一种基于3D SMPL模型的步态识别框架 SMPLGait, 并构建了3D步态数据库 Gait3D

(<https://github.com/Gait3D/Gait3D-Benchmark>)。相比于面向非受控场景的步态数据库 GREW (gait recognition in the wild) (Zhu 等, 2021) (<https://github.com/GREW-Benchmark/GREW-Benchmark>), Gait3D 为真实场景的步态识别提供了更丰富的 3D 步态表示, 更贴合实际应用。Meng 等人(2023)提出了一种端到端的步态识别方法, 该方法基于人体 3D 重建, 以应对行人外观和着装变化带来的挑战。

人体 3D 重建能够由 2D 输入有效地构建人体 3D 模型。但是, 基于人体 3D 重建的方法建模过程耗时较长, 建模的计算量与特征维数较高, 阻碍了性能的进一步提升。

2 3D 步态数据集

2.1 3D 步态数据

3D 步态数据在 2D 步态数据的基础上, 增加了一维深度信息, 从而更全面地表示步态特征。3D 步态数据包括深度图像、3D 骨骼、3D 人体网格、3D 点云等, 其可视化效果如图 1 所示。图 1 包括以下 3 个部分:

1) 图 1(b) 的 3D 人体网格是通过步态数据集 CASIA-B (Institute of Automation Chinese Academy of Sciences gait database, dataset B) (Yu 等, 2006) (<http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>) 中图 1(a) 的 RGB 图像进行 3D 人体重建而成。

2) 图 1(c) 的深度图像和图 1(d) 3D 骨骼是由本文作者使用 Kinect 深度相机第三代版本 Azure Kinect 在实验室环境中采集而得。其中, 深度图像通过 Azure Kinect 深度相机查看器进行展示, 而 3D 骨骼则借助 Azure Kinect 的人体跟踪查看器来显示。

3) 图 1(e) 的 RGB 图像和图 1(f) 的 3D 点云是由数据集 SUSTech1K (Shen 等, 2023a) 的采集者分别用 RGB 相机和 3D 设备 LiDAR (light detection and ranging) 采集而得。

3D 数据通常作为 3D 识别方法的输入。因此, 图 1 中的 3D 步态数据与前文的 3D 步态识别方法是一一对应的关系: 图 1(b) 的 3D 人体网格对应于基于人体 3D 重建的识别方法; 图 1(c) 的深度图像对应于基于深度图像的识别方法; 图 1(d) 的 3D 骨骼对应于基于 3D 骨骼的识别方法; 图 1(f) 的 3D 点云对应于基于 3D 点云的识别方法。

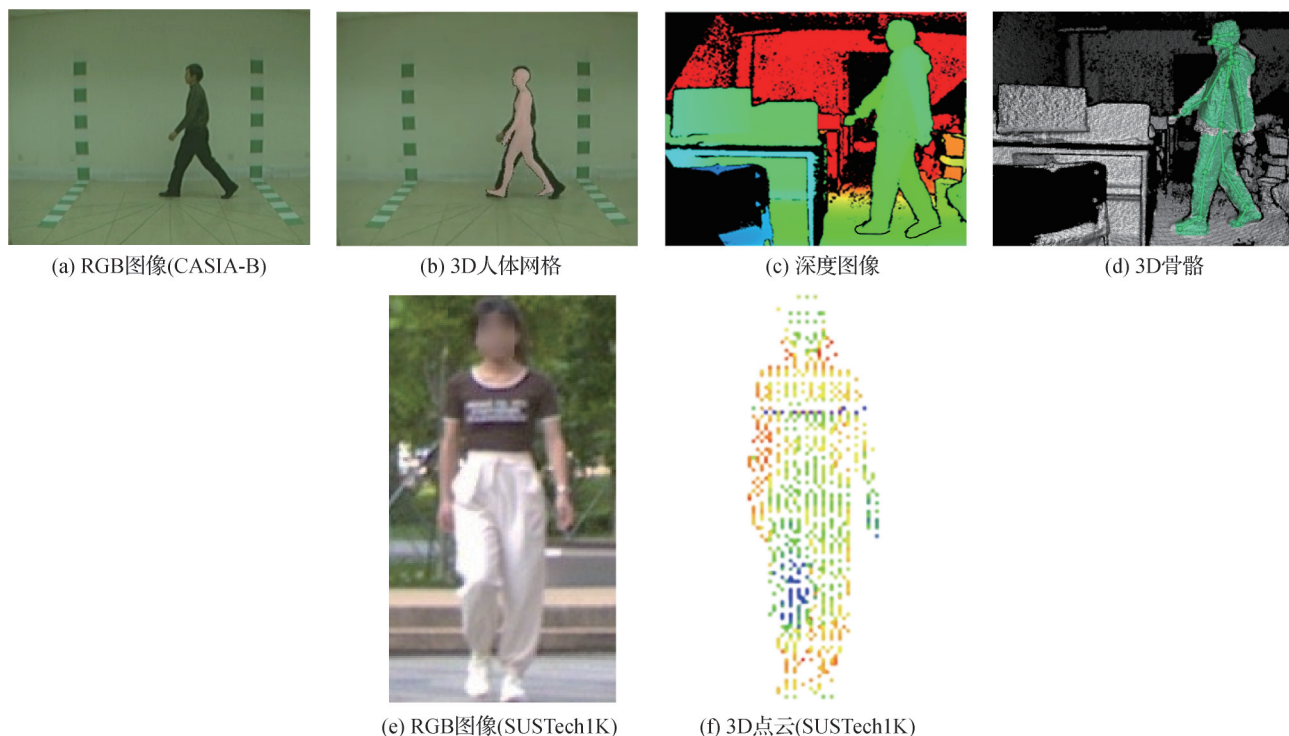


图 1 3D 步态数据可视化

Fig. 1 Visualization of 3D gait data ((a) RGB image (CASIA-B dataset); (b) 3D human mesh; (c) depth image; (d) 3D skeleton; (e) RGB image (SUSTech1K dataset); (f) 3D point cloud (SUSTech1K))

2.2 3D步态数据集分类

3D步态数据集按采集环境可分为室内数据集和室外数据集,如表3所示。其中,Gait3D(Zheng等,2022)(<https://github.com/Gait3D/Gait3D-Benchmark>)是一个室内数据集,在无约束的真实场景中进行采

集。该数据集中包含的3D人体网格能够支持从3D视角对步态动态信息进行建模。大规模的室外数据集包括 GREW(Zhu等,2021)和 CASIA-E(CASIA gait database, dataset E)(Song等,2023)(<http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>)。

表3 3D步态数据集
Table 3 3D gait dataset

名称	建立机构	采集环境	数据类型	样本容量/人
Gait3D(Zheng等,2022)	杭州电子科技大学	室内	3D网格(mesh)、轮廓图(silhouette)、人体关键点数据(key-points of human bodies)	4 000
GREW(Zhu等,2021)	清华大学	室外	轮廓图、步态能量图(gait energy image)、光流(optical flow)、2D和3D人体姿态数据(2D/3D human pose)	26 345
OUMVLP-Mesh(Li等,2022)	OU-ISIR 大阪大学产业	室内	3D人体网格(human mesh)	10 307
OUMVLP-Pose(An等,2020)	科学研究所	室内	人体姿态序列(pose sequence)	10 307
SZU RGB-D Gait(Yu等,2013)	深圳大学	室内	RGB图像、深度图像(depth image)	99
SDUGait(Wang等,2016)	山东大学	室内	3D骨骼节点(skeleton joint)、轮廓图	52
CASIA-B(Yu等,2006)	中国科学院自	室内	RGB图像(可提取3D骨骼)、轮廓图	124
CASIA-E(Song等,2023)	动化研究所	室外	RGB图像(可提取3D骨骼)、轮廓图、步态能量图(GEI)、步态热图像(gait thermal image)	1 014
SUSTech1K(Shen等,2023a)	南方科技大学	室外	RGB图像、轮廓图、3D点云(point cloud)	1 050

GREW(Zhu等,2021)数据集在复杂的室外场景中采集,示例如图2所示,主要面临以下挑战:1)视角和服装的变化:摄像头往往拍摄到同一个人不同视角、不同服装下的图像。跨视角、跨着装的应用场景下,行人的步态特征不易被准确捕捉,这对步态识别研究提出了挑战。2)复杂背景的干扰:图像的背景颜色可能与图像中行人的衣服颜色相近,因此难以将行人从背景中分割出来,这对行人的身份识别造成了干扰。3)密集人群的遮挡:实际应用中行人常常结伴行走,人体轮廓的提取容易受到其他行人的遮挡与干扰,从而影响了步态识别算法的精度。

CASIA-E(Song等,2023)数据集包含3个室外场景:简单静态背景(场景1)、复杂静态背景(场景2)和复杂动态背景(场景3),如图3所示。其中,在一



图2 GREW数据集中多种识别挑战的示例
Fig. 2 Examples of multiple recognition challenges in GREW dataset

天中不同时间段(上午和下午)采集时,光照条件(光照方向和光照强度)不同,因此增加了CASIA-E数据集的识别难度。

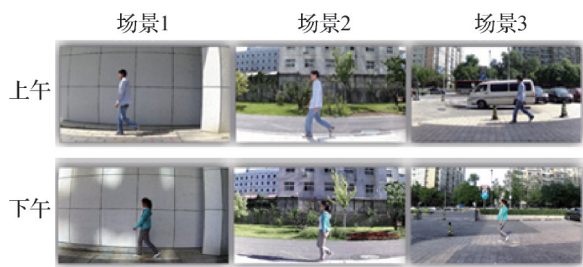


图3 CASIA-E数据集中不同室外场景的RGB图像

Fig. 3 RGB images in different outdoor scenes on CASIA-E dataset

3 3D识别方法与2D识别方法的对比

传统的2D步态识别方法主要分为基于外观的方法(一般以步态轮廓图为输入)和基于模型的方法(通常以人体骨骼为输入)。而新型的3D步态识别方法则主要采用基于人体3D表示的方法(往往以3D人体网格为输入)。为了比较3D识别方法与2D识别方法在真实场景中的识别性能,本文整理和对比了这两种识别方法在3D步态数据集Gait3D和GREW上的结果。这两个数据集分别在室内和室外的真实场景中采集得到。

3.1 室内数据集Gait3D上的识别结果

室内数据集Gait3D上的识别结果如表4所示。其中,输入尺寸为 128×88 像素和 64×44 像素。识

别结果的评估指标包括准确度Rank- n ($n = 1, 5$) (Rank- n 准确度表示在前 n 个候选项中正确识别的概率)、均值平均精度(mean average precision, mAP)以及平均反向负惩罚(mean inverse negative penalty, mINP)。

表4显示,以步态轮廓图为输入的方法在3D步态数据集Gait3D上,与在CASIA-B(Yu等,2006)和OU-MVLP(Takemura等,2018)等经典步态数据集上的性能相比,整体性能有所下降。这反映了实验室研究与更具挑战性的现实应用之间存在较大差距。

同时,基于外观的方法之间在识别准确率等性能指标上存在显著差异。例如,基于步态能量图(gait energy image, GEI)的方法GEINet(Shiraga等,2016)表现出较低的识别准确率,这表明GEI可能未能保留所有对步态识别重要的信息。同时,GaitPart(Fan等,2020)、GLN(gait lateral network)(Hou等,2020)、GaitGL(Lin等,2021)和CSTL(context-sensitive temporal feature learning)(Huang等,2021)方法考虑了步态序列帧的时间顺序信息,识别精度较低,这意味着室外场景中时间信息的学习难度较高,因为在无约束的环境中,人们可能会在行走过程中停下脚步,继续行走时速度和路线会发生变化。作为对比,GaitSet(Chao等,2022)通过无序集合的方式处理步态帧,取得了更优的识别结果。

表4 步态识别方法在Gait3D数据集上的结果

Table 4 Results of gait recognition method on Gait3D dataset

方法	输入特征	/%							
		Rank-1		Rank-5		mAP		mINP	
		128×88	64×44	128×88	64×44	128×88	64×44	128×88	64×44
GEINet(Shiraga等,2016)	步态轮廓图	7.00	5.40	16.30	14.20	6.05	5.06	3.77	3.14
GaitSet(Chao等,2022)	步态轮廓图	42.60	36.70	63.10	58.30	33.69	30.01	19.69	17.30
GaitPart(Fan等,2020)	步态轮廓图	29.90	28.20	50.60	47.60	23.34	21.58	13.15	12.36
GLN(Hou等,2020)	步态轮廓图	42.20	31.40	64.50	52.90	33.14	24.74	19.56	13.58
GaitGL(Lin等,2021)	步态轮廓图	23.50	29.70	38.50	48.50	16.40	22.29	9.20	13.26
CSTL(Huang等,2021)	步态轮廓图	12.20	11.70	21.70	19.20	6.44	5.59	3.28	2.59
PoseGait(Liao等,2020)	人体骨骼	0.24	-	1.08	-	0.47	-	0.34	-
GaitGraph(Teepe等,2021)	人体骨骼	6.25	-	16.23	-	5.18	-	2.42	-
SMPLGait(Zheng等,2022)	步态轮廓图、3D人体网格	53.20	46.30	71.00	64.50	42.43	37.16	25.97	22.23
GaitBase(Fan等,2023)	步态轮廓图等2D和3D数据	-	64.60	-	81.50	-	55.30	-	31.60

注:“-”表示无数据。

对于以人体骨骼为输入的方法,在Gait3D数据集上的实验结果表明,其识别精度相对于以步态轮廓图为输入的方法有所下降。这表明在真实场景中,人体骨骼携带的信息可能不如步态轮廓图丰富,因此对识别精度产生了影响。此外,在真实场景中,行走速度和路线具有不确定性,这也极大地影响了这些方法对人体动态建模的性能。

3D 识别方法 SMPLGait(Zheng等,2022)和 GaitBase(Fan等,2023)在Gait3D数据集上与现有的步态识别方法相比,具有更高的识别准确率,显示出3D特征表示在真实场景中提升步态识别性能的潜力。

3.2 室外数据集 GREW 上的识别结果

室外数据集 GREW 上的识别结果如表 5 所示。识别结果的评估指标为准确度 Rank- n ($n = 1, 5, 10, 20$)。

表 5 中,数据集 GREW 上的结果表明,GaitSet(Chao等,2022)和 GaitPart(Fan等,2020)在室外环境下具有较好的步态识别性能,与室内数据集 CASIA-B(Yu等,2006)和 OU-MVLP(Takemura等,2018)上的性能一致。其中,GaitSet 和 GaitPart 方法的 Rank-1 指标分别为 46.28% 和 44.01%。对于 Rank-5 和 Rank-20 指标,两种方法分别超过 60% 和 70%。与之相对的是,基于卷积神经网络(convolutional neural network, CNN)的方法 TS-CNN(two-stream CNN)(Wu等,2017)和 GEINet(Shiraga等,2016)将步态能量图(GEI)作为输入,而且模型层数相对较少,因此在 GREW 数据集上取得的准确率较

低。GEI 丢失了部分在无约束步态识别中重要的时间信息。与 GEINet 相比,TS-CNN 采用双流度量学习网络结构来缓解过拟合,识别率相对更高。

相比于 GaitSet 和 GaitPart 等经典的识别方法,新型的识别方法 GaitGL(Lin等,2021)和 CSTL(Huang等,2021)优化了先前方法的模型结构,以提高从输入的步态轮廓图中提取特征的能力,从而进一步提升了步态识别性能。

基于人体骨骼的方法 PoseGait(Liao等,2020)和 GaitGraph(Teepe等,2021)在与基于步态轮廓图的方法 GaitSet 和 GaitPart 相比较时,识别性能不足。这表明在真实场景中,人体骨骼携带的信息量可能不如步态轮廓图丰富,因此影响了识别精度。其中,基于图卷积神经网络(graph convolutional network, GCN)的 GaitGraph 识别率略优于基于卷积神经网络(CNN)的 PoseGait,彰显出图卷积神经网络比普通的卷积神经网络更适用于真实场景中的步态识别。总体而言,基于人体骨骼的方法在室外环境下进行步态识别,仍面临挑战并有改进空间。

3D 识别方法 GaitBase(Fan等,2023)通过输入三维步态数据,并运用数据增强技术,如水平翻转、旋转和透视变换,实现了比 2D 识别方法更优的识别性能。

为了综合比较 3D 识别方法与 2D 识别方法在不同场景中的识别性能,本文总结了步态识别方法在 3D 数据集 GREW 和 Gait3D 上的识别结果(Rank-5),如图 4 所示。从图 4 可以看出,3D 识别方法

表 5 步态识别方法在 GREW 数据集上的结果

Table 5 Results of gait recognition method on GREW dataset

方法	输入特征	Rank-1	Rank-5	Rank-10	Rank-20
GEINet(Shiraga等,2016)	步态轮廓图	6.82	13.42	16.97	21.01
TS-CNN(Wu等,2017)	步态轮廓图	13.55	24.55	30.15	37.01
GaitSet(Chao等,2022)	步态轮廓图	46.28	63.58	70.26	76.82
GaitPart(Fan等,2020)	步态轮廓图	44.01	60.68	67.25	73.47
GaitGL(Lin等,2021)	步态轮廓图	47.30	63.60	69.30	74.20
CSTL(Huang等,2021)	步态轮廓图	50.60	65.90	71.90	76.90
PoseGait(Liao等,2020)	人体骨骼	0.23	1.05	2.23	4.28
GaitGraph(Teepe等,2021)	人体骨骼	1.31	3.46	5.08	7.51
GaitBase(Fan等,2023)	2D和3D人体姿态数据、步态轮廓图	60.10	75.50	80.40	84.20

SMPLGait (Zheng 等, 2022) 在 Gait3D 数据集上的识别性能优于 2D 识别方法, 包括基于外观的方法 GaitSet (Chao 等, 2022)、GaitPart (Fan 等, 2020) 以及基于模型的方法 GaitGraph (Teepe 等, 2021)。Zheng 等人 (2022) 未能提供 SMPLGait 方法在 GREW 数据集上的结果, 因为 GREW 数据集缺少 SMPLGait 方法所需的输入数据—3D 人体网格。而 3D 识别方法 GaitBase (Fan 等, 2023) 进一步提升了 GREW 和 Gait3D 数据集上的识别性能。

由上述可知, 相比于传统的 2D 识别方法, 新型的 3D 识别方法能够提升真实场景中的识别性能。

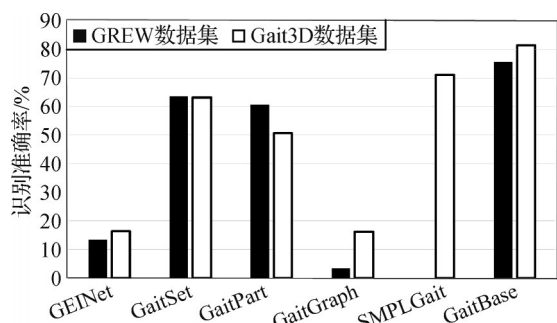


图4 步态识别方法在 GREW 和 Gait3D 数据集上的准确率
Fig. 4 Accuracy of gait recognition methods on GREW and Gait3D datasets

4 潜在的研究方向

近年来, 随着人体 3D 重建等 3D 视觉技术的发展, 步态识别等生物特征识别正从 2D 迈向 3D。然而, 在真实场景中应用 3D 步态识别仍面临一些挑战, 其中包括如何提升身份识别性能以适应复杂多变的环境。本文针对 3D 身份识别领域提供了一些潜在的研究方向。

1) 性能提升和模型优化。不同于 2D 步态, 3D 步态的性能与 3D 模型密切相关。为了提升 3D 步态在真实场景中的识别性能, 需要优化 3D 深度学习模型, 例如将视觉转换器 (vision Transformer, ViT) 提升性能的训练技巧应用于 3D 卷积神经网络等 3D 模型, 从而提高模型的泛化性和鲁棒性。引入 ViT 思想的 3D 模型有望从 3D 步态数据中学习更有判别力的特征, 从而提升身份识别的准确性。

2) 3D 数据集的采集和整理。相比于 2D 步态数据集, 3D 步态数据集的数量较少、数据种类不够丰

富, 极大地限制了 3D 步态在身份识别领域的应用和发展。因此, 进一步采集和整理高质量的 3D 步态数据集成为迫切需求。在采集过程中, 应充分考虑不同场景、不同姿态和不同个体的变化, 以构建更具代表性的数据集。在数据整理时, 除了划分训练集与测试集, 还应注重开发集与验证集的划分, 以便更准确地评估模型在身份识别任务上的性能。评估指标方面, 除了首位命中率 (rank-1 accuracy) 和均值平均精度 (mAP) 等, 还可以引入其他针对身份识别任务的特定指标, 以全面评估模型的性能。

3) 2D 和 3D 数据的多模态融合。相对于 2D 数据, 3D 数据包含的信息更丰富, 因此有效利用 3D 数据可以提高真实场景中的身份识别性能。在步态识别领域, 目前的研究主要集中在 2D 数据 (人体 2D 骨骼、步态轮廓图等), 近年来逐渐转向 3D 数据 (人体 3D 骨骼、人体 3D 网格、深度图像等)。因此, 探索基于多模态融合的多维度步态识别网络具有重要的研究价值。通过融合 2D 和 3D 步态数据, 能够充分利用两者的优点, 从而提升身份识别的准确性和鲁棒性。在室外复杂场景下, 多模态融合有助于应对光照变化、遮挡等挑战, 提高身份识别的稳定性。

4) 3D 视觉技术的推广和应用。本文主要讨论了 3D 视觉技术在新型的生物特征识别方法——步态识别中的应用。除此之外, 传统的生物特征识别, 如人脸识别和指纹识别, 也正从 2D 转向 3D。目前, 虽然已有研究探索了 3D 人脸识别和 3D 指纹采集等方面, 但在这些领域尚存在较大的探索空间。

例如, 3D 人脸识别相比于 2D 人脸识别, 在安全性与信息量上更具优势, 能提升不同光照和遮挡条件下的身份识别性能, 但 3D 人脸的数据集规模不如 2D 人脸 (孙哲南 等, 2021)。因此, 在数据驱动的深度时代, 3D 人脸识别的精度难以完全超越 2D 人脸识别, 在应用场景上也比较有限。期待未来的 3D 人脸识别将拥有更丰富的数据集, 并扩展到更广阔的应用领域。

此外, 在指纹采集方面, 3D 非接触指纹与 2D 接触式指纹相比, 增加了 3D 信息, 更符合手指的自然结构 (孙哲南 等, 2021)。非接触指纹采集能避免接触设备引起的卫生问题和按压手指导致的指纹变形, 因此采集过程更便捷、指纹识别率更高。这对身份识别至关重要, 因为准确的指纹信息有助于实现安全的身份识别。然而, 非接触采集的指纹图像在

不同部分的对比度较低,且细节点特征提取存在困难,这对身份识别的准确性构成挑战。因此,如何提升非接触指纹图像的质量,以便更准确地进行身份识别,是一个亟待探索和深入研究的问题。

因此,未来的研究可以进一步关注3D视觉技术在身份识别领域的推广和应用,探索更多创新性的方法,以提升身份识别的准确性和可靠性。

5 结 语

在实际复杂场景中,新型的三维(3D)步态识别相比于传统的二维(2D)步态识别,往往能达到更高的识别性能。本文对3D步态识别领域进行了梳理和总结,内容涵盖了3D步态识别方法、3D步态数据集以及3D识别方法与2D识别方法的对比。同时,本文探讨了3D身份识别领域潜在的研究方向,包括2D和3D数据的多模态融合、3D视觉技术的推广和应用。

参考文献(References)

- An W Z, Yu S Q, Makihara Y, Wu X H, Xu C, Yu Y, Liao R J and Yagi Y. 2020. Performance evaluation of model-based gait on multi-view very large population database with pose sequences. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2(4): 421-430 [DOI: 10.1109/TBIOM.2020.3008862]
- Ben X Y, Xu S and Wang K J. 2012. Review on pedestrian gait feature expression and recognition. *Pattern Recognition and Artificial Intelligence*, 25(1): 71-81 (黄晔焯, 徐森, 王科俊. 2012. 行人步态的特征表达及识别综述. *模式识别与人工智能*, 25(1): 71-81) [DOI: 10.16451/j.cnki.issn1003-6059.2012.01.010]
- Caesar H, Bankiti V, Lang A H, Vora S, Liong V E, Xu Q, Krishnan A, Pan Y, Baldan G and Beijbom O. 2020. nuScenes: a multi-modal dataset for autonomous driving//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 11618-11628 [DOI: 10.1109/CVPR42600.2020.01164]
- Chao H Q, Wang K, He Y W, Zhang J P and Feng J F. 2022. GaitSet: cross-view gait recognition through utilizing gait as a deep set. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7): 3467-3478 [DOI: 10.1109/TPAMI.2021.3057879]
- Chen C H and Ramanan D. 2017. 3D human pose estimation = 2D pose estimation + matching//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE: 7035-7043 [DOI: 10.1109/CVPR.2017.610]
- Deng M Q, Wang C and Chen Q F. 2016. Human gait recognition based on deterministic learning through multiple views fusion. *Pattern Recognition Letters*, 78: 56-63 [DOI: 10.1016/j.patrec.2016.04.004]
- Deng M Q, Wang C, Cheng F J and Zeng W. 2017. Fusion of spatial-temporal and kinematic features for gait recognition with deterministic learning. *Pattern Recognition*, 67: 186-200 [DOI: 10.1016/j.patcog.2017.02.014]
- dos Santos C F G, de Souza Oliveira D, Passos L A, Pires R G, Santos D F S, Valem L P, Moreira T P, Santana M C S, Roder M, Papa J P and Colombo D. 2023. Gait recognition based on deep learning: a survey. *ACM Computing Surveys*, 55(2): #34 [DOI: 10.1145/3490235]
- Fan C, Liang J H, Shen C F, Hou S H, Huang Y Z and Yu S Q. 2023. OpenGait: revisiting gait recognition toward better practicality//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, Canada: IEEE: 9707-9716 [DOI: 10.1109/CVPR52729.2023.00936]
- Fan C, Peng Y J, Cao C S, Liu X, Hou S H, Chi J N, Huang Y Z, Li Q and He Z Q. 2020. GaitPart: temporal part-based model for gait recognition//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 14213-14221 [DOI: 10.1109/CVPR42600.2020.01423]
- Han J and Bhanu B. 2006. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(2): 316-322 [DOI: 10.1109/TPAMI.2006.38]
- He Y W and Zhang J P. 2018. Deep learning for gait recognition: a survey. *Pattern Recognition and Artificial Intelligence*, 31(5): 442-452 (何逸炜, 张军平. 2018. 步态识别的深度学习: 综述. *模式识别与人工智能*, 31(5): 442-452) [DOI: 10.16451/j.cnki.issn1003-6059.201805006]
- Hofmann M, Bachmann S and Rigoll G. 2012. 2.5D gait biometrics using the depth gradient histogram energy image//*Proceedings of the 5th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. Arlington, USA: IEEE: 399-403 [DOI: 10.1109/BTAS.2012.6374606]
- Hofmann M, Geiger J, Bachmann S, Schuller B and Rigoll G. 2014. The TUM gait from audio, image and depth (GAID) database: multimodal recognition of subjects and traits. *Journal of Visual Communication and Image Representation*, 25(1): 195-206 [DOI: 10.1016/j.jvcir.2013.02.006]
- Hou S H, Cao C S, Liu X and Huang Y Z. 2020. Gait lateral network: learning discriminative and compact representations for gait recognition//*Proceedings of the 16th European Conference on Computer Vision*. Glasgow, UK: Springer: 382-398 [DOI: 10.1007/978-3-030-58545-7_22]
- Huang X H, Zhu D W, Wang H, Wang X G, Yang B, He B T, Liu W Y and Feng B. 2021. Context-sensitive temporal feature learning for gait recognition//*Proceedings of 2021 IEEE/CVF International Con-*

- ference on Computer Vision. Montreal, Canada: IEEE: 12889-12898 [DOI: 10.1109/ICCV48922.2021.01267]
- Kim W, Kim Y and Lee K Y. 2020. Human gait recognition based on integrated gait features using Kinect depth cameras//Proceedings of the 44th IEEE Annual Computers, Software, and Applications Conference (COMPSAC). Madrid, Spain: IEEE: 328-333 [DOI: 10.1109/COMPSAC48688.2020.0-225]
- Li X, Makihara Y, Xu C and Yagi Y. 2022. Multi-view large population gait database with human meshes and its performance evaluation. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 4(2): 234-248 [DOI: 10.1109/TBIOM.2022.3174559]
- Liao R J, Yu S Q, An W Z and Huang Y Z. 2020. A model-based gait recognition method with body pose and human prior knowledge. *Pattern Recognition*, 98: #107069 [DOI: 10.1016/j.patcog.2019.107069]
- Lin B B, Zhang S L and Yu X. 2021. Gait recognition via effective global-local feature representation and local temporal aggregation//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 14628-14636 [DOI: 10.1109/ICCV48922.2021.01438]
- Loper M, Mahmood N, Romero J, Pons-Moll G and Black M J. 2015. SMPL: a skinned multi-person linear model. *ACM Transactions on Graphics*, 34(6): #3248 [DOI: 10.1145/2816795.2818013]
- López-Fernández D, Madrid-Cuevas F J, Carmona-Poyato A, Muñoz-Salinas R and Medina-Carnicer R. 2016. A new approach for multi-view gait recognition on unconstrained paths. *Journal of Visual Communication and Image Representation*, 38: 396-406 [DOI: 10.1016/j.jvcir.2016.03.020]
- Meng C Y, He X B, Tan Z and Luan L. 2023. Gait recognition based on 3D human body reconstruction and multi-granular feature fusion. *The Journal of Supercomputing*, 79(11): 12106-12125 [DOI: 10.1007/s11227-023-05143-0]
- Nunes J F, Moreira P M and Tavares J M R S. 2019. Benchmark RGB-D gait datasets: a systematic review//Proceedings of 2019 VII ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing. Porto, Portugal: Springer: 366-372 [DOI: 10.1007/978-3-030-32040-9_38]
- Owaidah E M, Aloufi K S and Alkhatib J H. 2019. Gait recognition for Saudi Costume using Kinect skeletal tracking//Proceedings of the 2nd International Conference on Computer Applications and Information Security (ICCAIS). Riyadh, Saudi Arabia: IEEE: 1-5 [DOI: 10.1109/CAIS.2019.8769552]
- Ryu J and Kamata S I. 2011. Front view gait recognition using spherical space model with human point clouds//Proceedings of the 18th IEEE International Conference on Image Processing. Brussels, Belgium: IEEE: 3209-3212 [DOI: 10.1109/ICIP.2011.6116351]
- Sepas-Moghaddam A and Etemad A. 2023. Deep gait recognition: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1): 264-284 [DOI: 10.1109/TPAMI.2022.3151865]
- Shen C F, Fan C, Wu W, Wang R, Huang G Q and Yu S Q. 2023a. LidarGait: benchmarking 3D gait recognition with point clouds//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE: 1054-1063 [DOI: 10.1109/CVPR52729.2023.00108]
- Shen S, Sun S S, Li W J, Wang R C, Sun P, Wang S and Geng X Y. 2023b. A classifier based on multiple feature extraction blocks for gait authentication using smartphone sensors. *Computers and Electrical Engineering*, 108: #108663 [DOI: 10.1016/j.compeleceng.2023.108663]
- Shiraga K, Makihara Y, Muramatsu D, Echigo T and Yagi Y. 2016. Geinet: view-invariant gait recognition using a convolutional neural network//Proceedings of 2016 International Conference on Biometrics (ICB). Halmstad, Sweden: IEEE: 1-8 [DOI: 10.1109/ICB.2016.7550060]
- Sivapalan S, Chen D, Denman S, Sridharan S and Fookes C. 2011. Gait energy volumes and frontal gait recognition using depth images//Proceedings of 2011 International Joint Conference on Biometrics (IJCB). Washington, USA: IEEE: 1-6 [DOI: 10.1109/IJCB.2011.6117504]
- Song C F, Huang Y Z, Wang W N and Wang L. 2023. CASIA-E: a large comprehensive dataset for gait recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3): 2801-2815 [DOI: 10.1109/TPAMI.2022.3183288]
- Sun Z N, He R, Wang L, Kan M N, Feng J J, Zheng F, Zheng W S, Zuo W M, Kang W X, Deng W H, Zhang J, Han H, Shan S G, Wang Y L, Ru Y W, Zhu Y H, Liu Y F and He Y. 2021. Overview of biometrics research. *Journal of Image and Graphics*, 26(6): 1254-1329 (孙哲南, 赫然, 王亮, 阚美娜, 冯建江, 郑方, 郑伟诗, 左旺孟, 康文雄, 邓伟洪, 张杰, 韩琥, 山世光, 王云龙, 茹一伟, 朱宇豪, 刘云帆, 何勇. 2021. 生物特征识别学科发展报告. *中国图象图形学报*, 26(6): 1254-1329) [DOI: 10.11834/jig.210078]
- Takemura N, Makihara Y, Muramatsu D, Echigo T and Yagi Y. 2018. Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. *IPSJ Transactions on Computer Vision and Applications*, 10(1): #4 [DOI: 10.1186/s41074-018-0039-6]
- Teepe T, Khan A, Gilg J, Herzog F, Hörmann S and Rigoll G. 2021. Gaitgraph: graph convolutional network for skeleton-based gait recognition//Proceedings of 2021 IEEE International Conference on Image Processing (ICIP). Anchorage, USA: IEEE: 2314-2318 [DOI: 10.1109/ICIP42928.2021.9506717]
- Uy M A, Pham Q H, Hua B S, Nguyen T and Yeung S K. 2019. Revisiting point cloud classification: a new benchmark dataset and classification model on real-world data//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 1588-1597 [DOI: 10.1109/ICCV.2019.00167]
- Wang K J, Ding X N, Xing X L and Liu M C. 2019. A survey of multi-

- view gait recognition. *Acta Automatica Sinica*, 45(5): 841-852 (王科俊, 丁欣楠, 邢向磊, 刘美辰. 2019. 多视角步态识别综述. *自动化学报*, 45(5): 841-852) [DOI: 10.16383/j.aas.2018.c170559]
- Wang K J and Hou B B. 2007. A survey of gait recognition. *Journal of Image and Graphics*, 12(7): 1152-1160 (王科俊, 侯本博. 2007. 步态识别综述. *中国图象图形学报*, 12(7): 1152-1160) [DOI: 10.3969/j.issn.1006-8961.2007.07.002]
- Wang Y F, Sun J D, Li J and Zhao D. 2016. Gait recognition based on 3D skeleton joints captured by Kinect//*Proceedings of 2016 IEEE International Conference on Image Processing (ICIP)*. Phoenix, USA: IEEE: 3151-3155 [DOI: 10.1109/ICIP.2016.7532940]
- Wang Z Y, Sun H Y and Sun X P. 2023. Survey on large scale 3D point cloud processing using deep learning. *Computer Systems and Applications*, 32(2): 1-12 (王振燕, 孙红岩, 孙晓鹏. 2023. 基于深度学习的大规模三维点云处理综述. *计算机系统应用*, 32(2): 1-12) [DOI: 10.15888/j.cnki.csa.008743]
- Wu Z F, Huang Y Z, Wang L, Wang X G and Tan T N. 2017. A comprehensive study on cross-view gait based human identification with deep CNNs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(2): 209-226 [DOI: 10.1109/TPAMI.2016.2545669]
- Xu W Z, Huang T H, Ben X Y, Zeng Y and Zhang J P. 2023. Cross-view gait recognition: a review. *Journal of Image and Graphics*, 28(5): 1265-1286 (许文正, 黄天欢, 贲晔, 曾翌, 张军平. 2023. 跨视角步态识别综述. *中国图象图形学报*, 28(5): 1265-1286) [DOI: 10.11834/jig.220458]
- Yu S Q, Tan D L and Tan T N. 2006. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition//*Proceedings of the 18th International Conference on Pattern Recognition*. Hong Kong, China: IEEE: 441-444 [DOI: 10.1109/ICPR.2006.67]
- Yu S Q, Wang Q and Huang Y Z. 2013. A large RGB-D gait dataset and the baseline algorithm//*Proceedings of the 8th Chinese Conference on Biometric Recognition*. Jinan, China: Springer: 417-424 [DOI: 10.1007/978-3-319-02961-0_52]
- Zeng Y, Wu L F and Xie D L. 2021. Gait analysis based on azure Kinect 3D human skeleton//*Proceedings of 2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI)*. Kunming, China: IEEE: 1059-1062 [DOI: 10.1109/CISAI54367.2021.00212]
- Zhang Z Y. 2012. Microsoft Kinect sensor and its effect. *IEEE Multimedia*, 19(2): 4-10 [DOI: 10.1109/MMUL.2012.24]
- Zhao G Y, Liu G Y, Li H and Pietikainen M. 2006. 3D gait recognition using multiple cameras//*Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR06)*. Southampton, UK: IEEE: 529-534 [DOI: 10.1109/FGR.2006.2]
- Zheng J K, Liu X C, Liu W, He L X, Yan C G and Mei T. 2022. Gait recognition in the wild with dense 3D representations and a benchmark//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, USA: IEEE, 20196-20205 [DOI: 10.1109/CVPR52688.2022.01959]
- Zhu Z, Guo X D, Yang T, Huang J J, Deng J K, Huang G, Du D L, Lu J W and Zhou J. 2021. Gait recognition in the wild: a benchmark//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*. Montreal, Canada: IEEE, 14769-14779 [DOI: 10.1109/ICCV48922.2021.01452]

作者简介

沈澍,男,副教授,主要研究方向为人工智能物联网和人机交互。E-mail:shens@njupt.edu.cn

张文昊,男,硕士研究生,主要研究方向为计算机视觉和步态识别。E-mail:1021041224@njupt.edu.cn

丁浩,男,副教授,主要研究方向为步态识别和物证检验。E-mail:dinghao@jspi.cn

张浩,男,教授,主要研究方向为模式识别与人工智能、智能装备设计与研发。E-mail:zhanghaowo79@njtech.edu.cn

沙超,男,教授,主要研究方向为泛在网协同信息感控。E-mail:shac@njupt.edu.cn

王森,女,硕士研究生,主要研究方向为步态识别和计算机视觉。E-mail:1222045737@njupt.edu.cn

陈书军,男,硕士研究生,主要研究方向为计算机视觉和步态识别。E-mail:chnshujun1114@gmail.com