

# 基于人体姿态的多视角步态识别技术

杨东亮<sup>1</sup>, 宋昌江<sup>1</sup>, 孙思文<sup>1</sup>, 邢向磊<sup>2</sup>, 刘晓曦<sup>1</sup>

(1. 黑龙江省科学院智能制造研究所, 黑龙江 哈尔滨 150090;

2. 哈尔滨工程大学, 黑龙江 哈尔滨 150090)

**摘要:** 针对步态识别中视角受限导致的识别率低问题, 提出一种基于人体姿态估计模型的多视角步态生成方法。该方法通过模型提取视频帧中的人体姿态参数, 并利用Rodrigues旋转矩阵生成跨视角步态序列。同时设计了包含姿态平均和角度校正的校准网络, 通过姿态参数均值化和根节点修正模型优化姿态参数。实验结果表明, 经校准后的跨视角步态序列在视觉效果和识别精度上均显著提升。通过Gaitset验证, 该方法在定性和定量评估中均表现出色, 能够有效扩展数据库视角并生成更准确的跨视角步态模型。

**关键词:** 步态识别; 人体姿态; 多视角步态; 视角生成

中图分类号: TP391.41 文献标识码: A 文章编号: 1003-7241(2025)08-0075-04

## Multi-view Gait Recognition Technology Based on Human Posture

YANG Dongliang<sup>1</sup>, SONG Changjiang<sup>1</sup>, SUN Siwen<sup>1</sup>, XING Xianglei<sup>2</sup>, LIU Xiaoxi<sup>1</sup>

(1. Intelligent Manufacturing Institute, Heilongjiang Academy of Sciences, Harbin 150090, China;

2. Harbin Engineering University, Harbin 150090, China)

**Abstract:** Aiming at the low recognition rate caused by limited viewing angle in gait recognition, a multi-view gait generation method based on human pose estimation model is proposed in this paper. In this method, human pose parameters in video frames are extracted by the model, and cross-view gait sequences are generated by Rodrigues rotation matrix. In order to solve the attitude parameter error problem, this paper designs a calibration network including attitude average and angle correction, and optimizes the attitude parameters by means of attitude parameter averaging and root node correction model. The experimental results show that the visual effect and recognition accuracy of the calibrated cross-viewing angle gait sequence are significantly improved. Through Gaitset validation, the proposed method performs well in both qualitative and quantitative evaluation, and can effectively extend the database perspective and generate more accurate cross-perspective gait models.

**Keywords:** gait recognition; human posture model; multiple view gait; perspective generate

## 0 引言

生物特征识别技术通过个体固有的生理特征(如指纹、人脸、虹膜)或行为特征(如笔迹、语音、步态)进行身份认证。近年来,视频监控技术因其突破时空限制和证据固定优势,逐渐成为维护社会稳定、服务民生、打击犯罪和精准执法的重要手段。在视频监控应用中,目标识别与追踪是破案的核心环节,其效率与准确性至关重要。与其他人脸、指纹等生物识别技术相比,基于视频的步态识别具有显著优势:首先,步态特征可以通过低分辨率摄像头远距离获取,而人脸和虹膜特征则需要近距离接触式采集;其次,步态识别无需受试者主动配合,能够在非接触、非配合的情况下完成识别;最后,步态特征难以伪装或隐藏,而人脸识别系统易受伪造图像或口罩等干

扰。步态识别技术主要包括视频数据采集、步态分割、特征提取和步态识别四个步骤。研究学者针对步态识别的技术流程提出了多种处理方法。周航等<sup>[1]</sup>提出了结合风格编码和双通道时空卷积生成对抗网络模型的步态识别方法。该模型通过双通道时空卷积网络分别提取步态图像中的复杂时空特征与步态特征,并引入风格编码器来减少生成图像与原图之间的语义差异,从而提升生成器的稳定性。刘庆华等<sup>[2]</sup>采用Transformer模型用于步态识别,融合轮廓特征与骨骼特征,实现跨模态的互补,同时使用了多尺度时间特征融合器,采用Transformer结构的编码部分提取全局时空特征,并通过残差结构提取局部时空特征,以此来提高模型识别的准确率。魏永超等<sup>[3]</sup>提出了融合全局多尺度和局部细粒度特征的点云步态识别方法。该方法将点云投影为深度步态图,使用跨视角数据变换模块,增强步态特征的表征力。许振齐等<sup>[4]</sup>将注意力模块集成到模型中,自适应地关注输入步态序列关键

\*基金项目:黑龙江省省属科研院所科研业务费项目(CZKYF2024-1-C020)

收稿日期:2025-04-29

通道和空间位置,构造了一种融合基于卷积注意力机制的模型(convolutional block attention module, CBAM)和跨时间尺度特征分析的步态识别方法,提升了步态识别性能。郭坛等<sup>[5]</sup>针对帕金森步态识别问题,将时域卷积网络(temporal convolutional network, TCN)层替换深度可分离卷积中的逐通道卷积,减少模型的参数量和参数复杂度。通过分析了帕金森步态数据的冗余性,降低了模型训练所需的存储空间,进一步提升了模型在轻量设备中的可部署能力。郭剑等<sup>[6]</sup>针对毫米波雷达点云的步态识别问题,提出了一种基于点体素交叉注意力机制的步态识别方法,使用双分支网络分别从点数据和体素数据协作提取并融合全局与局部特征,以更好地提取人体步态信息。霍丽娜等<sup>[7]</sup>使用空洞卷积扩大神经元感受野,采用了扩张重参数化模块,通过重参数化方法融合多尺度卷积核参数,使模型捕获更多的全局上下文信息;提升了步态特征的判别性。霍威等<sup>[8]</sup>提出一种由形变场和步态剪影构成网络模型。采用形变场数据的稀疏性,以获取形变场的多层次空间结构信息。采用步态剪捕捉上下文信息增强模块,双分支网络的输出特征经过特征融合得到最终的步态表示。李浩森等<sup>[9]</sup>提出一种基于Transformer提取时间上下文信息的步态识别方法,可以提取多分辨率细粒度时间的特征,提高识别的准确率。陈万志等<sup>[10]</sup>通过行人的轮廓图,计算其均值,合成步态GEI能量图,增强了人体外观的表示。使用注意力机制,增强模型的特征提取能力,从而提高步态识别的精度。钱兴等<sup>[11]</sup>提出一种基于STM32单片机的步态识别系统,解决占用内存资源过大,导致实时性差的问题。张红颖等<sup>[12]</sup>建立一种多级分块结构的步态识别方法,解决了衣着与背包的遮挡导致识别准确率不高的问题。彭小波等<sup>[13]</sup>提出以视频流为输入,设计一种步态时空网络作为特征学习,在网络末端采用softmax分类器进行分类。刘瑞华等<sup>[14]</sup>建立边缘运动捕捉模块用于提取包含时间信息的边缘运动特征,并且建立多层次特征提取模块用于提取包含不同粒度全局和局部信息的多层级精细特征,达到跨视角步态识别的目的。王晓路等<sup>[15]</sup>建立随机裁剪随机遮挡的数据增强方法,采用结合注意力机制的两路复合卷积层提取步态特征,将两个分支的特征信息相加融合,再通过全连接层实现步态识别。

然而,现有步态识别系统的鲁棒性仍有待提高,其性能易受视角变化、服装差异、携带物品等协变量影响<sup>[16-18]</sup>。在步态剪影序列生成过程中,行人衣物、背包等额外信息不仅会掩盖个体的行走模式,还会引入冗余信息,尤其在跨状态识别场景中,这些冗余信息会显著影响后续学习效果。因此,本文开展多视角步态识别研究。首先,设计深度神经网络估计三维人体姿态与形状参数,进而从任

意视角渲染二维步态序列,有效弥补传统二维模型在视角信息上的不足,提升多视角步态识别性能。

## 1 方法设计

### 1.1 整体框架

多视角步态识别研究,首先,引入蒙皮多人线性模型(skinned multi person linear, SMPL)模型,这是一种基于姿态参数生成人体姿态三维图像的方法。接着,估计人体姿态参数。随后,利用罗德里格斯公式和校正网络将步态旋转至其他角度。

### 1.2 人体姿态估计网络

SMPL模型用于表示人体姿态参数和身体形状参数。通过控制这两个参数,我们可以在任意给定帧中渲染关节化的人体模型。构建一种利用视频序列估计参数的新方法。该方法在很大程度上解决了步态识别中数据库有限和识别率不足的问题。

(1) 利用卷积神经网络(convolutional neural network, CNN)对视频中的步态序列特征进行提取

(2) 利用双向门控循环单元(gated recurrent unit, GRU)提取过去和未来时间关联的潜在信息,其中GRU是一种循环神经网络的变体,它被引入以解决传统RNN在处理长序列数据时存在的梯度消失和梯度爆炸问题。

(3) 利用提取特征在每个时间的实例回归SMPL身体模型的参数,其中,SMPL模型是一个参数化的人体模型,通过改变SMPL模型参数来控制人体的形状和人体的姿态。

总的来说,所提出的损失由 $2D(x)$ ,姿态( $\theta$ )和形状( $\beta$ )损失组成。总损失为

$$L_G = L_{2D} + L_{SMPL}$$

其中每一项计算如下:

$$L_{2D} = \sum_t \|x_t - \hat{x}_t\|_2$$

$$L_{SMPL} = \|\beta - \hat{\beta}\|_2 + \sum_t \|\theta - \hat{\theta}_t\|_2$$

式中, $L_G$ 是整体损失, $L_{2D}$ 是关键点损失, $L_{SMPL}$ 是姿态损失, $x_t$ 代表关键点。

### 1.3 罗德里格斯公式

建立通过当前视角生成其他视角下的步态序列的方法:通过模型训练得到SMPL参数,通过罗德里格斯公式来旋转到其他角度下生成其他角度的步态序列,具体流程如下:

(1) 利用模型来预估当前步态视频下的SMPL模型参数,得到当前步态视频姿态参数 $\theta \in R^{72}$ 和形态参数 $\beta \in R^{10}$ 。

(2) 对生成的参数进行旋转, $\beta \in R^{10}$ 是控制人体的高

矮胖瘦的形状参数等。 $\theta \in R^{72}$ 为控制人体的各个关节的姿态参数。然后利用罗德里格斯公式对根节点旋转矩阵进行旋转,得到旋转后的SMPL参数。

(3) 最后利用SMPL模型对其进行渲染成步态图像。

#### 1.4 校正网络

首先输入来自当前视角和期望视角的相同长度的步态序列。接着,将这一步态序列输入到模型,以预测期望视角下个体的SMPL参数:姿态参数和形状参数。

接下来,将两个视角的渲染图像输入角度校正模型进行处理。如图1所示,角度校正模型输出一个根节点增量,将该增量与期望视角的姿态参数中的根节点向量相加,从而得到校正后的根节点向量。至此,校正过程完成。最后,所使用的损失函数是均方误差(MSE)损失函数。

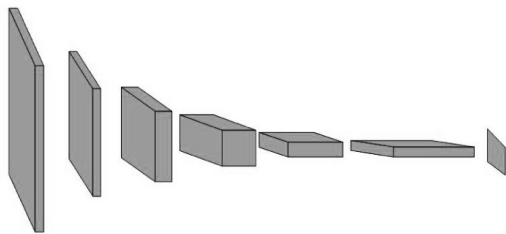


图1 角度校正网络

由于模型以灰度图像作为输入,而灰度图像仅有两个通道,因此模型的输入尺寸为(2,128,128)。输出尺寸为(1,3,3)。该模型由六个卷积层构成。在每一卷积层之后,应用了批归一化(batch normalization, BN)和线性整流函数(rectified linear unit, ReLU)激活函数。这种方法极大地提升了训练效率,并增强了模型的非线性能力。采用均方误差(mean-square error, MSE)损失函数作为模型的损失函数,其公式如下:

$$MSE = \frac{1}{n} \sum_{i=1}^n (f(x) - y)^2$$

均方误差(MSE)计算的是模型预测值与真实标签之间的接近程度。MSE损失函数能够有效反映预测值与标签之间的差异。因此,将利用旋转至期望视角的校正后的根节点向量,并将其与真实的姿态参数进行比较,以计算损失值。

## 2 实验分析

### 2.1 实验数据集

为了实现高效的步态识别,考虑采用CASIA-B数据集。CASIA-B包含11种不同的视角(0°,18°,⋯,180°),共计124人。该数据集适用于跨视角实验。使用CASIA-B数据库中前74名参与者的数据进行训练,其余50名参与者的数据用于测试。

实验参数设置:本实验使用Adam优化器,共进行

80 000轮训练,学习率设置为0.000 1。

### 2.2 实验分析

利用提出的方法,将步态序列旋转90°,从上到下分别生成:90°时的步态(标签)、原始角度时的步态、利用罗德里格斯公式旋转得到的90°步态、位姿平均后的步态和经过角度校正网络的步态渲染。实验结果如图2所示。通过实验结果,可以看到,经过提出的方法得到的姿态有了明显的提高。

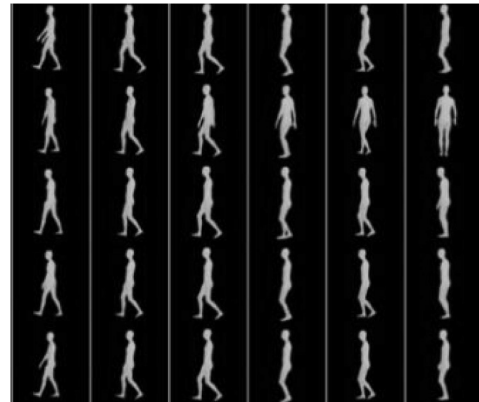


图2 校正网络生成的图像

表1 识别方法对比实验

group	0°	72°	90°	162°	180°	mean
1	34.00	44.00	41.20	43.00	32.50	44.48
2	23.90	43.60	41.00	35.50	24.00	37.96
3	80.41	82.34	83.50	79.18	80.51	81.48
4	87.96	89.70	89.60	88.37	86.33	88.34

4组实验分别是使用原始角度时的步态,直接使用旋转后的步态,多角度姿态参数均值模型和角度修正模型。如表1所示,可以看出Gaitset直接使用Rodriguez公式旋转到90°的准确率低于组1的结果,这说明直接通过估算的人体姿态参数存在误差,直接使用该参数进行物理旋转会放大误差。组3采用多角度下姿态参数均值作为最终姿态参数,在对姿态参数取均值后,明显准确率从37.96%提升至81.48%。组4通过角度修正模型对根节点进行修正。通过实验对比可以看出,准确率相比组3提升了6.86%,各个角度下的准确率均有所提升。我们观察到经过角度修正后,人物的角度有了较好的修正效果,说明提出的角度修正模型是有效的

## 3 结束语

本文提出一种基于三维姿态估计的多视角步态生成方法,通过模型提取姿态参数,结合Rodrigues旋转与校正网络优化跨视角步态序列。实验表明,该方法能够有效扩展数据库视角,生成高精度步态模型,为复杂场景下的步态识别提供新思路。

## 参考文献:

- [1] 周航, 宋珂蔚, 徐源昊, 等. 融合风格编码的双通道生成对抗网络步态识别[J]. 吉林大学学报(工学版), 2025(4): 1-14.
- [2] 刘庆华, 王正勇, 秦予祺, 等. 结合多特征融合和Transformer的步态识别[J]. 计算机技术发展, 2025(4): 1-10.
- [3] 魏永超, 谢卫鑫, 张娅岚, 等. 基于多尺度特征融合的跨视角云步态识别[J]. 电子测量技术, 2025(4): 1-9.
- [4] 许振齐, 朴燕, 康继元, 等. 融合3D-CBAM和跨时间尺度特征分析的步态识别方法[J]. 电子测量与仪器学报, 2025, 39(2): 32-40.
- [5] 郭坛, 时文雅, 郁战, 等. 基于轻量化网络的帕金森步态识别方法[J]. 传感器与微系统, 2025, 44(4): 143-147.
- [6] 郭剑, 鲁捷敏, 韩崇, 等. 基于点与体素融合的人体步态识别方法[J]. 小型微型计算机系统, 2025(4): 1-9.
- [7] 霍丽娜, 薛乐仁, 戴钰俊, 等. 基于扩张重参数化和空洞卷积架构的步态识别方法[J]. 计算机应用, 2025, 45(4): 1285-1292.
- [8] 霍威, 王科, 唐俊, 等. 一种基于人体轮廓形变场的双流网络步态识别方法[J]. 电子与信息学报, 2024, 46(10): 4062-4071.
- [9] 李浩森, 张含笑, 邢向磊. 联合局部多尺度和全局上下文特征的步态识别[J]. 智能系统学报, 2024, 19(4): 853-862.
- [10] 陈万志, 唐浩博, 王天元. 融合轮廓增强和注意力机制的改进GaitSet步态识别方法[J]. 电子测量与仪器学报, 2024, 38(1): 203-210.
- [11] 钱兴, 张晓明. 基于轻量卷积神经网络的步态识别系统[J]. 中北大学学报(自然科学版), 2022, 43(5): 474-480.

- [12] 张红颖, 田鹏华. 结合残差网络与多级分块结构的步态识别方法[J]. 电子测量与仪器学报, 2022, 36(6): 66-72.
- [13] 彭小波, 黄海娜, 杨辉跃, 等. 基于行走特征矢量图的步态识别方法[J]. 深圳大学学报(理工版), 2021, 38(5): 528-535.
- [14] 刘瑞华, 郝子赫, 邹洋杨. 基于多层级精细特征融合的步态识别算法[J]. 计算机应用, 2024, 44(7): 2250-2257.
- [15] 王晓路, 千王菲. 基于双支路卷积网络的步态识别方法[J]. 计算机应用, 2024, 44(6): 1965-1971.
- [16] 喻龙, 单紫薇, 赵冬, 等. 武术难点动作图像多目标自动提取方法研究[J]. 自动化技术与应用, 2023, 42(8): 54-57.
- [17] 许庆, 陈文涛, 刘志文. 基于3D视觉分析技术的医院患者跌倒识别方案设计[J]. 自动化技术与应用, 2025, 44(4): 99-102.
- [18] 张博文, 周丽丽. 基于多模态数据建模的生理电人体运动意图预测的研究方法简述[J]. 自动化技术与应用, 2022, 41(3): 192-194.

作者简介: 杨东亮(1984—), 男, 硕士, 助理研究员, 研究方向: 计算机视觉、目标检测。

通信作者: 宋昌江(1983—), 男, 硕士, 研究员, 研究方向: 自动控制。

(上接第36页)

## 3 结束语

综上所述, 本文设计的算法在水电系统的集控预案匹配方面具有广阔的应用前景。通过整合和分析多样化的时控信息和风险因子, 该算法为水电系统提供了科学、高效且自动化的预案匹配策略。它不仅提升了水电系统的运行效率和安全性, 还降低了人为操作误差, 并推动了水电行业的可持续发展。在实际应用中, 这一算法将持续发挥在多源信息整合、风险评估和预案匹配等方面的优势, 为梯级水电系统的远程集控提供强有力的技术支持。随着技术的不断进步和创新, 这一算法将在保障水电系统安全稳定运行、优化资源配置和提升能源利用率方面扮演更加重要的角色。

## 参考文献:

- [1] 辛瑞, 张霄力, 彭侠夫, 等. UFormer: 基于Transformer和U-Net结构的端到端特征点景象匹配算法[J]. 计算机科学, 2023, 50(增刊2): 346-351.
- [2] 盖绍彦, 黄妍妍, 达飞鹏. 基于通道注意力和特征切片的图像快速匹配算法[J]. 光学学报, 2023, 43(22): 158-166.
- [3] 籍雯煊, 崔建业, 冯斌, 等. 基于视觉字符增强的电力调度故障预案匹配[J]. 中国电机工程学报, 2022, 42(15): 5439-5448.
- [4] 赵季红, 罗兴刚, 曲桦, 等. 工业物联网中基于多维特征业务的网络切片匹配算法[J]. 计算机应用研究, 2023, 40(2): 549-553, 570.

- [5] LV Z, PENG R. A novel periodic learning ontology matching model based on interactive grasshopper optimization algorithm[J]. Knowledge-Based Systems, 2021, 44(1): 1320-1334.
- [6] XU X, JIANHUA H, HONG Z, et al. An optimal stability matching algorithm for dag blockchain based on matching theory[J]. Chinese Journal of Electronics, 2021, 25(54): 568-579.
- [7] 王帅, 韩锁, 肖宏宇. 顾及几何与场景结构相似度的管线数据多步骤匹配算法[J]. 测绘科学, 2023, 48(8): 210-219.
- [8] 谷学静, 杨宝上, 刘秋月. 基于高斯滤波和AKAZE-LATCH的图像匹配算法[J]. 半导体光电, 2023, 44(4): 639-644.
- [9] 行芳仪, 徐成, 高宏伟. 高效高精度光照自适应的ORB特征匹配算法[J]. 电子测量与仪器学报, 2023, 37(7): 140-147.
- [10] 宋佳璇, 范大昭, 董杨, 等. 神经网络学习与灰度信息结合的跨视角影像线特征匹配算法[J]. 测绘学报, 2023, 52(6): 990-999.
- [11] 邓建波. 基于图像处理分析的真空泵刮片安装缺陷检测[J]. 自动化技术与应用, 2024, 43(1): 44-48.
- [12] 肖凡雨, 刘泽, 胡泓. 同轴理线机视觉系统设计与分析[J]. 自动化技术与应用, 2023, 42(10): 1-4, 14.

作者简介: 徐伟(1984—), 男, 本科, 工程师, 研究方向: 水电站水库调度。