



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021

04

VOL.26

ISSN1006-8961
CN11-3758/TB



细粒度图像分类 P847



第26卷第4期 (总第300期)

2021年4月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿, 均视为同意在本刊网站及CNKI等全文数据库出版, 所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用, 所有内容, 未经许可, 不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院空天信息创新研究院

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 吴一戎

编辑出版 《中国图象图形学报》编辑部

通信地址 北京市海淀区北四环西路19号

邮 编 100190

电子信箱 jig@radi.ac.cn

电 话 010-58887035

网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

(邮政信箱: 北京399信箱 邮编: 100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Aerospace Information Research Institute, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong

Editor, Publisher Editorial Office of Journal of
Image and Graphics

Address No. 19, North 4th Ring Road West, Haidian District,
Beijing, P. R. China

Zip code 100190

E-mail jig@radi.ac.cn

Telephone 010-58887035

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

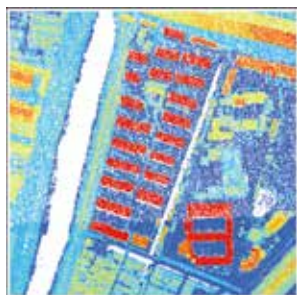
国内定价 60.00元



YOLOv3剪枝模型的多人姿态估计(第0837页)



融入时序和速度信息的自适应更新目标跟踪(第0883页)



可变半径Alpha Shapes提取机载LiDAR点云建筑物轮廓(第0910页)

综述

从图像到语言: 图像标题生成与描述

谭云兰, 汤鹏杰, 张丽, 罗玉盘 0727

CT影像肺结节分割研究进展

董婷, 魏珑, 聂生东 0751

图像处理和编码

全局注意力门控残差记忆网络的图像超分重建

王静, 宋慧慧, 张开华, 刘青山 0766

多层次感知残差卷积网络的单幅图像超分重建

何蕾, 程佳豪, 占志钰, 杨雯博, 刘沛然 0776

采用双尺度图像分解的水下彩色图像增强

李健, 张显斗, 李熹飞, 吴子朝 0787

图像分析和识别

姿态特征结合2维傅里叶变换的步态识别

王新年, 胡丹丹, 张涛, 白桂欣 0796

复杂背景下的手势识别

王银, 陈云龙, 孙前来 0815

结合形变模型与图像修复的人脸姿态矫正

吴从中, 郑荣生, 臧怀娟, 刘明威, 徐甲甲, 詹曙 0828

YOLOv3剪枝模型的多人姿态估计

蔡哲栋, 应娜, 郭春生, 郭锐, 杨鹏 0837

YOLOv3和双线性特征融合的细粒度图像分类

闫子旭, 侯志强, 熊磊, 刘晓义, 余旺盛, 马素刚 0847

深度多模态融合服装风格检索

苏卓, 柯司博, 王若梅, 周凡 0857

超像素条件随机场下的RGB-D视频显著性检测

李贝, 杨铀, 刘琼 0872

融入时序和速度信息的自适应更新目标跟踪

尹宽, 李均利, 胡凯, 李丽 0883

融合逆密度函数与关系形状卷积神经网络的点云分析

邱云飞, 朱梦影 0898

图像理解和计算机视觉

可变半径Alpha Shapes提取机载LiDAR点云建筑物轮廓

伍阳, 王丽妍, 胡春霞, 程亮 0910

3D遮挡模型引导的光场图像深度获取

吴迪, 张旭东, 张骏, 范之国, 孙锐 0924

计算机图形学

基于全正基的非均匀三次加权B样条

姬佩佩, 张贵仓, 汪凯, 孟建军 0939

遥感图像处理

LBP特征分类的极化SAR图像机场跑道检测

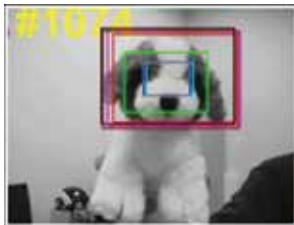
韩萍, 万义爽, 刘亚芳, 韩宾宾 0952

CONTENTS

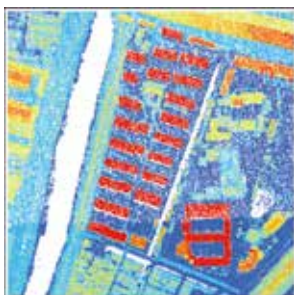
JOURNAL OF IMAGE AND GRAPHICS



Research on multiperson pose estimation combined with YOLOv3 pruning model(P0837)



Adaptive update object tracking algorithm incorporating timing and speed information(P0883)



Extraction of building contours from airborne LiDAR point cloud using variable radius Alpha Shapes method(P0910)

Review

From image to language: image captioning and description

Tan Yunlan, Tang Pengjie, Zhang Li, Luo Yupan 0727

Research progress of lung nodule segmentation based on CT images

Dong Ting, Wei Long, Nie Shengdong 0751

Image Processing and Coding

Learning global attention-gated multi-scale memory residual networks for single-image super-resolution

Wang Jing, Song Huihui, Zhang Kaihua, Liu Qingshan 0766

Single image super-resolution reconstruction based on multi-level perceptual residual convolutional network

He Lei, Cheng Jiahao, Zhan Zhiyu, Yang Wenbo, Liu Peiran 0776

Underwater color image enhancement based on two-scale image decomposition

Li Jian, Zhang Xiandou, Li Shangfei, Wu Zizhao 0787

Image Analysis and Recognition

Gait recognition using pose features and 2D Fourier transform

Wang Xinnian, Hu Dandan, Zhang Tao, Bai Guixin 0796

Hand gesture recognition in complex background

Wang Yin, Chen Yunlong, Sun Qianlai 0815

Face pose correction based on morphable model and image inpainting

Wu Congzhong, Zheng Rongsheng, Zang Huaijuan, Liu Mingwei, Xu Jiajia, Zhan Shu 0828

Research on multiperson pose estimation combined with YOLOv3 pruning model

Cai Zhedong, Ying Na, Guo Chunsheng, Guo Rui, Yang Peng 0837

Fine-grained classification based on bilinear feature fusion and YOLOv3

Yan Zixu, Hou Zhiqiang, Xiong Lei, Liu Xiaoyi, Yu Wangsheng, Ma Sugang 0847

Fashion style retrieval based on deep multimodal fusion

Su Zhuo, Ke Sibao, Wang Ruomei, Zhou Fan 0857

RGB-D video saliency detection via superpixel-level conditional random field

Li Bei, Yang You, Liu Qiong 0872

Adaptive update object tracking algorithm incorporating timing and speed information

Yin Kuan, Li Junli, Hu Kai, Li Li 0883

Point cloud analysis model combining inverse density function and relation-shape convolution neural network

Qiu Yunfei, Zhu Mengying 0898

Image Understanding and Computer Vision

Extraction of building contours from airborne LiDAR point cloud using variable radius Alpha Shapes method

Wu Yang, Wang Liyan, Hu Chunxia, Cheng Liang 0910

Light field depth estimation guided by 3D occlusion model

Wu Di, Zhang Xudong, Zhang Jun, Fan Zhiguo, Sun Rui 0924

Computer Graphics

Non uniform weighted cubic B-spline based on all positive basis

Ji Peipei, Zhang Guicang, Wang Kai, Meng Jianjun 0939

Remote Sensing Image Processing

Airport runway detection based on LBP feature classification in PolSAR images

Han Ping, Wan Yishuang, Liu Yafang, Han Binbin 0952

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2021)04-0796-19

论文引用格式: Wang X N, Hu D D, Zhang T and Bai G X. 2021. Gait recognition using pose features and 2D Fourier transform. Journal of Image and Graphics, 26(04): 0796-0814 (王新年, 胡丹丹, 张涛, 白桂欣. 2021. 姿态特征结合 2 维傅里叶变换的步态识别. 中国图象图形学报, 26(04): 0796-0814) [DOI:10.11834/jig.200061]

姿态特征结合 2 维傅里叶变换的步态识别

王新年¹, 胡丹丹¹, 张涛², 白桂欣¹

1. 大连海事大学, 大连 116026; 2. 辽宁师范大学, 大连 116029

摘要: **目的** 针对现有步态识别方法易受携带物品、衣服变化等影响的问题, 提出了将无肩姿态能量图、步态参数等姿态特征与步态参数的 2 维傅里叶变换相结合的步态识别算法。 **方法** 基于姿态关节点序列提出忽略肩膀宽度信息的无肩姿态能量图, 用以减弱衣服变化的影响; 由于下肢受衣物及背包影响较小, 提取 3 个或 3 个以上的下肢关节点局部结构参数, 即提取中臀点与左右膝关节点、中臀点与左右踝关节点构成的两个三角形面积以及所有下肢关节点构成的多边形面积作为步态参数, 增强下肢参数在步态识别中的作用; 人在行走时, 单肢体的运动具有一定的周期性, 且肢体之间运动具有一定的协调性, 用步态参数的 2 维幅度谱来表示单肢体运动的周期性与肢体之间运动的协调性, 以提高步态参数的可区别性; 在现有典型步态特征的基础上, 融合本文提出的无肩姿态能量图、步态参数及其 2 维傅里叶变换幅度谱, 采用多特征表示步态的方法, 充分利用各特征的优点, 提出加权平均与最大池化相结合的两层分数融合策略进行步态识别, 提高了步态识别算法在携带物品、衣服变化和跨视角等条件下的正确率。 **结果** 实验结果表明, 在中国科学院自动化研究所发布的步态数据集 CASIA-B 上, 本文方法在相同视角条件下, 正常状态、背包状态和穿大衣状态的平均识别率分别为 99.56%、99.23% 和 94.25%; 在跨视角条件下, 正常状态、背包状态和穿大衣状态的平均识别率分别为 91.32%、85.34% 和 69.51%。与典型算法相比, 穿大衣状态的识别率有显著提升。 **结论** 本文方法采用加权平均与最大池化相结合的两层分数融合策略, 综合利用各特征的优点及其适用场景, 有效提高了步态识别的准确率, 特别是减弱了衣服厚度、样式等变化对步态识别的影响。 **关键词:** 步态识别; 姿态特征; 无肩姿态能量图; 运动的协调性与周期性; 局部结构步态参数; 步态参数的 2 维傅里叶变换; 两层分数融合策略

Gait recognition using pose features and 2D Fourier transform

Wang Xinnian¹, Hu Dandan¹, Zhang Tao², Bai Guixin¹

1. Dalian Maritime University, Dalian 116026, China; 2. Liaoning Normal University, Dalian 116029, China

Abstract: **Objective** Gait recognition aims to identify and verify individuals on the basis of walking postures. The performance of existing gait recognition methods is easily influenced by factors such as viewing variances, clothing changes, and types of objects carried by a person. Furthermore, none of these methods consider that the coordination and periodicity of human walking are also important features for gait recognition. Therefore, we propose the pose energy map without considering shoulders (PEMoS) to reduce the effect of clothing changes and 2D Fourier transform magnitude spectrum of gait parameters (2DFoMS) to enhance the effect of the coordination and periodicity of human movements. As these features have a close relationship with the human pose, we call them pose features. Moreover, the proposed pose features are fused

收稿日期: 2020-03-16; 修回日期: 2020-07-31; 预印本日期: 2020-08-07

基金项目: 大连市科技创新基金项目 (2019J12GX036)

Supported by: Dalian Science and Technology Innovation Fund (2019J12GX036)

together with other excellent features such as GaitSet to improve the overall performance of gait recognition. **Method** Clothing changes can affect the detected positions of body joints, especially the shoulder joints. Therefore, we propose PEMoS, which ignores the shoulder width, to reduce the effect of clothing changes. The construction process of PEMoS is as follows: First, body joints in each frame are detected by pose estimation methods. Second, six upper limb joints, namely, RShoulder (right shoulder), RElbow (right elbow), RWrist (right wrist), LShoulder (left shoulder), LElbow (left elbow), and LWrist (left wrist), are horizontally shifted with the displacement between the neck joint and the right or left shoulder joint, whereas the rest remain unchanged. Third, the pose binary map is formatted by connecting the corrected joints in a predefined order and width. Then, it is resized to 128×88 pixels centering on the MidHip joint. Fourth, PEMoS is computed by averaging pose binary maps within a period that include at least one complete gait cycle. Finally, PEMoS is activated by gamma transformation to improve the performance. 2DFoMS uses the coordination between human movements and the periodicity of one single movement to enhance the gait recognition performance. As the lower limbs are less affected by clothing or bags, three new gait parameters computed from the lower limbs are proposed, which include the area of the triangle formed by MidHip (middle hip), LKnee (left knee), and RKnee (right knee); the area of the triangle formed by MidHip, LAnkle (left ankle), and RAnkle (right ankle); and the area of the polygon enclosed by all the lower limb joints. Unlike the existing gait parameters that only consider the relationship between two joints, the proposed area parameters consider the local structural relationship of more than three points, which can enhance the effect of lower limb joints to the gait. The proposed three parameters are concatenated with the other 16 gait parameters extracted by regular methods to form a gait parameter column vector in each frame. The gait parameter column vectors of successive frames over time form a two-dimensional gait parameter matrix. As the gait parameters vary with time, they should be aligned to facilitate training and matching. On the basis of the observation that the width between two ankles varies with time much more regularly than other gait parameters, we propose to use the peak value position of the ankle width curve as base points to align all other gait parameters. 2DFoMS is computed by applying 2D Fourier transform on the registered gait parameter matrix. To fully utilize the advantages of the proposed features and the state-of-the-art features, such as GaitSet, two-level score fusion based on weighted average-max pooling is proposed to compute the matching score. The first level is the weighted average scores of multiple features. At this level, three groups of weights applicable to three different scenarios are proposed to compute the scores. At the second level, the max pooling of the first level scores is used as the final matching score. **Result** We evaluate the proposed method in terms of four aspects: overall performance, performance under different walking conditions, performance under cross-view conditions, and ablation study. In CASIA Gait Database B, the first 62 subjects are used for training, and the remaining 62 subjects are used for testing. The experimental results show that the proposed method achieves average accuracies of 99.56%, 99.23%, and 94.25% under walking normally, walking with a bag, and walking with a different coat, respectively, when the views of the probe sequence and its counterpart in the gallery set are the same. For cross-view recognition, the proposed method achieves average accuracies of 91.32%, 85.34%, and 69.51% under walking normally, walking with a bag, and walking with a different coat, respectively. Compared with the state-of-the-art methods, the average accuracy of the proposed method increases by about 6.98% under clothing changing conditions. Ablation study shows that PEMoS and 2DFoMS are effective in improving the gait recognition accuracy. **Conclusion** The proposed PEMoS, which ignores the shoulder width, can increase the robustness and accuracy of gait recognition under clothing changing conditions. The proposed local structure gait parameters can enhance the effect of lower limb points to the gait, which are more robust to clothing or bag changes than the upper limb parameters. 2DFoMS can emphasize the effect of coordination between human movements and the periodicity of one single moment on the performance of gait recognition. The experimental results show that the proposed algorithm has achieved state-of-the-art performance, especially under clothing or bag changing conditions.

Key words: gait recognition; pose feature; pose energy map without considering shoulders; coordination and periodicity of human movement; local structure gait parameters; 2D Fourier transform on gait parameters; two level score fusion

0 引言

步态是指人走路的方式,与传统的生物特征如指纹、虹膜和人脸相比,具有非接触性、难伪装和远距离等优点,在视频监控中受到越来越多的重视。步态识别的方法可分为基于人体外观形状的方法和基于人体结构模型的方法。

基于人体外观形状的方法是指直接从步态视频序列提取人体外观轮廓、步态侧影图等特征进行步态识别。Kale 等人(2004)提出隐马尔可夫模型,该方法需要逐帧比较,计算量大。Han 和 Bhanu (2006)将一个周期内的步态侧影图求和取平均得到的步态能量图(gait energy image, GEI)作为步态特征,充分利用步态特征之间的连续性和动态性。Lam 等人(2011)改进了 GEI 特征,提出步态光流图(gait flow image)。基于 GEI 的方法具有较好的性能,但人体的外观形状易受衣服变化和携带物品等影响,为此,一些研究人员致力于研究减少形状变化和遮挡的干扰。Huang 和 Boulgouris (2012)融合步态结构轮廓(gait structure profile)和重心偏移校正能量图(shifted energy image)特征,提高携带物品可能导致身体重心发生位移等变化时的识别性能。Yu 等人(2016, 2017a)提出的堆叠式自动编码器模型(stacked progressive AutoEncoders, SPAE), Yu 等人(2017b, 2019)提出的步态生成式对抗网络模型(gait generative adversarial network, GaitGAN),将任意视角、不同衣服或携带物品的 GEI 序列转化为指定角度、正常状态的 GEI 序列。Chao 等人(2019)提出基于步态侧影图集合(GaitSet)的方法,首次将不同视角、不同行走状态的步态侧影图组合成无时序关系的集合,利用深度神经网络优化集合中步态图像的内在关系。上述步态识别方法在外观形状没有改变的条件下取得了较好的识别结果,但人体的外观形状易受诸多因素的影响,如携带物品和衣服等,进而影响步态识别的准确性。

基于人体结构模型的方法是指对人体结构或运动方式进行建模,提取模型拟合的参数作为步态特征进行识别,在处理噪声、遮挡等方面的性能更好。Cunado 等人(1997)提出钟摆模型,使用霍夫变换提取腿部线条作为步态特征。Zhang 等人(2007)提出“五连杆双足”运动模型,使用 Metropolis-Hasting 方

法从步态视频序列提取步态特征。大多数基于人体结构模型的方法需预先手动或使用特定设备来获取人体关键点,如 Yang 等人(2016)、Kang 等人(2017)和 Zhen 等人(2018)使用 Kinect 传感器获取,但因其精度较低,步态识别率不是很高。Feng 等人(2016)利用姿态估计方法提取身体关节热图描述步态信息。Liao 等人(2017)提出基于人体 2D 姿态关节节点的时空网络(pose-based temporal-spatial network, PTSN), PTSN 包括卷积神经网络(convolutional neural network, CNN)和长短期记忆(long short-term memory, LSTM)两路网络,网络中 CNN 主要提取步态中的静态特征, LSTM 提取步态中的动态特征。之后, Liao 等人(2020)提出基于人体 3D 姿态的 PoseGait 模型,通过姿态特征解决视角方面的问题。不同于上述将步态视频序列每帧的姿态关节节点直接输入进行步态识别, Yao 等人(2018)将预先定义的关节点和线宽表示的骨架步态能量图(skeleton gait energy image)用于步态识别。基于姿态估计方法检测到的姿态关节节点易受衣服厚度、样式等影响,导致衣服改变条件下步态识别的性能不是很高。

生活中,人在行走过程中衣服状态(如厚度、长度等)和携带物品类型(如双肩包、斜挎包等)都会发生改变,从而造成步态识别的性能不是很高。因此,本文提出了将无肩姿态能量图、步态参数等姿态特征与步态参数的 2 维傅里叶变换相结合的步态识别算法。主要工作包括以下 4 个方面:1) 基于人体关节节点序列,提出忽略肩膀宽度信息的无肩姿态能量图,用以减弱衣服变化对步态识别的影响;2) 由于下肢参数受衣服或携带包裹影响较小,提出基于下肢关节节点的局部结构参数,用来增强下肢行为在步态识别中的作用;3) 提出用步态参数的 2 维傅里叶变换的幅度谱来增强行走过程中单肢体运动的周期性和肢体之间运动的协调性关系在步态表示中的贡献;4) 采用多特征表示人体步态,在基于外观形状的典型方法 GaitSet 网络特征和基于姿态关节节点的有肩姿态能量图基础上,结合本文提出的无肩姿态能量图、步态参数及其 2 维傅里叶变换幅度谱特征,综合利用各特征的优点,提出了加权平均与最大池化相结合的两层分数融合策略进行步态识别,有效提高了步态识别的准确率。

1 基于无肩姿态能量图的步态表示

基于姿态估计的步态识别方法易受衣服变化影响。图1是正常、背包和穿大衣状态的有肩姿态能量图比较。其中,图1(a)是不同状态的步态视频;图1(b)是采用姿态估计方法(Cao等,2019)提取的人体姿态关节点示意图;图1(c)是将姿态关节点以2个像素线宽并按照设定的顺序连接形成的有肩姿态二值图;图1(d)是有肩姿态能量图(pose energy map considering shoulders, PEMS),其定义为一段时间内有肩姿态二值图的均值;图1(e)(f)是将正常

状态与其他两种状态的有肩姿态能量图的边缘图叠加。可以看出,背包状态的肩膀宽度特征与正常状态较相似,而穿大衣状态下的有肩姿态能量图上身的外观形状整体偏宽,主要因为衣服变化导致检测的上身关节点(尤其肩膀关节点)位置发生改变,从而造成穿大衣状态的有肩姿态能量图比正常状态偏宽。为此,提出忽略肩膀宽度特征的无肩姿态能量图,用以减弱衣服变化对步态识别的影响。

基于无肩姿态能量图的步态表示分为提取无肩姿态能量图和伽马变换两个阶段,前者用于减弱衣服改变对步态识别的影响,后者用于增强图像低灰度区域的对比度。

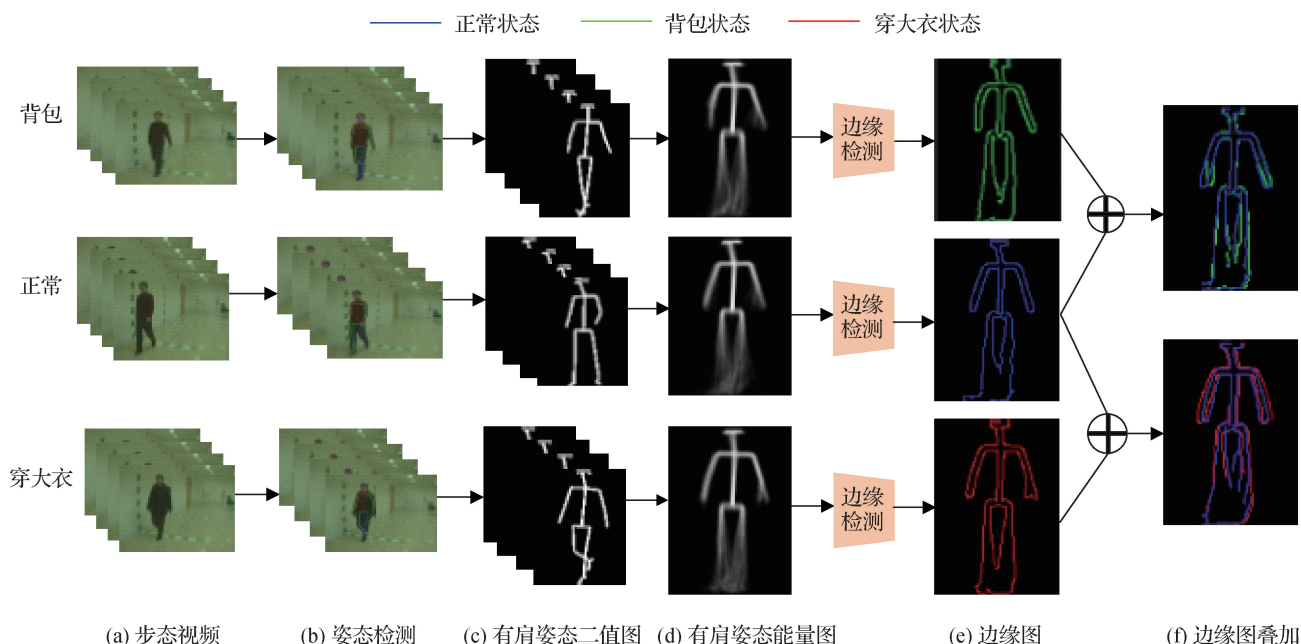


图1 正常、背包和穿大衣状态有肩姿态能量图比较

Fig. 1 Comparison of PEMSs among walking normally, walking with a bag and walking with a different coat

((a) gait videos; (b) pose estimation; (c) binary maps with considering shoulders; (d) pose energy maps considering shoulders; (e) edge maps; (f) overlay of edge maps)

1.1 无肩姿态能量图的提取

首先,使用姿态估计方法(Cao等,2019)提取25个有肩姿态关节点,如图2所示。

然后,以颈部关节点(Neck)和中臀点(MidHip)连线为轴,将右肩(RShoulder)、右手肘(RElbow)、右手腕(RWrist)、左肩(LShoulder)、左手肘(LElbow)和左手腕(LWrist)6个关节点的坐标加上肩膀关节点和颈关节点的差向量进行坐标变换,即

$$\begin{aligned} x'_i &= x_i + x_{\text{Neck}} - x_{\text{RShoulder}} \\ x'_j &= x_j + x_{\text{Neck}} - x_{\text{LShoulder}} \end{aligned} \quad (1)$$

式中, x 为关节点的横坐标, x' 为移动后的关节点的横坐标, $i \in \{\text{RShoulder}, \text{RElbow}, \text{RWrist}\}$, $j \in \{\text{LShoulder}, \text{LElbow}, \text{LWrist}\}$ 。

坐标变换后,得到23个无肩姿态关节点示意图,如图3所示。

将图3中的所有姿态关节点以2个像素线宽的直线连接,并以MidHip为中心,归一化为 128×88 像素,得到无肩姿态二值图 $B(r, c)$ 。

无肩姿态能量图(pose energy maps without considering shoulders, PEMoS)为一段时间内各帧提取

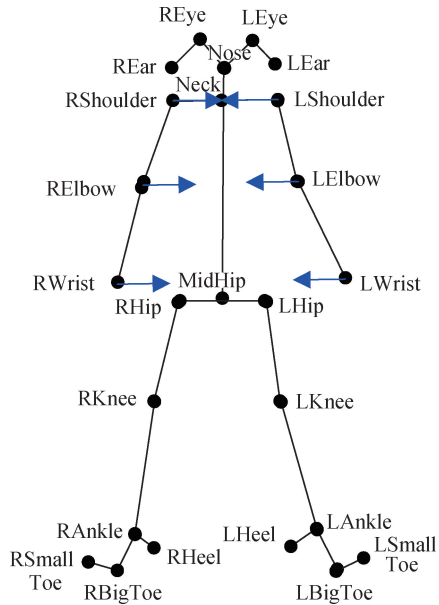


图2 25个有肩关节示意图

Fig. 2 Illustration of twenty-five body joints with shoulders

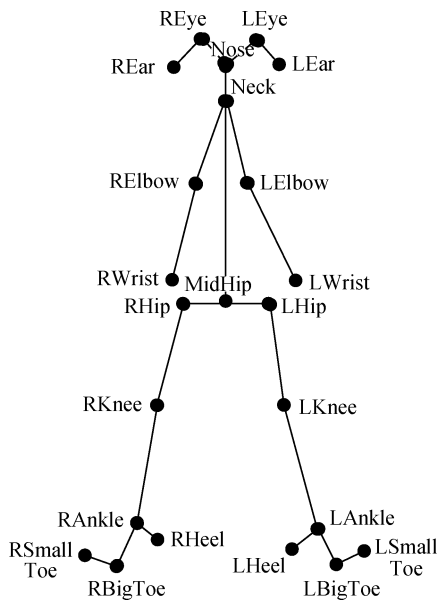


图3 23个无肩关节示意图

Fig. 3 Illustration of twenty-three body joints without shoulders

的无肩姿态二值图的均值,即

$$G(r, c) = \frac{1}{T} \sum_{t=1}^T B_t(r, c) \quad (2)$$

式中, $t \in [1, T]$, T 为步态视频序列的帧数, $B_t(r, c)$ 表示第 t 帧的无肩姿态二值图, (r, c) 为图像中像素的坐标。无肩姿态能量图反映人体胳膊和腿部的形状及其在步态周期中的变化,灰度值越高,表示相应部位在该位置的运动频率越高。图4为不同状态的无肩姿态二值图,图中最右端为对应的无肩姿态能量图。

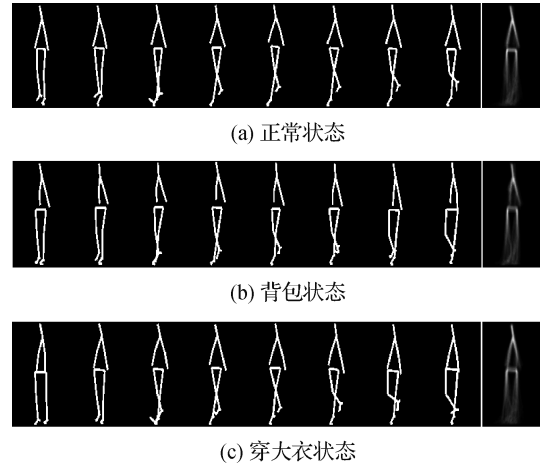


图4 不同状态的无肩姿态二值图及对应的无肩姿态能量图

Fig. 4 Examples of binary maps and PEMoSs under different walking states ((a) walking normally; (b) walking with a bag; (c) walking with a different coat)

图5为不同状态下有肩和无肩两种姿态能量图的边缘。图5(a)是按图1方式将正常状态和穿大衣状态下的无肩姿态能量图的边缘叠加。与图5(b)相比,图5(a)中人体上身的形状特征在正常状态和穿大衣状态下更加相似。

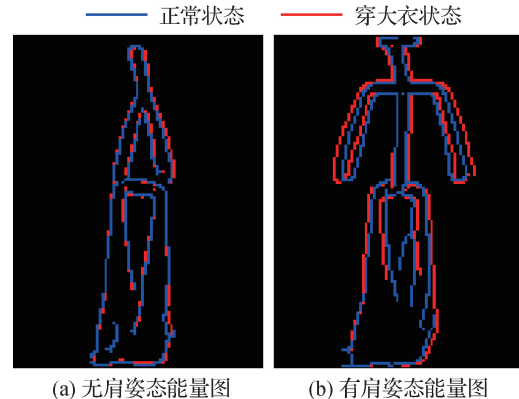


图5 不同状态下两种姿态能量图的边缘

Fig. 5 Edge maps of two kinds of pose energy maps under different walking states ((a) PEMoS; (b) PEMS)

1.2 无肩姿态能量图的伽马变换

为增强无肩姿态能量图中低灰度区域的对比度,将无肩姿态能量图进行伽马变换,具体为

$$G^*(r, c) = C \times G(r, c)^\lambda \quad (3)$$

式中, $G(r, c)$ 和 $G^*(r, c)$ 分别为伽马变换前后的无肩姿态能量图, C 为常数, λ 为伽马变换的指数。实验结果表明,当 $C = 1$, $\lambda = 0.4$ 时,图像增强效果最好。图6为中国科学院自动化研究所发布的步

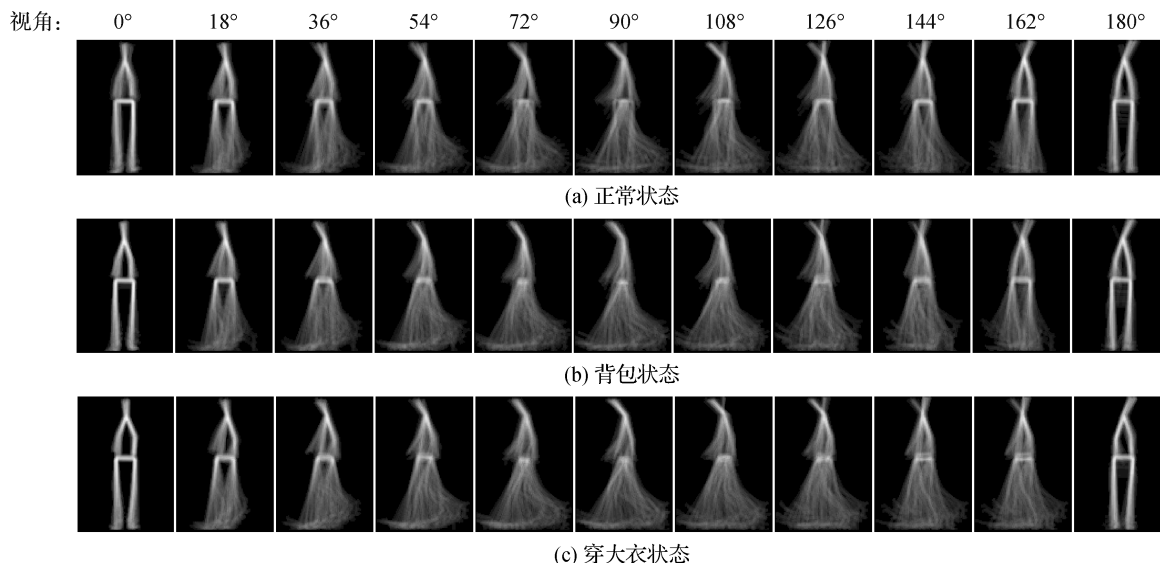


图6 不同状态下11个视角的无肩姿态能量图

Fig. 6 PEMoS in eleven view angles under different walking states

((a) walking normally; (b) walking with a bag; (c) walking with a different coat)

态识别数据集 CASIA-B 中正常状态、背包状态和穿大衣状态下,11 个视角经过伽马变换后的无肩姿态能量图。

2 基于步态参数 2 维傅里叶变换的步态表示

基于步态参数的方法通常考虑步态参数随时间变化的关系,没有考虑行走过程中多肢体运动之间的协调性关系。为此,本文提出步态参数的 2 维傅里叶变换幅度谱特征,用以表示单步态参数的周期性以及步态参数之间的协调性关系。

基于步态参数 2 维傅里叶变换的步态表示分为步态参数提取与对齐和步态参数的 2 维傅里叶变换幅度谱计算两个阶段。第 1 阶段的目的在于提取步态参数并去除同一人不同序列存在的时间差影响;第 2 阶段的目的在于表示人在行走过程中单肢体运动的周期性和各肢体之间运动的协调性关系。

2.1 步态参数提取与对齐

2.1.1 步态参数提取

Kang 等人(2017)用基于两点之间关系的步态参数表示人体步态,即左右两个胳膊关节角度、左右腿部关节角度、关节间高度和距离参数,分别记为 $\theta^1, \theta^2, \theta^3, \theta^4, \theta^5, \theta^6, \beta^1, \beta^2, \beta^3, \beta^4, \beta^5, \beta^6, W^1, W^2, H^1$ 和 H^2 ,如图 7(a)(b)所示。其中 W^1 表示 RAnkle 与

LAnkle 的横坐标之差, W^2 表示 RKnee 与 LKnee 的横坐标之差,由于行走过程中左右腿交替在前, W^1 和 W^2 会出现负值。

由于下肢关节受衣服或携带包裹的影响较小,本文提出基于 3 个及以上的下肢关节局部结构参数来增强下肢运动行为在步态识别中的作用。3 个下肢局部结构参数包括中臀点(MidHip)与左右膝关节点(LKnee、RKnee)、中臀点(MidHip)与左右踝关节点(LAnkle、RAnkle)构成的两个三角形面积和所有下肢关节点右臀(RHip)、左臀(LHip)、右膝盖(RKnee)、左膝盖(LKnee)、右脚踝(RAnkle)、左脚踝(LAnkle)构成的多边形面积,分别记为 S^1 、 S^2 和 S^3 。图 7(c)(d)为下肢局部结构参数示意图(阴影部分面积)。

步态参数的真实匹配和虚假匹配分数分布图如图 8 所示。图 8(a)是参照 Kang 等人(2017)提取的 16 个步态参数的真实匹配和虚假匹配分数分布曲线,图 8(b)是结合增加的 3 个下肢局部结构参数的真实匹配和虚假匹配分数分布曲线。真实匹配分数与虚假匹配分数分布曲线下的重叠面积越小,表示该特征的可区分度越高。图 8(b)曲线重叠区域(阴影)面积小于图 8(a),说明增加的 3 个局部结构参数有助于提高特征的区分度。

2.1.2 步态参数对齐

步态参数对齐包括参考点选择、规定长度的步

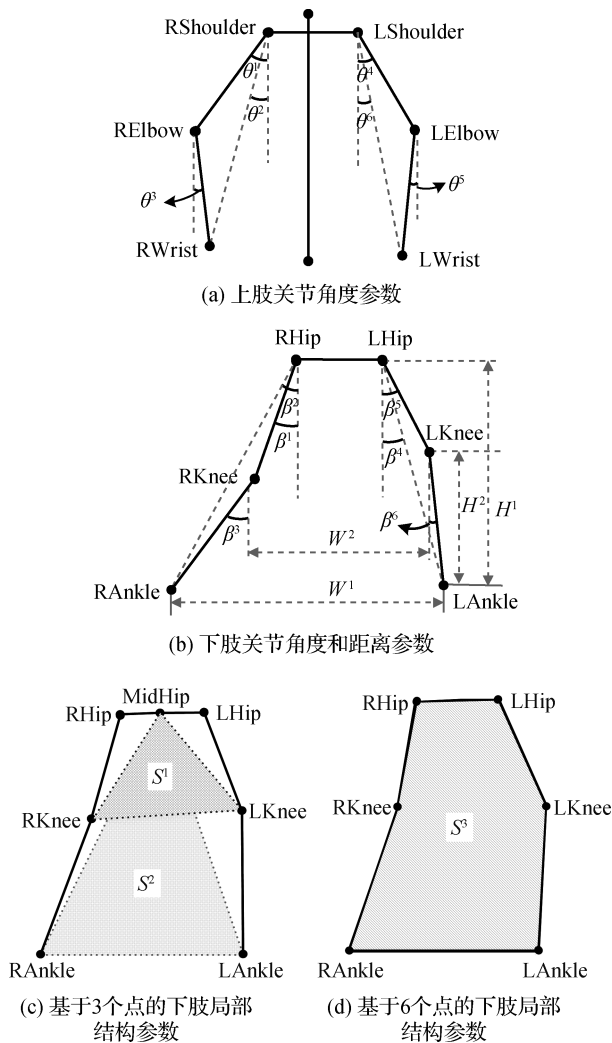


图7 步态参数示意图

Fig. 7 Illustration of gait parameters ((a) joint angle parameters extracted from upper limbs; (b) joint angle and distance parameters extracted from lower limbs; (c) the proposed local structure parameters extracted from three lower limb joints; (d) the proposed local structure parameters extracted from six lower limb joints)

态参数序列提取和步态参数取值范围归一化。由于步态参数随着时间变化而变化,因此在下文中用粗斜体表示步态参数对应的时间序列,如步态参数 W^1 的时间序列用 $\mathbf{W}^1$ 表示,由 W^1 在各时刻的值构成,即 $\mathbf{W}^1 = [W_1^1, \dots, W_t^1, \dots]$,其中 W_t^1 表示 W^1 在 t 时刻的值。示意图如图9所示。

对步态参数进行对齐处理,主要原因是:1) 每个人行走的步伐速率不同,收集的步态数据长度可能不同,导致提取的步态参数长度也不同,为便于后续模型的训练与使用,需将数据根据一定规则选择相同数目的数据点;2) 在进行特征匹配时,同一个

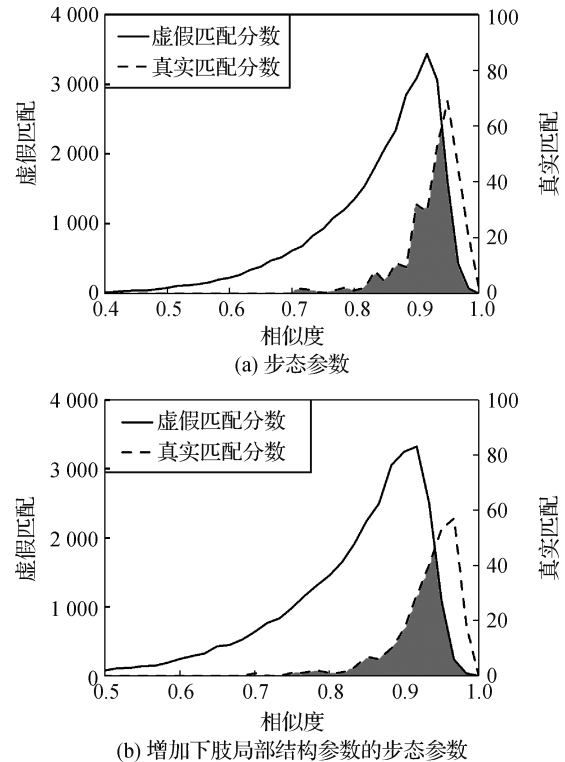


图8 步态参数的真实匹配和虚假匹配分数分布

Fig. 8 The genuine and imposter matching score distributions of gait parameters ((a) gait parameters; (b) gait parameters with the proposed local structure parameters from lower limb joints)

人不同序列的步态参数可能存在时间差。如图9(a)所示, A 点表示序列1的波峰点, A_1 点表示序列2中与 A 点相同时刻对应的数据点, A_2 点表示序列2中与 A 对应的波峰点,由于时移影响,导致 A 和 A_1 点匹配,但实际上 A 和 A_2 点特征匹配度更高,为此也需要对步态参数进行对齐。

具体步骤如下:

1) 参考点选择。由于人的脚部受外界遮挡影响较小,姿态检测结果相对稳定,因此基于表示两脚踝关节横坐标之差的时间序列 $\mathbf{W}^1$ 提取步态参数对齐的参考点。对时间序列 $\mathbf{W}^1$ 进行中值和均值滤波处理,去除异常点和高斯噪声;在滤波处理后的序列中,找出局部极大值点对应的时刻序列,记为 $\mathbf{a} = (a_1, a_2, \dots, a_k)$,其中 $\mathbf{a}$ 中各分量按时间顺序先后排列。由于开始行走时步态参数往往不稳定,因此将时间最晚的 a_k 作为参考点。

2) 规定长度的步态参数序列提取。根据选择的参考点,从各个步态参数时间序列中,均匀提取 T 个步态参数,形成对应的步态参数序列。图9(b)中 B 点表示参考点, C 点表示参考点向前(时间早

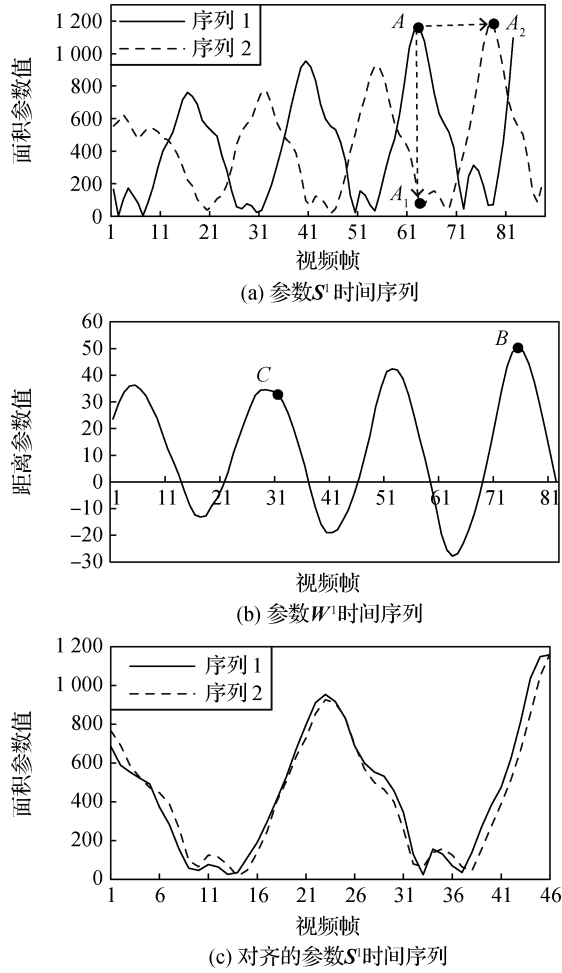


图9 步态参数对齐示意图

Fig. 9 Illustration of aligned gait parameters

((a) the time series of S^1 ; (b) the time series of W^1 ; (c) the aligned time series of S^1)

的方向)移动 $T-1$ 个采样点对应时刻。若步态参数个数不足 $T-1$ 个,则采用外插的方法估计出缺少的数据点。外插的方法分为线性方法和非线性方法,尽管非线性方法估计的精度略高,但实验发现对识别率影响不大,因此为便于计算,采用线性插值方法。 T 表示一个步态周期包含的采样点个数,与步频和帧频有关。据统计,成人平均自然步频约为 95~125 步/min,即正常行走一步需要 0.48~0.63 s,一个步态周期包括 3 步,因此 T 可选择 1.4~1.9 倍帧频之间的一个整数,本文实验基于 CASIA-B 数据集,其帧频是 25 帧/s,经过实验,选择 1.84 倍的帧频,即 $T=46$ 。

3) 步态参数取值范围归一化。由于各个步态参数序列的量纲不一致,为便于后续步态参数的

2 维幅度谱特征提取和匹配,采用最值归一化方法 (min-max normalization) 将每个步态参数序列归一化到 $[0,1]$ 之间。图 9(c) 是图 9(a) 对齐步态参数后的两个面积参数序列。可以看出,与图 9(a) 相比,对齐后的步态参数去除了时移的影响。

2.2 步态参数的 2 维傅里叶变换幅度谱特征

由于要考虑单步态参数的周期性关系以及步态参数之间的协调性关系,将胳膊关节角度、腿部关节角度、关节间高度、宽度和面积等 19 个步态参数合并形成一个 2 维步态参数矩阵 P_{pose} , 即

$$P_{pose} = [p_1 \cdots p_t \cdots p_T] \quad (4)$$

$$p_t = \begin{bmatrix} \theta_t \\ \beta_t \\ W_t \\ H_t \\ S_t \end{bmatrix} \quad (5)$$

$$\theta_t = [\theta_t^1, \theta_t^2, \theta_t^3, \theta_t^4, \theta_t^5, \theta_t^6]^T \quad (6)$$

$$\beta_t = [\beta_t^1, \beta_t^2, \beta_t^3, \beta_t^4, \beta_t^5, \beta_t^6]^T \quad (7)$$

$$W_t = [W_t^1, W_t^2]^T \quad (8)$$

$$H_t = [H_t^1, H_t^2]^T \quad (9)$$

$$S_t = [S_t^1, S_t^2, S_t^3]^T \quad (10)$$

式中, $t \in [1, T]$, $T=46$ 为对齐后步态参数的长度, $\theta_t, \beta_t, W_t, H_t, S_t$ 表示各步态参数对应的时间序列。

2 维步态参数矩阵中的每一行表示每一种对齐和规格化后的步态参数序列。矩阵中元素 $P_{pose}(m, t)$ 的值表示第 m 种步态参数、第 t 个采样点的值。本文提取了 19 种步态参数,每种步态参数提取连续 46 个采样点构成了步态参数序列,因此 2 维步态参数的维数为 19×46 。图 10 是两个人的不同步态序列的 2 维参数矩阵,矩阵横轴(列)表示采样点序号,纵轴(行)是步态参数种类标识,图例表示归一化的步态参数矩阵的值。图 11 是部分运动参数的 1 维示意图,图 11(a) 为腿部角度 ($\beta^1, \beta^2, \beta^3$) 的参数示意图,图 11(b) 为不同类型的步态参数 ($\theta^1, \beta^1, W^1, H^1, S^1$) 的示意图。

从图 10 和图 11 可以看出:1) 每个步态参数呈一定的周期性变化,体现了人走路动作的周期性,本文用行方向的频率和幅度表示;2) 步态参数之间或呈同步递增/减变化或呈正交变化(变化方向相反)或差值相对固定,这体现了肢体之间运动的协调性,

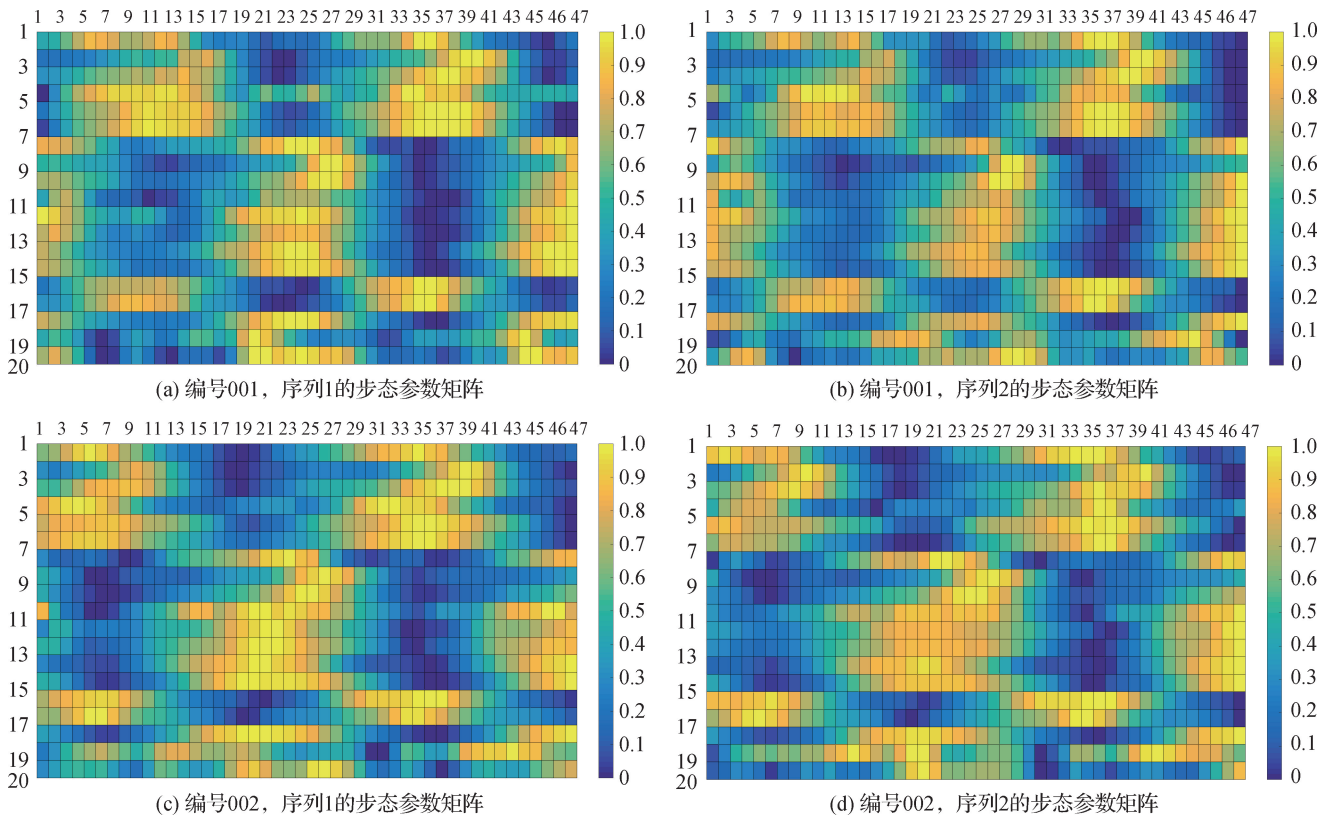


图 10 两个人的 2 维步态参数矩阵示意图

Fig. 10 Illustration of two-dimensional gait parameters matrix of two persons((a) gait parameters matrix of sequence #1 from person #001; (b) gait parameters matrix of sequence #2 from person #001; (c) gait parameters matrix of sequence #1 from person #002; (d) gait parameters matrix of sequence #2 from person #002)

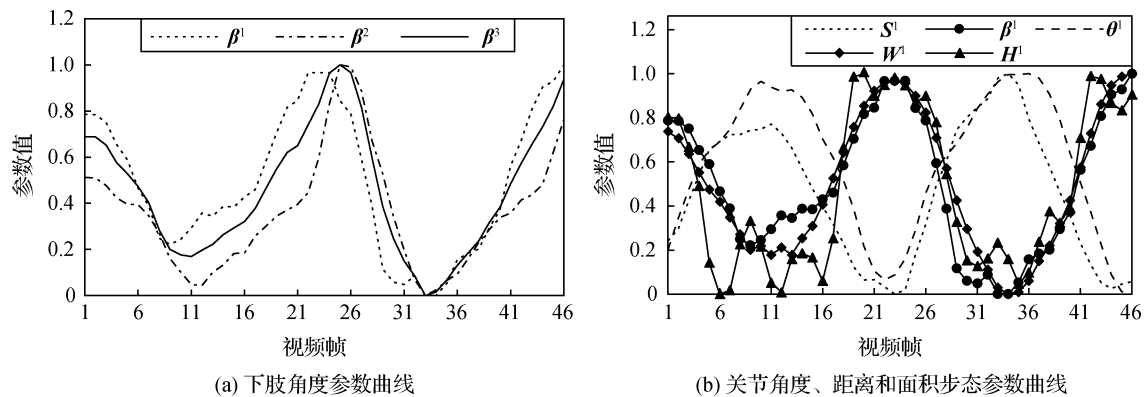


图 11 部分步态参数的时间序列曲线

Fig. 11 Time series curves of some gait parameters((a) curves of joint angle parameters from lower limbs; (b) curves of joint angle, distance and area parameters)

本文用列方向的频率和幅度表示。基于此,提出步态参数矩阵 $\mathbf{P}_{\text{pose}}$ 的 2 维傅里叶变换幅度谱 ($\mathbf{F}_{\text{pose}}$) 表示步态的周期性和协调性,即

$$\mathbf{F}_{\text{pose}}(u, v) = \left| \frac{1}{MT} \sum_{m=0}^{M-1} \sum_{t=0}^{T-1} \mathbf{P}_{\text{pose}}(m, t) e^{-j2\pi(\frac{um}{M} + \frac{vt}{T})} \right| \quad (11)$$

式中, $M = 19$ 为矩阵中步态参数的个数, u, v 分别表示垂直和水平方向频率, j 表示复数。

2 维傅里叶变换幅度谱 $\mathbf{F}_{\text{pose}}$ 既考虑同一步态参数的时间序列(如 $[S_1, S_2, \dots, S_{T-1}, S_T]$)之间的周期性关系,又考虑同一时刻不同步态参数(如 $[\theta_1, \beta_1, W_1, H_1, S_1]$)之间的协调性关系,增加了可区分

性。步态参数幅度谱的真实匹配和虚假匹配分数分布如图12所示。图12(a)(b)分别为2维和1维傅里叶变换幅度谱的真实匹配和虚假匹配分数分布曲线。可以看出,图12(a)中曲线的重叠区域(阴影)面积小于图12(b),说明2维傅里叶变换幅度谱特征相对1维傅里叶变换幅度谱有较好的可区分度。

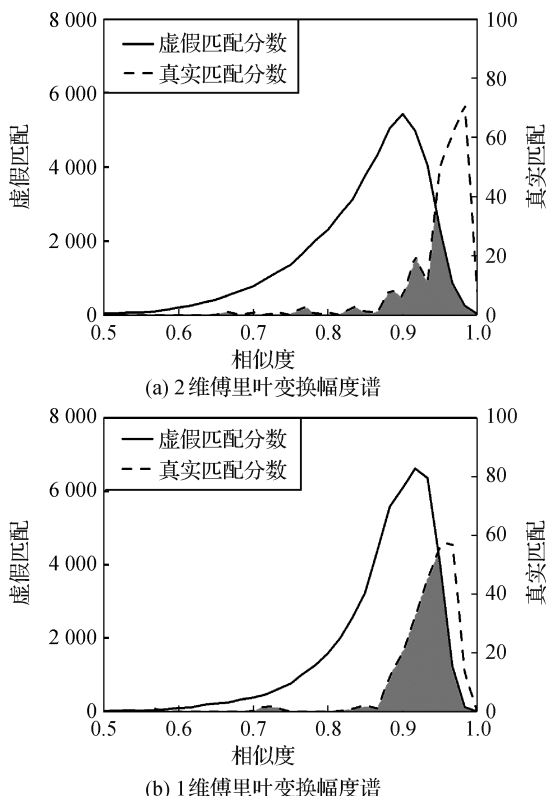


图12 步态参数幅度谱的真实匹配和虚假匹配分数分布
Fig. 12 The genuine and imposter matching score distributions of magnitude spectrums of gait parameters((a) two-dimensional Fourier transform magnitude spectrum; (b) one-dimensional Fourier transform magnitude spectrum)

3 基于两层分数融合的步态识别

为了充分利用本文所提特征在不同行走状态下步态识别的性能,提出了加权平均与最大池化相结合的两层分数融合策略进行步态识别。第1层为得分水平的加权融合,采用适用于不同行走条件的权重分别计算得分;第2层将第1层加权融合得到的分数的最大值作为最终匹配分数。具体过程如图13所示,图中 s_1 、 s_2 、 s_3 和 s_4 分别表示待匹配的两个序列对应特征的余弦相似度。

结合 GaitSet 网络特征(Chao 等,2019),用4种

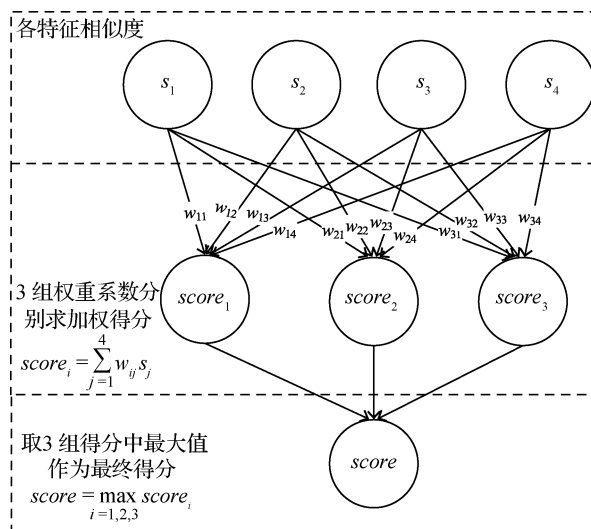


图13 两层分数融合策略

Fig. 13 Two level score fusion strategy

特征表示步态序列,即 GaitSet、无肩姿态能量图、有肩姿态能量图和步态参数及其2维傅里叶变换幅度谱,后3种特征采用视图一致性的多视角判别分析算法(multi-view discriminant analysis with view-consistency, MvDA-VC)(Kan 等,2016)进行特征映射。

图13中, $w_{ij}(i=1,2,3; j=1,2,3,4)$ 为权重系数,取值范围为 $[0,1]$,且满足 $\sum_{j=1}^4 w_{ij} = 1$ 。由于有肩姿态能量图和无肩姿态能量图适用于解决不同的行走场景问题,因此采用3组不同的权重系数加权融合,每组权重参数适用于不同的行走场景,有效利用各步态特征的优点。在正常行走状态,无肩姿态能量图与有肩姿态能量图的权重系数相同,即 $w_{12} = w_{13}$;在携带包裹状态,无肩姿态能量图的权重系数小于有肩姿态能量图的权重系数,即 $w_{22} < w_{23}$;在衣服改变状态,无肩姿态能量图的权重系数大于有肩姿态能量图的权重系数,即 $w_{32} > w_{33}$;在3组权重系数中,GaitSet、步态参数及其2维傅里叶变换幅度谱所占比列相同,即 $w_{11} = w_{21} = w_{31}$, $w_{14} = w_{24} = w_{34}$ 。

4 实验与结果分析

4.1 CASIA-B 步态数据集

CASIA-B 数据集(Yu 等,2006)选自中国科学院自动化研究所提供的 CASIA 步态数据库,共124人(001—124)11个视角($0^\circ, 18^\circ, \dots, 180^\circ$)的步态序列,其中,每个人每个视角有10个步态序列,包括

6 个正常条件(01—06)、2 个穿大衣条件(01—02)和 2 个背包条件(01—02)行走的序列。将 CASIA-B 数据集中前 62 人的步态序列放入训练集,其余 62 人的步态序列放入测试集。具体测试集又分为模型图像集和测试图像集,模型图像集由正常条件 01—04 组成,测试图像集包含以下 3 个部分:正常条件 05—06、背包条件 01—02,穿大衣条件 01—02,具体的实验设置如表 1。

4.2 CASIA-B 数据集上的总体性能比较

在 CASIA-B 数据集上,将本文方法在正常、背包和穿大衣状态下分别进行实验,Rank-1 识别率如表 2—表 4 所示。每种状态有 121 个识别率,平均

表 1 CASIA-B 数据集上的实验设置
Table 1 Experimental setting on CASIA-B dataset

训练集	测试集	
	模型图像集	测试图像集
ID:001—062	ID:063—124	ID:063—124
正常:01—06	正常:01—04	正常:05—06
背包:01—02		背包:01—02
穿大衣:01—02		穿大衣:01—02

识别率为 121 个结果的平均值,用于衡量对应状态下不区分相同或跨视角进行步态识别的总体性能,后续实验分析均依据表 2—表 4 进行。

表 2 正常状态下 CASIA-B 数据集上的 Rank-1 识别率
Table 2 Rank-1 accuracies of our method in normal walking mode on CASIA-B dataset

模型图像集视角 (正常:01—04)	测试图像集视角(正常:05—06)											/%
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
0°	100.00	100.00	99.19	87.10	70.97	66.94	75.81	86.29	95.97	97.58	99.19	
18°	98.39	100.00	100.00	99.19	94.35	87.90	84.68	93.55	96.77	99.19	89.52	
36°	91.94	100.00	100.00	100.00	99.19	94.35	91.94	95.97	96.77	97.58	80.65	
54°	83.06	97.58	97.58	98.39	97.58	97.58	96.77	96.77	95.97	92.74	75.00	
72°	79.03	95.16	100.00	100.00	100.00	99.19	98.39	98.39	94.35	87.90	61.29	
90°	69.35	83.87	95.97	97.58	100.00	99.19	99.19	99.19	95.16	80.65	62.90	
108°	72.58	84.68	96.77	100.00	100.00	100.00	100.00	100.00	98.39	94.35	67.74	
126°	84.68	91.13	95.97	96.77	97.58	97.58	97.58	98.39	98.39	95.97	81.45	
144°	89.52	94.35	99.19	97.58	94.35	95.16	99.19	99.19	99.19	100.00	87.10	
162°	96.77	97.58	98.39	96.77	89.52	87.10	94.35	98.39	99.19	100.00	97.58	
180°	96.77	91.94	85.48	76.61	56.45	55.65	59.68	78.23	84.68	97.58	100.00	
平均识别率	92.07											

注:加粗字体(对角线)表示视角未改变情况下的识别率。

图 14 为本文方法与基于人体外观形状的方法和基于人体结构模型的方法的总体性能比较。基于人体外观形状的方法包括:步态能量图 + 元成分分析(GEI + PCA)(Han 和 Bhanu, 2006)、SPAЕ(Yu 等, 2017a)、GaitGANv1(Yu 等, 2017b)、GaitGANv2(Yu 等, 2019)和 GaitSet(Chao 等, 2019);基于人体结构模型的方法包括 PoseGait(Liao 等, 2020)。除 GaitSet 外,其他方法的结果均取自对应原文。GaitSet 文献中未列出 CASIA-B 数据集的总体性能,为便于比较,本文按照文献设置的参数训练 5 次模型,取 5 次测试结果

的平均值作为后续 GaitSet 的识别率(包括 4.3 和 4.5 节中的结果)。可以发现,正常状态和背包状态下,本文方法的总体性能与 GaitSet 接近,均高于其他方法;穿大衣状态下,本文方法比 GaitSet 提高了 6.26%。

4.3 行走状态改变条件下的性能评估

为对本文方法在行走状态改变条件下的性能进行评估,将本文方法与 SPAЕ、GaitGANv1、GaitGANv2、PoseGait、GaitSet 和 PTSN(Liao 等, 2017)方法在视角不变、行走状态改变(背包和穿大衣状态)条件下的识别率进行比较,结果如表 5 和表 6 所示。

表3 背包状态下 CASIA-B 上的 Rank-1 识别率

Table 3 Rank-1 accuracies of our method in carrying variation mode on CASIA-B dataset

模型图像集视角 (正常:01—04)	测试图像集视角(背包:01—02)											/%
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
0°	99.60	96.77	85.89	72.98	62.90	50.00	57.26	71.37	84.68	93.95	91.53	
18°	96.77	99.19	99.19	93.55	83.06	76.21	80.24	86.69	88.71	92.74	87.90	
36°	83.47	98.79	99.19	98.79	97.98	89.11	92.74	94.35	93.55	90.73	75.81	
54°	72.18	93.95	97.18	97.98	96.77	93.95	93.15	95.56	93.55	82.26	67.34	
72°	62.10	88.31	96.77	97.98	100.00	97.98	95.56	97.18	93.15	77.42	57.66	
90°	60.48	78.63	93.55	97.18	99.19	98.39	96.77	97.18	91.53	73.39	49.60	
108°	62.50	82.26	92.74	95.56	98.39	97.18	99.60	100.00	100.00	84.68	58.06	
126°	72.58	84.68	92.34	94.76	94.35	95.16	97.58	98.39	98.39	91.94	71.77	
144°	82.26	93.55	95.16	92.74	91.94	89.52	93.95	99.19	99.60	97.98	80.65	
162°	89.92	90.73	90.73	84.68	72.58	66.94	77.82	91.94	97.58	100.00	93.95	
180°	89.11	84.68	79.44	60.89	48.79	44.35	48.79	61.69	81.45	93.95	99.60	
平均识别率	86.60											

注:加粗字体(对角线)表示视角未改变情况下的识别率。

表4 穿大衣状态下 CASIA-B 上的 Rank-1 识别率

Table 4 Rank-1 accuracies of our method in clothing variation mode on CASIA-B dataset

模型图像集视角 (正常:01—04)	测试图像集视角(穿大衣:01—02)											/%
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
0°	90.93	78.02	69.96	61.69	47.38	35.08	35.89	46.17	55.04	66.13	68.55	
18°	73.59	97.38	91.53	81.05	65.93	54.44	52.02	60.89	69.76	68.15	57.86	
36°	62.30	91.13	96.57	92.14	81.45	71.17	67.74	77.42	73.79	72.38	54.84	
54°	57.46	83.67	94.96	93.95	89.11	81.65	80.65	82.06	74.80	63.51	45.97	
72°	45.56	69.96	87.10	93.15	96.77	91.73	91.13	90.12	80.44	60.28	40.52	
90°	42.74	63.51	78.23	86.69	91.33	93.15	90.12	86.69	77.22	51.21	39.72	
108°	45.36	64.52	79.84	85.08	88.71	90.73	97.18	94.56	82.66	67.54	38.71	
126°	63.51	71.37	80.85	81.25	80.65	85.48	89.72	94.15	89.52	80.04	49.40	
144°	65.93	71.37	79.64	74.40	72.58	74.60	80.65	87.90	92.74	86.29	56.65	
162°	71.77	77.62	77.02	69.15	58.67	49.19	57.46	73.79	85.28	93.15	73.19	
180°	65.52	62.10	53.43	47.58	34.68	28.43	33.47	45.36	56.45	71.37	90.73	
平均识别率	71.76											

注:加粗字体(对角线)表示视角未改变情况下的识别率。

可以看出,在背包状态下,本文方法的识别率均高于其他方法。在穿大衣状态下,本文方法在11个视角下的识别率均有不同程度提高,平均识别率为

94.25%,比GaitSet提高了9.39%,高于其他方法。同时,本文方法的识别率在前视角(0°—90°)情况下比后视角(108°—180°)情况下的更高,原因在于

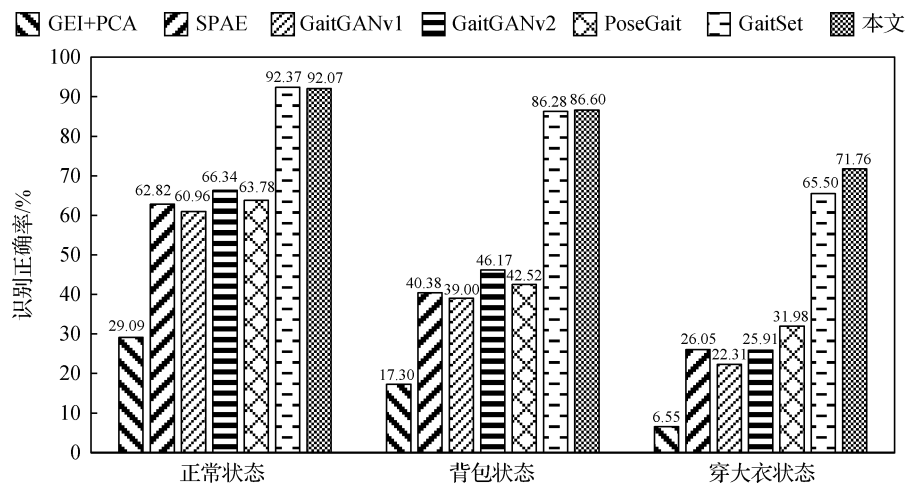


图 14 不同状态下不同方法的平均识别率

Fig. 14 The average accuracies of different methods when the probe data is in different modes

表 5 不同方法在背包状态下的 Rank-1 识别率

Table 5 Rank-1 accuracies of different methods in carrying variation mode

方法	测试图像集视角											/%
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
SPAE(Yu 等,2017a)	79.84	81.45	70.16	66.94	74.19	65.32	62.10	75.81	72.58	68.55	74.19	71.92
GaitGANv1(Yu 等,2017b)	79.03	76.61	75.81	68.55	65.32	64.52	69.35	73.39	73.39	76.61	77.42	72.73
GaitGANv2(Yu 等,2019)	84.68	82.26	74.19	79.64	72.58	70.97	65.32	78.23	74.19	79.03	77.42	76.23
PoseGait(Liao 等,2020)	74.19	75.81	77.42	76.61	69.35	70.16	70.97	69.35	74.19	65.32	60.48	71.26
PTSN(Liao 等,2017)	89.52	95.16	92.74	87.90	83.87	79.03	84.68	83.06	83.06	90.32	74.19	85.78
GaitSet(Chao 等,2019)	98.39	98.79	96.77	95.56	99.60	97.98	97.98	97.18	99.19	98.79	97.97	98.02
本文	99.60	99.19	99.19	97.98	100.00	98.39	99.60	98.39	99.60	100.00	99.60	99.23

注:加粗字体表示各列最优结果。

表 6 不同方法在穿大衣状态下的 Rank-1 识别率

Table 6 Rank-1 accuracies of different methods in clothing variation mode

方法	测试图像集视角											/%
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
SPAE(Yu 等,2017a)	44.35	49.19	46.77	46.77	49.19	42.74	46.77	43.55	40.32	41.13	42.74	44.87
GaitGANv1(Yu 等,2017b)	25.81	37.90	45.16	55.65	43.55	48.35	47.58	45.97	43.55	35.48	27.42	41.49
GaitGANv2(Yu 等,2019)	33.06	45.97	44.35	44.35	55.65	43.55	45.97	44.35	40.32	41.13	33.06	42.89
PoseGait(Liao 等,2020)	46.77	48.39	57.26	61.29	58.06	56.45	59.68	54.84	55.65	58.06	39.52	54.18
PTSN(Liao 等,2017)	53.23	83.87	87.90	72.58	61.29	61.29	75.00	66.94	70.97	70.16	45.97	68.11
GaitSet(Chao 等,2019)	82.06	85.89	86.90	83.67	90.12	87.70	88.71	86.49	86.09	80.44	75.40	84.86
本文	90.93	97.38	96.57	93.95	96.77	93.15	97.18	94.15	92.74	93.15	90.73	94.25

注:加粗字体表示各列最优结果。

前视角包含了更多肢体摆动信息。

综上所述,本文提出的姿态特征与2维傅里叶变换相结合的步态识别方法有助于提高行走状态改变条件下步态识别的性能。

4.4 跨视角条件下的识别性能评估

为验证本文方法在跨视角条件下不同状态(正常、背包和穿大衣)的性能,与目前在跨视角条件下取得最好性能的 GaitSet(Chao 等,2019)进行识别率比较,结果(不包括相同视角情况)如表7所示,其中 GaitSet 的结果取自原文。可以看出,在正常状态下,本文方法的平均识别率为 91.32%,比 GaitSet 低

0.73%,原因在于本文方法的融合策略中,步态特征的加权系数可能不是最优的,具体原因需要进一步研究。在背包和穿大衣状态下,本文方法的平均识别率为 85.34% 和 69.51%,分别比 GaitSet 提高了 1.08% 和 6.98%。表明本文方法进一步提高了背包和穿大衣状态下跨视角步态识别的性能。原因在于本文提出的将无肩姿态能量图、步态参数等姿态特征与步态参数的2维傅里叶变换相结合的步态识别算法融合了 GaitSet 网络特征的优点,有助于提高行走状态改变条件下的步态识别性能。

表7 不同状态下本文方法和 GaitSet 在 CASIA-B 上的 Rank-1 平均识别率

Table 7 Average rank-1 accuracies of our method and GaitSet on CASIA-B under different walking conditions

														/%
行走状态	方法	测试图像集视角												
		0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	平均	
正常	GaitSet	86.80	95.20	98.00	94.50	91.50	89.10	91.10	95.00	97.40	93.70	80.20	92.05	
	本文	86.21	93.63	96.85	95.16	90.00	88.15	89.76	94.60	95.56	94.35	80.24	91.32	↓0.73
背包	GaitSet	79.90	89.80	91.20	86.70	81.60	76.70	81.00	88.20	90.30	88.50	73.00	84.26	
	本文	77.14	89.23	92.30	88.91	84.60	80.04	83.39	89.52	92.26	87.90	73.43	85.34	↑1.08
穿大衣	GaitSet	52.00	66.00	72.80	69.30	63.10	61.20	63.50	66.50	67.50	60.00	45.90	62.53	
	本文	59.38	73.33	79.25	77.22	71.05	66.25	67.88	74.50	74.50	68.69	52.54	69.51	↑6.98

注:加粗字体表示性能较高的结果,“↑”、“↓”表示提高或下降的幅度,模型图像集视角为 0°—180°。

4.5 消融实验分析

为验证不同特征对识别性能的贡献,将无肩姿态能量图、有肩姿态能量图、步态参数及其2维傅里叶变换幅度谱和 GaitSet 进行组合实验。表8是穿大衣、背包和正常行走状态下,不同组合特征的平均识别率(121个视角组合(包括相同视角和跨视角)识别结果的平均值),其中1—9行是没有结合 GaitSet 特征、仅基于姿态信息的特征消融实验结果。“正常+背包”和“正常+穿大衣”状态下的边缘图如图15和图16所示。

从表8可以看出,如第1—3行所示,在3种状态下,步态参数及其2维傅里叶变换幅度谱的识别率比单一特征步态参数的识别率均有提高,说明本文提取的步态参数2维傅里叶变换幅度谱进一步提高了步态参数的可区分性。如第4—5行所示,在背包状态下,无肩姿态能量图的平均识别率为 31.44%,比有肩姿态能量图低 2.8%,原因在于与

正常状态相比,背包状态下的肩宽信息改变较小(图15(a)),但胳膊摆动易受携带包裹的影响,如有人背斜挎包时,一侧的胳膊摆动幅度较小或不摆动(图15(b)黄色框),而无肩姿态能量图忽略了肩宽信息,保留了胳膊摆动信息,因此造成识别率低。在穿大衣状态下,无肩姿态能量图单一特征的平均识别率为 28.87%,比有肩姿态能量图高 2.03%,原因在于穿大衣改变了肩宽信息(图16(a)(b)),有肩姿态能量图保持了肩宽信息,而无肩姿态能量图去掉了肩宽的影响,因此无肩姿态能量图的性能更好。如第6行所示,加权融合有肩姿态能量图和无肩姿态能量图的性能比单一姿态能量图在穿大衣、背包和正常状态下,分别提高了 2.21%、2.26% 和 3.74%。如第7—9行所示,将步态参数及其2维幅度谱与有肩姿态能量图和无肩姿态能量图加权融合有助于提高步态识别的性能(对比第6行),在穿大衣、背包和正常状态下分别提升了 4.48%、

表 8 特征消融实验结果
Table 8 Results of feature ablation experiments

序号	GaitSet	步态参数	2 维傅里叶 变换幅度谱	步态参数及其 2 维 傅里叶变换幅度谱	有肩姿态 能量图	无肩姿态 能量图	平均识别率/%		
							穿大衣状态	背包状态	正常状态
1		✓					13.42	14.36	22.27
2			✓				11.78	12.12	19.56
3				✓			16.50	16.58	25.63
4					✓		26.84	34.24	42.96
5						✓	28.87	31.44	42.08
6					✓	✓	31.08	36.50	46.70
7				✓	✓		35.24	40.75	50.29
8				✓		✓	35.55	38.88	50.55
9				✓	✓	✓	35.56	40.54	53.87
10	✓						65.50	86.28	92.37
11	✓			✓			66.51 ↑1.01	86.68 ↑0.39	92.02 ↓0.35
12	✓				✓		68.55 ↑3.05	87.36 ↑1.07	92.38 ↑0.01
13	✓					✓	68.60 ↑3.10	87.18 ↑0.89	92.94 ↑0.57
14	✓				✓	✓	68.87 ↑3.37	85.21 ↓1.08	91.33 ↓1.04
15	✓			✓	✓		70.05 ↑4.55	87.81 ↑1.52	92.90 ↑0.53
16	✓			✓		✓	70.84 ↑5.34	86.07 ↓0.22	91.28 ↓1.09
17	✓			✓	✓	✓	71.76 ↑6.26	86.60 ↑0.32	92.07 ↓0.30

注:加粗字体表示相对 GaitSet 特征性能有所提升,“↑”、“↓”表示识别率相对只使用 GaitSet 提高或下降的幅度,“✓”表示使用当前特征加权融合。

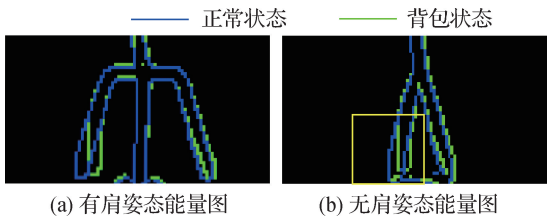


图 15 正常和背包状态下的边缘图
Fig. 15 Edge maps under walking normally and with a bag
((a) PEMS; (b) PEMoS)

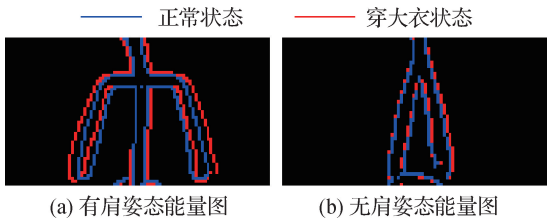


图 16 正常和穿大衣状态下的边缘图
Fig. 16 Edge maps under walking normally and with a different coat((a) PEMS; (b) PEMoS)

4.04% 和 7.17%。如第 10—17 行所示,将步态参数及其 2 维傅里叶变换幅度谱、无肩姿态能量图与 GaitSet 结合,有效利用了姿态特征和人体外观形状特征的优点,提升了步态识别的性能,在穿大衣和背包状态下比 GaitSet 分别提高了 6.26% 和 0.32%,原因在于本文方法在保持 GaitSet 优点的同时,融合了减弱衣服改变影响的无肩姿态能量图以及增强下肢结构参数和运动协调性的步态参数及其 2 维傅里叶变换幅度谱。在正常状态下,本文方法融合了 GaitSet 的特征优点,识别率也应该像其他状态下一样高于 GaitSet,但实际上却在 GaitSet 的 92.37% 识别率上下浮动,原因在于融合策略中步态特征的加权系数可能不是最优,使得跨视角情况下,本文方法的平均识别率低 0.73%。

4.6 采样点个数 T 对算法性能的影响

为验证采样点个数 T 对步态识别性能的影响,

统计了 T 取不同值时不同状态下的步态识别率, 其中 T 从 $\{16, 21, 26, 31, 36, 41, 46, 51, 56, 61, 66, 71\}$ 中取值, T 对与其关系紧密的步态参数及其2维傅里叶变换幅度谱特征性能和对本文算法总体性能的影响如图17所示。为便于细节比较, 表9列出了 T 在41~51之间取值时本文算法的总体性能。

从图17和表9可以看出, 步态识别率随着 T 的递增呈先增大、稳定(变化较小)、后减小的趋势, 原因在于: 1) 当 T 增大时, 步态参数序列包含的信息增加, 算法的识别率也随之提升。2) 当 T 增大到一定值时, 识别率呈稳定趋势, 不再有较大变化, 如在穿大衣状态下, 步态识别的总体性能在70.93%~71.67%之间, 波动幅度较小。原因在于步态参数具

有周期性, 当 T 超过一定值时, 就不会有新增信息, 所以识别率也不会有较大变化。3) 在 $T = 46$ 时, 识别率出现峰值, 原因在于正常行走一步需要0.48~0.63 s, 一个步态周期包括连续3步, 需要的时间是1.4~1.9 s, 即 T 取值在1.4~1.9倍帧频之间的整数, 即能覆盖一个步态周期的步态参数。本文实验基于CASIA-B数据集进行, 其帧频是25帧/s, 算法性能在 T 取1.84倍帧频, 即 $T = 46$ 帧步态识别率最优。4) 当 $T > 51$ 时(参考图17), 步态识别的识别率明显下降, 原因在于满足此条件的模型图像集和测试图像集中的视频帧数变少, 在步态参数对齐时, 需外插的点数增加, 同时增加了步态参数的误差, 从而造成识别率下降。

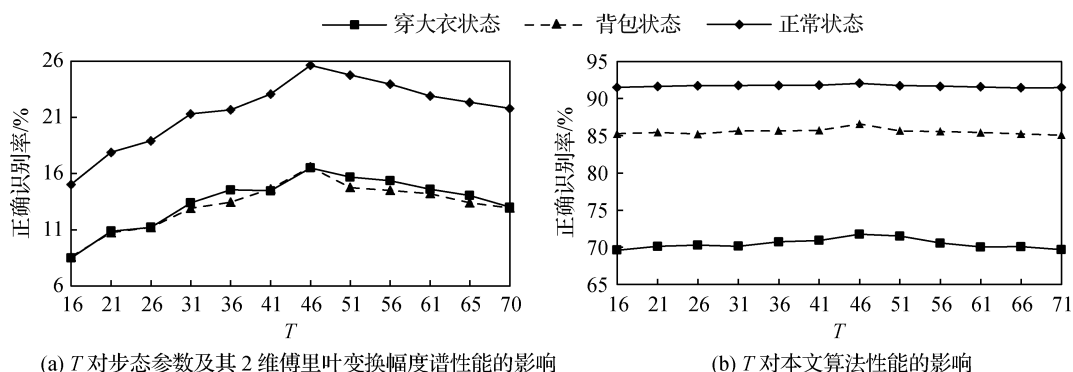


图17 步态参数采样点个数 T 对算法性能的影响

Fig. 17 The influence of the number of gait parameter sampling points T on the algorithm performance

((a) the influence of the number of gait parameter sampling points T on the performance of 2D Fourier transform magnitude spectrum of gait parameters; (b) the influence of the number of gait parameter sampling points T on the performance of ours)

表9 本文算法在 T 值为41~51时的性能

Table 9 The performance of the proposed algorithm when the value of T varies from 41 to 51

行走状态	T											%
	41	42	43	44	45	46	47	48	49	50	51	
穿大衣	70.93	71.43	71.45	71.57	71.67	71.76	71.39	71.65	71.60	71.53	71.52	
背包	85.73	85.89	85.95	85.92	86.21	86.60	85.82	85.72	85.71	85.67	85.67	
正常	91.82	91.85	91.87	91.85	91.87	92.07	91.95	91.87	91.81	91.76	91.76	

注: 加粗字体表示各行最优结果。

4.7 算法运行时间分析

本文算法的输入是待检视频以及已离线提取完特征的模型图像集(包含62人, 每人11个视角, 每个视角4个样本, 共 $62 \times 11 \times 4 = 2728$ 个视频), 输出为待检序列的预测标签。算法需要调用OpenPose提取每帧图像的关节点和GaitSet程序提

取特征, 这两个程序代码来源于原网站, 需要GPU支持; 其余程序采用本文自编的MATLAB程序, 不需要GPU支持。运行平台配置为主频Intel(R) Xeon(R) E5-2678 v3 @ 2.50 GHz CPU(2处理器)、32 GB内存、2块NVIDIA TITAN XP GPU、Windows 10操作系统、MATLAB R2016b和Python 3.6。

1)本文算法的运行时间及分析。本文算法分为特征提取、特征映射和匹配与识别3个阶段,各个阶段分别统计耗时,具体见表10。

在特征提取阶段,统计每帧图像提取特征的平均耗时(表10第1—6行)。姿态估计的帧数 T_p 需大于 T 才能提取 $T=46$ 帧时的数据,表征待检测序列,实验中 $T_p=100$;另外,GaitSet需要30帧

步态侧影图组成集合,因此本阶段的总耗时计算为

$$C_a = C_p \times T_p + G_p \times 30 + C_f \times T \tag{12}$$

式中, C_a 表示本阶段总耗时, T_p 表示姿态估计帧数, G_p 表示GaitSet方法每帧耗时, C_f 表示其余特征提取每帧耗时。最终本文方法总耗时约为 $57.1 \times 100 + 1.2 \times 30 + (7.9 + 2.4 + 2.6) \times 46 = 6\,339.4\text{ ms}$ 。

表 10 本文方法各步骤的平均耗时
Table 10 Average time consumption of each step of the proposed algorithm

序号	阶段	步骤	平均耗时	描述
1	特征提取	姿态估计	57.1 ms/帧	作者提供程序,GPU 和 OpenPose
2	特征提取	GaitSet	1.2 ms/帧	作者提供程序,GPU 和 Python
3	特征提取	步态参数及其2维傅里叶变换幅度谱特征提取	7.9 ms/帧	自编程序,CPU 和 MATLAB
4	特征提取	无肩姿态能量图特征提取	2.4 ms/帧	自编程序,CPU 和 MATLAB
5	特征提取	有肩姿态能量图特征提取	2.6 ms/帧	自编程序,CPU 和 MATLAB
6	MvDA-VC	步态参数及其2维傅里叶变换幅度谱特征映射	1.4 ms	自编程序,CPU 和 MATLAB
7	特征映射	无肩姿态能量图特征映射	1.5 ms	自编程序,CPU 和 MATLAB
8	特征映射	有肩姿态能量图特征映射	1.8 ms	自编程序,CPU 和 MATLAB
9	匹配与识别	特征相似度计算、分数融合与分类	783.5 ms	自编程序,CPU 和 MATLAB,共2 728个视频序列
10	总计	$57.1 \times 100 + 1.2 \times 30 + (7.9 + 2.4 + 2.6) \times 46 + 1.4 + 1.5 + 1.8 + 783.5 = 7\,127.6\text{ ms}$		

在匹配与识别阶段,待检序列分别与模型图像集中的2 728个视频数据进行匹配,取排名第1的步态序列对应标签为待识别步态序列的类别,在此阶段,本文算法耗时约为783.5 ms。

综上所述,本文算法从输入一个待检原始视频到预测出其标签总耗时约为7 127.6 ms,其中姿态估计是主要耗时环节,约占总耗时的90%。需要说明的是,本文自编程序采用MATLAB编码,并且没有经过优化,若采用C++编写并进行优化,时间会快得多。

2)与典型算法的运行时间比较分析。本文步态表达包括GaitSet网络特征、有肩姿态能量图、无肩姿态能量图、步态参数及其2维幅度谱特征,因此特征表达的速度要比GaitSet慢。

GEI+PCA、SPAЕ、GaitGANv1和GaitGANv2方法均是直接基于步态侧影图的方法,而步态侧影图特征提取耗时较少,另外GaitGANv1在2 NVIDIA TITAN XP机器上特征提取的平均耗时约为3.6 ms(该方法的输入是步态能量图,平均耗时未包含步

态特征预处理部分),而本文算法需预先通过姿态估计方法检测关节点序列,单此一项每帧图像平均耗时约为57.1 ms。因此在速度方面,本文方法要比GEI+PCA、SPAЕ、GaitGANv1和GaitGANv2慢。

本文算法和PoseGait均是基于姿态估计方法OpenPose(Cao等,2019)检测关节点,因此该环节耗时基本相同;PoseGait中2D姿态关节点转化为3D姿态关节点,采用网上公开源码(Chen和Ramanan,2017)在本文算法相同环境下运行,一帧图像2D关节点转化为3D关节点的平均耗时约417 ms(CPU),而本文方法提取1帧图像特征的耗时约为1.2 ms(CPU)+12.9 ms(CPU),特征匹配与识别部分耗时基本相同,因此本文方法速度要快于或接近PoseGait的速度。

5 结 论

为了减弱行走状态改变对步态识别性能的影响,提出了姿态特征与2维傅里叶变换相结合的步

态识别方法。该方法的核心思想是:1)提出了忽略肩宽信息的无肩姿态能量图,在兼顾上肢、下肢对步态表示作用的同时,减弱了衣服改变对步态识别的影响;2)提出了基于3个及以上下肢关节的局部结构参数,用于增强下肢多个关节之间的协调性行为在步态识别中的作用;3)提出用步态参数的2维傅里叶变换的幅度谱增强行走过程中单肢体运动的周期性和肢体之间运动的协调性在步态表示中的作用;4)综合利用各特征的优点及其适用场景,提出了加权平均与最大池化相结合的两层分数融合策略进行步态识别。实验结果表明:本文方法有效提高了视角、携带物品和衣服改变条件下步态识别的准确性。

本文存在的问题是在正常状态下,跨视角识别性能要比 GaitSet 方法的性能稍低,原因在于融合策略中各步态特征加权系数是基于实验结果选择的固定值。在下一步的工作中,将开展加权系数自适应学习策略的研究。

参考文献 (References)

- Cao Z, Hidalgo G, Simon T, Wei S E and Sheikh Y. 2019. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(1): 172-186 [DOI: 10.1109/TPAMI.2019.2929257]
- Chao H Q, He Y W, Zhang J P and Feng J F. 2019. GaitSet: regarding gait as a set for cross-view gait recognition//*Proceedings of 2019 AAAI Conference on Artificial Intelligence*, 33(1): 8126-8133 [DOI: 10.1609/aaai.v33i01.33018126]
- 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*. Honolulu, USA: IEEE: 5759-5767 [DOI: 10.1109/CVPR.2017.610]
- Cunado D, Nixon M S and Carter J N. 1997. Using gait as a biometric, via phase-weighted magnitude spectra//*Proceedings of the 1st International Conference on Audio- and Video-Based Biometric Person Authentication*. Crans-Montana, Switzerland: Springer: 93-102 [DOI: 10.1007/BFb0015984]
- Feng Y, Li Y C and Luo J B. 2016. Learning effective gait features using LSTM//*Proceedings of the 23rd International Conference on Pattern Recognition*. Cancun, Mexico: IEEE: 325-330 [DOI: 10.1109/ICPR.2016.7899654]
- 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]
- Huang X X and Boulgouris N V. 2012. Gait recognition with shifted energy image and structural feature extraction. *IEEE Transactions on Image Processing*, 21(4): 2256-2268 [DOI: 10.1109/TIP.2011.2180914]
- Kale A, Sundaresan A, Rajagopalan A N, Cuntoor N P, Roy-Chowdhury A K, Krüger V and Chellappa R. 2004. Identification of humans using gait. *IEEE Transactions on Image Processing*, 13(9): 1163-1173 [DOI: 10.1109/TIP.2004.832865]
- Kan M N, Shan S G, Zhang H H, Lao S H and Chen X L. 2016. Multi-view discriminant analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(1): 188-194 [DOI: 10.1109/TPAMI.2015.2435740]
- Kang Z R, Deng M Q and Wang C. 2017. Frontal-view human gait recognition based on Kinect features and deterministic learning//*Proceedings of the 36th Chinese Control Conference*. Dalian, China: IEEE: 10834-10839 [DOI: 10.23919/ChiCC.2017.8029085]
- Lam T H W, Cheung K H and Liu J N K. 2011. Gait flow image: a silhouette-based gait representation for human identification. *Pattern Recognition*, 44(4): 973-987 [DOI: 10.1016/j.patcog.2010.10.011]
- Liao R J, Cao C S, Garcia E B, Yu S Q and Huang Y Z. 2017. Pose-based temporal-spatial network (PTSN) for gait recognition with carrying and clothing variations//*Proceedings of the 12th Chinese Conference on Biometric Recognition*. Beijing, China: Springer: 474-483 [DOI: 10.1007/978-3-319-69923-3_51]
- 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]
- Yang K, Dou Y, Lyu S H, Zhang F and Lyu Q. 2016. Relative distance features for gait recognition with Kinect. *Journal of Visual Communication and Image Representation*, 39: 209-217 [DOI: 10.1016/j.jvcir.2016.05.020]
- Yao L X, Kusakunniran W, Wu Q, Zhang J and Tang Z M. 2018. Robust CNN-based gait verification and identification using skeleton gait energy image//*Proceedings of 2018 Digital Image Computing: Techniques and Applications*. Canberra, Australia: IEEE: 1-7 [DOI: 10.1109/DICTA.2018.8615802]
- 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, Shen L L and Huang Y Z. 2016. View invariant gait recognition using only one uniform model//*Proceedings of the 23rd International Conference on Pattern Recognition*. Cancun, Mexico: IEEE: 889-894 [DOI: 10.1109/ICPR.2016.7899748]
- Yu S Q, Chen H F, Wang Q, Shen L L and Huang Y Z. 2017a. Invariant feature extraction for gait recognition using only one uniform

- model. *Neurocomputing*, 239: 81-93 [DOI: 10.1016/j.neucom.2017.02.006]
- Yu S Q, Chen H F, Reyes E B G and Poh N. 2017b. GaitGAN: invariant gait feature extraction using generative adversarial networks// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu, USA: IEEE: 532-539 [DOI: 10.1109/CVPRW.2017.80]
- Yu S Q, Liao R J, An W Z, Chen H F, García E B, Huang Y Z and Poh N. 2019. GaitGANv2: invariant gait feature extraction using generative adversarial networks. *Pattern Recognition*, 87: 179-189 [DOI: 10.1016/j.patcog.2018.10.019]
- Zhang R, Vogler C and Metaxas D. 2007. Human gait recognition at sagittal plane. *Image and Vision Computing*, 25(3): 321-330 [DOI: 10.1016/j.imavis.2005.10.007]
- Zhen H, Deng M Q, Lin P and Wang C. 2018. Human gait recognition based on deterministic learning and Kinect sensor//*Proceedings of 2018 Chinese Control and Decision Conference*. Shenyang, China: IEEE: 1842-1847 [DOI: 10.1109/CCDC.2018.8407426]

作者简介



王新年, 1973 年生, 男, 副教授, 博士生导师, 主要研究方向为数字图像处理、生物特征识别和模式识别。

E-mail: wxn@dlmu.edu.cn

胡丹丹, 女, 硕士研究生, 主要研究方向为图像处理与计算机视觉。E-mail: 17824855102@163.com

张涛, 女, 高级工程师, 主要研究方向为数字图像处理、生物特征识别和模式识别。E-mail: lnnuzt@lnnu.edu.cn

白桂欣, 男, 高级工程师, 主要研究方向为数字信号处理。

E-mail: bgx@dlmu.edu.cn