



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象学报

2014

09

VOL.19

ISSN1006-8961
CN11-3758/TB





第19卷第9期（总第221期）

2014年9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

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 中国科学院遥感与数字地球研究所

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 李小文

编辑出版《中国图象图形学报》编辑出版委员会

邮政信箱 北京9718信箱

邮 编 100101

电子信箱 jig@irsa.ac.cn

电 话 010-64807995

网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361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 Institute of Remote Sensing and Digital Earth, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen

Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics

P.O.Box 9718, Beijing, P.R.China

Zip code 100101

E-mail jig@irsa.ac.cn

Telephone 010-64807995

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

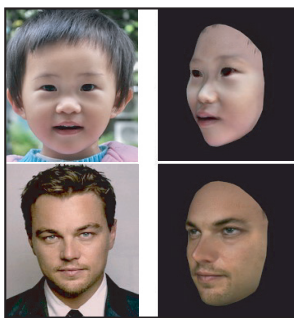
国内定价 50.00元



自适应对称自回归模型的压缩图像内插方法(第1268页)



加权局部特征结合判别式字典的目标跟踪算法(第1316页)



结合特征适配与拉普拉斯形变的3维人脸重建(第1349页)

综述

能量泛函正则化模型在图像恢复中的应用分析

李旭超, 刘海宽, 宋博 1247

图像处理和编码

融合最小化失真与模型保持的隐写算法

高瞻瞻, 汤光明, 张伟伟 1260

自适应对称自回归模型的压缩图像内插方法

刘婧, 干宗良, 崔子冠, 陈昌红, 朱秀昌 1268

基于能量泛函和视觉特性的全变分图像降噪模型

郭从洲, 秦志远, 时文俊 1282

分块策略实现图像椒盐噪声密度估计

林亚明, 李佐勇, 林叶郁 1288

图像分析和识别

Dempster-Shafer证据融合金字塔韦伯局部特征的表情识别

王晓华, 金超, 任福继, 胡敏 1297

运动捕捉数据中足迹的谱聚类检测方法

刘晓平, 陆劲挺, 谢文军 1306

加权局部特征结合判别式字典的目标跟踪

王飞, 房胜 1316

面向瓦当文字识别的改进水平集骨架提取

邓茜芸, 周明全, 武仲科, 王醒策 1324

改进复杂网络模型的形状特征提取

阮瑞, 江波, 汤进, 罗斌 1332

几何特性二元关系的直线匹配

胡海霞, 李钢 1338

图像理解和计算机视觉

结合特征适配与拉普拉斯形变的3维人脸重建

张剑, 何骅, 詹小四, 肖俊 1349

利用背景加权和选择性子模型更新的视觉跟踪算法

黄安奇, 侯志强, 余旺盛, 刘翔 1360

计算机图形学

对可控Bézier曲线的改进

严兰兰, 韩旭里 1368

医学图像处理

结合区域生长与灰度重建的CT图像肺气管树分割

彭双, 肖昌炎 1377

检测肺结节的3维自适应模板匹配

高婷, 龚敬, 王远军, 聂生东, 孙希文 1384

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Novel compression image interpolation method using adaptive symmetrical autoregressive models(P1268)



Visual tracking based on discriminative dictionary and weighted local features (P1316)



Three dimensional face reconstruction via feature adaptation and Laplace deformation (P1349)

Review

- Application analysis of regularization model of energy functional to image restoration
Li Xuchao, Liu Haikuan, Song Bo 1247

Image Processing and Coding

- Steganographic method combining minimum embedding distortion and model preservation
Gao Zhanzhan, Tang Guangming, Zhang Weiwei 1260
- Novel compression image interpolation method using adaptive symmetrical autoregressive models
Liu Jing, Gan Zongliang, Cui Ziguan, Chen Changhong, Zhu Xiuchang 1268
- TV image denoising model based on energy functionals and HVS
Guo Congzhou, Qin Zhiyuan, Shi Wenjun 1282
- Image salt-and-pepper noise estimation based on partitioning strategy
Lin Yaming, Li Zuoyong, Lin Yeyu 1288

Image Analysis and Recognition

- Research on facial expression recognition based on pyramid Weber local descriptor and the Dempster-Shafer theory of evidence
Wang Xiaohua, Jin Chao, Ren Fuji, Hu Min 1297
- Foot plant detection based on spectral clustering algorithm for motion capture data
Liu Xiaoping, Lu Jintong, Xie Wenjun 1306
- Visual tracking based on the discriminative dictionary and weighted local features
Wang Fei, Fang Sheng 1316
- Skeleton extraction using improved level set methods in eaves' text recognition
Deng Qianyun, Zhou Mingquan, Wu Zhongke, Wang Xingce 1324
- Improved shape feature extraction using complex network model
Ruan Rui, Jiang Bo, Tang Jin, Luo Bin 1332
- Line matching based on binary relations of geometric attributes
Hu Haixia, Li Gang 1338

Image Understanding and Computer Vision

- Three dimensional face reconstruction via feature adaptation and Laplace deformation
Zhang Jian, He Hua, Zhan Xiaosi, Xiao Jun 1349
- Visual object tracking method based on weighted background and selective sub-model update strategy
Huang Anqi, Hou Zhiqiang, Yu Wangsheng, Liu Xiang 1360

Computer Graphics

- Improvement of the modifiable Bézier curves
Yan Lanlan, Han Xuli 1368

Medical Image Processing

- Segmentation of pulmonary airway tree in CT images by combining region growing and grayscale reconstruction algorithms
Peng Shuang, Xiao Changyan 1377
- Three dimensional adaptive template matching algorithm for lung nodule detection
Gao Ting, Gong Jing, Wang Yuanjun, Nie Shengdong, Sun Xiwen 1384

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2014)09-1306-10

论文引用格式: Liu X P, Lu J T, Xie W J. Foot plant detection based on spectral clustering algorithm for motion capture data[J]. Journal of Image and Graphics, 2014, 19(9): 1306-1315. [刘晓平, 陆劲挺, 谢文军. 运动捕捉数据中足迹的谱聚类检测方法[J]. 中国图象图形学报, 2014, 19(9): 1306-1315.][DOI:10.11834/jig.20140907]

运动捕捉数据中足迹的谱聚类检测方法

刘晓平, 陆劲挺, 谢文军

合肥工业大学计算机与信息学院, 合肥 230009

摘要: **目的** 从运动捕捉数据中直接检测足迹非常困难。目前,已有的方法不能在无人工交互条件下实现对任意平地运动数据的足迹自动检测,为此提出一种基于谱聚类的足迹自动检测方法,可应用于任意角色的平地运动。**方法** 首先,提取角色的脚部运动特征并表示为样本。然后,分析样本模长的变化规律并自适应计算谱聚类参数。最后,使用谱聚类方法检测出足迹帧。**结果** 实验应用于混合运动数据集后,足迹检测的准确性良好,检测的整体准确率可达98.72%。**结论** 对实验结果的分析以及与基准线法的比较,证明了本文方法的普适性和有效性。**关键词:** 计算机动画;运动捕捉;运动编辑;足迹检测;谱聚类

Foot plant detection based on spectral clustering algorithm for motion capture data

Liu Xiaoping, Lu Jinting, Xie Wenjun

School of Computer and Information, Hefei University of Technology, Hefei 230009, China

Abstract: **Objective** It's hardly to directly detect the foot plants from motion capture data. Many previous works have already successfully found the foot plant constraints. However, none of these methods is completely automatic without any interaction. In this paper, we present a foot plant detection method based on spectral clustering for motion capture data. **Method** First, samples are represented by motion features of performer's feet. Second, parameters are selected by analyzing the norms of samples. Finally, foot plants are detected by spectral clustering algorithm. **Result** After applying on a blending motion data set, high accuracy rates of foot plant detection are acquired, the accuracy rates of foot plant detection can reach 98.72%. **Conclusion** The analysis of experimental results and the comparison with baseline methods demonstrate the generality and efficiency of our method.

Key words: computer animation; motion capture; motion editing; foot plant detection; spectral clustering

0 引言

足迹是生物运动的一种典型约束,表现为脚部接触地面并保持固定的一段时间^[1]。运动捕捉设备已经越来越广泛应用于制作计算机动画,采集的

运动捕捉数据中也包含了采样角色的足迹信息。然而,由于几乎所有的运动编辑方法都会破坏原始运动捕捉数据中的足迹^[2],因此通常在进行运动编辑前需要先记录原始数据的足迹,用作编辑后足迹修正的依据。

然而,想要直接通过判断脚部是否静止来提取

收稿日期:2014-03-14;修回日期:2014-05-09

基金项目:国家自然科学基金项目(61370167);高等学校博士学科点专项科研基金项目(20120111110003)

第一作者简介:刘晓平(1964—),男,教授,博士生导师,1998年于合肥工业大学获计算机应用技术专业工学博士学位,主要研究方向为计算机图形学、计算机辅助设计。E-mail: lilin_julia@163.com

出运动捕捉数据中的足迹几乎没有可能。这是由于运动捕捉数据中普遍存在一类典型的足迹噪声,常被描述为当脚部应固定于地面时却发生足迹丢失而呈现出脚部空间滑动的现象^[2],图1展示了3种常见的足迹噪声实例,图1中的曲线为一个脚尖在运动中的高度轨迹。由于足迹噪声的存在,无论在多么理想的运动捕捉数据中,足迹也不可能完全静止。因此,快速准确的足迹检测是进行运动编辑不可避免的前端工作。

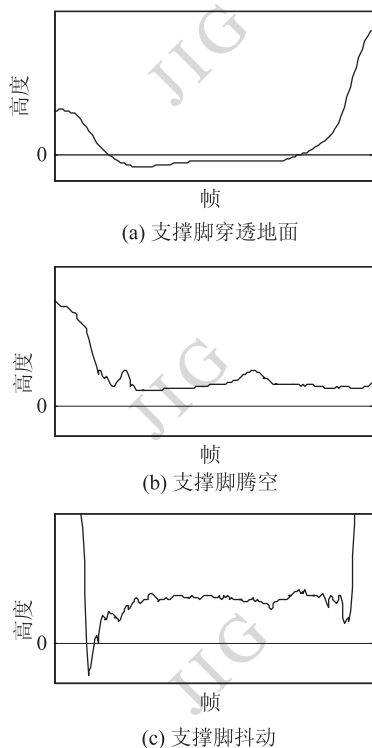


图1 运动捕捉数据中足迹噪声的实例

Fig. 1 Noises of foot plants in motion capture data

现有运动捕捉数据的足迹检测方法主要包括人工检测、基准线(baseline)法及分类器法等。

人工检测是准确率很高的传统方法,但检测代价大,效率不高。Kovar 等人^[1]通过人工指定存在足迹约束时刻的脚部空间位置以清除足迹噪声。Popović 等人^[3]根据用户指定的脚部足迹位置实现两足动物的地形自适应运动。

基准线方法一般通过对角色的空间位置、速度等信息人工指定阈值进行足迹划分。Ménardais 等人^[4]规定高度和速度同时小于各自的指定全局阈值的时段为足迹。然而这类方法容易导致足迹边界判定不准确,使得提取的足迹过短或过长,同时过大的噪声还可能引起足迹的断裂。Glardon 等人^[5]将

指定全局阈值改进为结合运动模式^[6]的自适应分段阈值,并在线检测足迹。此方法需要依赖预设的运动模式,对于除规律性行走、跑步等常规运动之外的无规律运动,由于难以总结出相应的运动模式,故处理能力有限。另外,不同条件下采集的数据所体现的噪声特征不一致,所以,使用基准线方法仍需每次重新设定阈值,难以实现自动化。

分类器方法则以 Ikemoto 等人^[2]提出的基于 K 近邻法的足部状态分类器为代表,可以实现无交互、高精度的足迹检测。然而该方法除了前期需要通过大量手工标定获得训练样本之外,其最大的局限性在于对采样角色的高敏感度。一旦采样对象发生变化,分类器便会失效,只能通过使用新对象的运动数据重新训练分类器得以解决。因此不能快速地适用于多样化的运动捕捉数据集合。

目前,仍没有一种真正意义上的自动化方法能够实现无人工干预的足迹检测。因此,本文以任意的平地运动数据为对象,提出一种可广泛应用的基于谱聚类的足迹检测方法。首先,逐帧提取脚部的运动特征作为样本,并对样本进行窗口化处理。然后,针对谱聚类方法中的高斯核函数,提出了一种根据不同运动数据自适应计算其尺度参数的方法。最后,根据计算出的尺度参数,使用谱聚类方法将样本划分为足迹帧和非足迹帧两类,其中连续的足迹帧区间即为运动数据的足迹。

1 谱聚类方法

谱聚类方法是基于谱图理论的一类聚类方法,目前广泛应用于模式识别和图像分割领域^[7]。它将待聚类的样本看做无向图的顶点,样本之间的相似度看做顶点之间边的权重,从而将样本的聚类问题转化为无向图的划分问题。对于 n 类别的聚类问题,其目标即为寻找将样本构成的无向图划分为 n 个子图的最小割问题。

谱聚类的优点在于能够处理复杂或未知形状的样本空间上的聚类^[8],而传统经典聚类算法,如 K-means 等,大多只能处理凸样本空间问题。同时谱聚类能够保证全局收敛^[9],并且对于高维数据,由于存在降维的过程,通常比直接运行 K-means 计算复杂度低。本文采用的是经典的谱聚类算法之一的 Ng-Jordan-Weiss (NJW) 谱聚类算法^[10],该算法中的样本相似度函数通常采用高斯核函数。

2 本文方法

运动捕捉设备所用的人体骨架模型通常是由关节连接骨骼构成的层次结构模型,其中根关节一般位于人体髋骨中心,具有6个自由度,其他关节各有3个自由度,如图2所示。实际中不同设备采用的人体标准骨架模型存在轻微差异。分析、编辑运动通常需要先选择合适的关节点作为末端效应器(end-effector)^[11],末端效应器的选择一般需根据具体的骨架模型和运动特征而定。脚部末端效应器一般选择脚趾或脚跟关节,本文选择脚部仅连接一根骨骼的脚尖关节为脚部末端效应器。

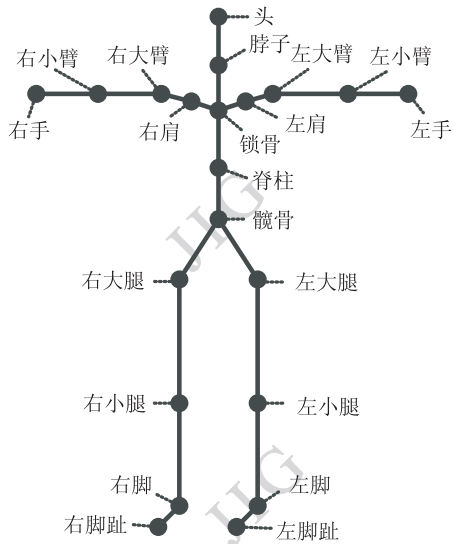


图2 人体骨架层次结构示意图

Fig. 2 Human skeleton structure

对于一个给定的运动序列,足迹检测问题可描述为判断脚部在每一帧的状态是否为接触地面且静止。若一只脚在某帧被认定为接触地面且静止,则称该帧为该脚的足迹帧。一段连续的足迹帧则构成一个足迹。因此,足迹检测可以描述为脚部末端效应器在每一帧是否为足迹帧的两类别判定问题。

由于运动捕捉技术的数据来源广泛、采样对象众多、采样的运动类型多变,因此难以找到普适的先验知识支撑如Ikemoto等人^[2]提出的分类器方法。在没有较好的先验知识的前提下,足迹检测可以看做是一个无需先验知识的两类别的聚类问题。选择适当的特征向量作为样本以描述脚部末端效应器在每一帧的状态,再利用谱聚类方法划分出该末端效应器在当前运动序列中所有的足迹帧和非足迹帧。

由于不能确定样本空间是凸空间,因此谱聚类方法比其他聚类方法更合适。

定义一个两位的标签 $lbl_i = [L_i, R_i]^T$ 记录双脚在运动序列中第 i 帧的状态, L_i 和 R_i 分别用于记录左、右脚在第 i 帧是否为足迹帧。则当前运动序列的足迹标签即可描述为 $LBL = [lbl_1, \dots, lbl_i, \dots, lbl_F]$,其中 $i = 1, 2, \dots, F$, F 是运动序列的总帧数。

2.1 聚类特征选择

聚类特征的选择会直接影响聚类样本相似度测量的有效性。合适的聚类特征构造的样本相似度能够使相同类别内的样本具有高相似度,而不同类别间的样本具有低相似度^[12]。

若令地面高度为0,“脚部接触地面并保持静止”这一约束等价于脚部的速度和高度同时为0。然而由于噪声的存在,实际上运动数据中的脚部在足迹状态时多表现为高度与速度接近0的状态,而在非足迹状态时多表现为高度或速度至少一项远大于0。传统的人工标定和基准线方法也主要以高度和速度为判定足迹状态的标准。可见,高度和速度是判断脚部状态的直观且有力的依据,因此,选择末端效应器的高度和速度为聚类特征子集。

从运动数据中可直接获取的信息只有根节点的位移和各关节的旋转量,之后通过正向运动学(FK)计算出脚部在每帧的空间位置,进而由离散的空间位置通过曲线拟合近似计算出脚部在每帧的速度。

对于运动序列 M ,定义其一个脚部末端效应器 E 的聚类样本集合 $X = \{x_i\} = \{[h_i, v_i]^T\}$, $i = 1, 2, \dots, F$,其中 h_i, v_i 表示 E 在第 i 帧的高度和速度。

由于足迹帧的高度和速度具有均接近0这一显著特征,可有效区别非足迹帧,故可以对聚类结果的类别标签设定相应规则:聚类后均值较小的一类样本的标签为足迹帧,反之另一类为非足迹帧。

实验中发现,标签在噪声较大处易产生短时抖动,如图3所示。运动捕捉数据的采样频率通常在20~120帧/s范围内,并且人体的反应时间通常介于0.1~0.4s。基于这两个事实,显然,标签相同的连续帧区间不应帧数过少,否则不足以支撑人体的正常反应,违反了人体的生理极限,属于标签错误标记。为了减少这种情况,借鉴数字滤波算法中邻域窗口的概念,对聚类样本进行窗口化平滑处理。使用部分相邻帧为当前帧定义一个窗口化特征向量,即令样本 $x_i = [x_{i-w}^T, x_{i-w+1}^T, \dots, x_i^T, \dots, x_{i+w}^T]^T$,其中 w 是预设

的窗口尺度参数,单位为帧,平滑窗口的宽度为 $2w+1$ 帧。经过实验验证,在其他参数取值合适的情况下, $w=1$ 或 $w=2$ 是比较合适的取值。

需要强调的是,窗口化处理仅能够在一定程度上降低噪声干扰,尤其对于消除足迹边界的跳帧现象效果较好。但对于尺度较大或者运动捕捉系统未良好校准所遗留的严重噪声仍然略显乏力。

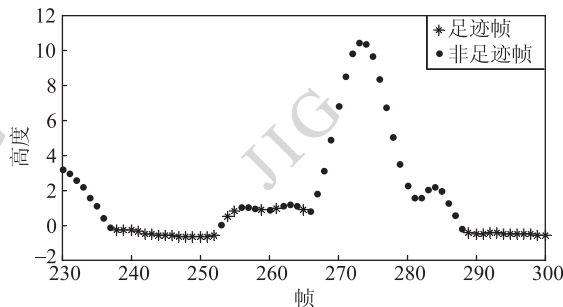


图3 未窗口化样本聚类结果中的足迹标签抖动实例

Fig. 3 The fluctuation of foot plant labels in spectral clustering results without neighborhood window

2.2 高斯核函数尺度参数自适应选择

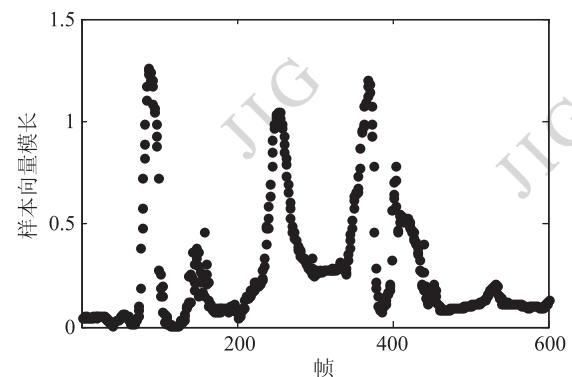
谱聚类算法通常需要涉及的一个参数是用来进行样本相似度测量的核函数中的尺度参数。选用高斯核函数,其形式化表达为 $A_{ij} = \exp(-\|x_i - x_j\|^2 / \sigma^2)$, σ 为其尺度参数。尺度参数 σ 的本质是一个度量标准,当两个样本之间的距离小于尺度参数时则被认为是相似的,反之被认为是不相似的。由于谱聚类方法对尺度参数非常敏感,不同的值会得到差异很大的聚类结果^[9],因此尺度参数的选择一直被认为是领域内的一个开放议题^[8]。目前尺度参数的确定除了一些自适应算法^[8]外,主要仍依靠人工经验尝试。

如果对每个不同的运动数据均需人工确定 σ 的值,则无法实现真正意义上的自动化。然而,使用文献[7]的尺度参数自适应算法,其效果也并不理想,这是由于聚类的目标并非是同类样本之间距离最小的全局最优解,而是将样本显著区分为模长远大于0和模长接近0的两类集合即可。

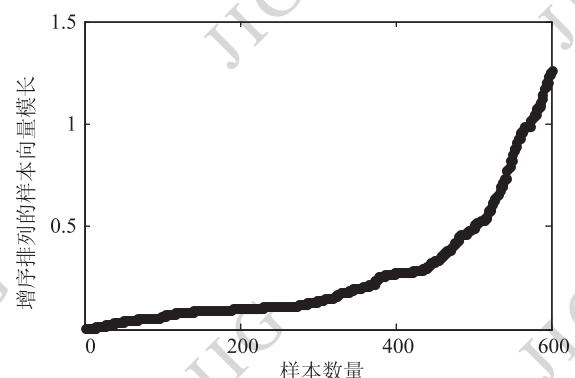
对于一段有效的平地运动数据,一般认为不会出现不包含任何足迹的极端情况。于是样本集合中必存在模长十分接近0的样本,该样本到其他样本的距离则近似于其他样本的模长。因此, σ 可近似取值为某样本的模长,该模长则为区分所有样本模长是否为远大于0或接近0的状态的临界点。

对一个运动序列 M 的左、右脚的两个脚部末端

效应器 E_L, E_R 的聚类样本集 $X = \{x_i\}$,归一化 X 后得到新的聚类样本集 $Y = \{y_i\}$,依据其模长 $|y_i|$ 的大小递增排序后,得到一个样本模长的递增数组,记为 $Z = \{z_j\}$ 。 Z 通常始于一组接近0的 z_j ,其对应的 y_i 多数是足迹帧,其数量一般不少于样本总数的20%,根据人类的步态规律,无外力辅助时常规运动中双脚着地的时间不可能过短,即使是世界级职业短跑运动员双脚着地的时间亦大于整个运动周期的20%^[13]。之后 z_j 会在某个时刻迅速增大,表现出一个斜率显著递增的拐点区域,其对应的 y_i 多数是足迹边界附近的帧,该区域内的样本模长即为 σ 的合适取值。图4(a)为一个步行运动数据的归一化样本 Y 模长的散点图,图4(b)为对 Y 的模长递增排序后得到的递增数组 Z 的散点图,可见其变化趋势符合上述分析的特征。



(a) 按帧排列的归一化样本模长



(b) 增序排列的归一化样本模长

图4 一个步行运动的归一化样本模长图示

Fig. 4 Normalized sample vectors' norms of a walking motion

然而,当尝试为实验数据集中的所有运动数据设定统一的 σ 值时,发现难以找到固定的 σ 值使得所有运动序列的足迹检测均获得可以接受的准确率。因而,仍需为每个运动数据确定独立的 σ 值。

从实验中发现,由于噪声影响, Z 的散点图经常存在细微抖动,使得通过分析其变化趋势的斜率寻找其中的拐点区域不可行。于是,尝试使用衡量数据离散程度的样本方差以寻找 Z 中的拐点区域。

Z 的子集 $\{z_1, z_2, \dots, z_j\}$ 的样本方差表现出一定规律:当 z_j 接近0时, $\{z_1, z_2, \dots, z_j\}$ 的样本方差较小;而当 z_j 位于拐点区域时,则 $\{z_1, z_2, \dots, z_j\}$ 的样本方差会显著增大。根据这一特征,即可通过 $\{z_1, z_2, \dots, z_j\}$ 的样本方差的变化寻找到拐点区域。首先,令 $s_1=0$,并分别计算 $j=2, 3, \dots, F$ 时 $\{z_1, z_2, \dots, z_j\}$ 的样本方差 s_j ,并记 $S=\{s_j\}$ 。而后,设定一个样本方差阈值 r ,若 s_j 是不大于 r 的最后一个 S 中的元素,则 z_j 即为 σ 的值,即,当 $\{z_1, z_2, \dots, z_j\}$ 的样本方差超过容忍程度时,认为此处是足迹边界。图5所示的是图4(b)中的 Z 对应的样本方差 S 的散点图,可以明显观察到 s_j 开始显著增加的位置正位于 z_j 的拐点区域。

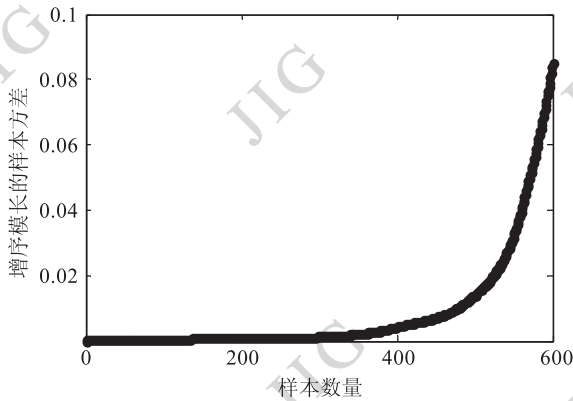


图5 增序排列的归一化样本模长的样本方差分析

Fig. 5 Sample variances of ascending sorted normalized sample vectors' norms

由于运动数据中的足迹边界具有模糊性,即使肉眼也难以精确标定^[2],因此样本方差阈值 r 的确定无需追求单一值。经实验测试,当窗口尺度参数 $w=1$ 时, r 的取值在0.0006~0.002时,实验样本集能取得较高的足迹检测准确率。

2.3 基于谱聚类的足迹检测算法

给定一个长度为 F 的运动序列 M , w 是窗口尺度参数, r 是样本方差阈值,则基于谱聚类的足迹检测算法如下:

1) 计算 M 的左、右脚末端效应器的高度和速度 $h_f^L, v_f^L, h_f^R, v_f^R, f=1, 2, \dots, F$;

2) 令 $X_L = \{[h_f^L, v_f^L]^T\}$, $X_R = \{[h_f^R, v_f^R]^T\}$, 并且令 $X = \{X_L, X_R\} = \{[h_i, v_i]^T\}, i=1, 2, \dots, 2F$;

3) 将 X 归一化处理为 X' , 令 $X' = \{x'_i\} = \left\{ \left[\frac{h_i}{\max(h_i) - \min(h_i)}, \frac{v_i}{\max(v_i) - \min(v_i)} \right]^T \right\}$;

4) 令 $X'_L = \{x'_1, x'_2, \dots, x'_F\}$, $X'_R = \{x'_{F+1}, x'_{F+2}, \dots, x'_{2F}\}$, 用参数 w 对 X'_L, X'_R 进行窗口化处理, 得到新的样本集合 X''_L, X''_R , 令 $Y = \{y_i\} = \{X''_L, X''_R\}$;

5) 计算 $|y_i|$ 并按其大小增序排列, 得到递增数组 $Z = \{z_j\}, j=1, 2, \dots, 2F$, 其中 $|y_i|$ 与 z_j 一一对应;

6) 记 $S = \{s_j\}$, 当 $j=1$ 时令 $s_j=0$, 当 $j>1$ 时令 s_j 为集合 $\{z_1, z_2, \dots, z_j\}$ 的样本方差;

7) 若 $s_j \leq r$ 且 $s_{j+m} > r, m=1, 2, \dots, F-j$, 则高斯核函数的尺度参数 $\sigma = z_j$;

8) 对样本集 Y , 令 $n=2F$, 计算距离矩阵 $A \in \mathbf{R}^{n \times n}$, 当 $i \neq j$ 时, $A_{ij} = \exp\left(-\frac{d^2(y_i, y_j)}{\sigma^2}\right)$, 并且 $A_{ii} = 0$, 其中 $d(y_i, y_j)$ 是 y_i 和 y_j 之间的欧氏距离, $i, j=1, 2, \dots, n$;

9) 定义一对角矩阵 D , 使得 $D_{ii} = \sum_{j=1}^n A_{ij}$, 并构造标准化距离矩阵 $L = D^{-1/2} A D^{-1/2}$;

10) 查找 L 最大的特征向量 $u \in \mathbf{R}^{n \times 1}$, 将 u 的每行看做一个 $\mathbf{R}^1$ 空间的点, 并进行K-means聚类得到两类点集 c_1 和 c_2 ;

11) 指定原始数据 y_i 到类别 c_i , 当且仅当 u 中对应的第 i 行属于类别 c_i 。

12) 分别计算 c_1 和 c_2 的样本均值 μ_1 和 μ_2 , 若 $\mu_1 < \mu_2$, 则 c_1 的标签为足迹帧, c_2 的标签为非足迹帧, 否则, c_2 的标签为足迹帧, c_1 的标签为非足迹帧。

3 实验与分析

3.1 实验设计

实验环境是 Intel Core i3 CPU、2 G 内存、Matlab 7.0 平台。实验数据集由一组 bvh 格式的运动捕捉数据文件组成。这些文件源自3种不同来源: 其中一部分使用 Xsen moven 惯性运动捕捉系统自行采集; 另一部分来自 CMU 运动捕捉数据库(mocap.cs.cmu.edu), 由 Vicon 光学运动捕捉系统采集; 其余部分来自互联网, 采用的运动捕捉系统不详。实验数据集共包含8个角色的16 613帧运动数据, 共35个文件, 采样频率包括24 Hz、30 Hz、60 Hz、120 Hz共4种, 运动类型涉及行走、跑步、跳跃、体育运动、舞蹈等。

实验方法被设计为对比运动数据的手工标记的

标签与算法计算的标签过程,定义相同标签占标签总数的比值为该实验方法的准确率。该方法的结果直观明了,有利于快速确定合适的算法参数。由于足迹边界固有的模糊性,即使人工标记也难以精确划分^[2],所以手工标签本身也存在一定误差,因此实验中使用一小段连续时间替代手工标签中的精确时刻作为足迹边界。

实验过程主要包括3个步骤。首先,对实验数据集进行人工标记,记录每个运动文件的足迹标签 **LBL**。之后,对 **LBL** 中每个独立足迹的边界进行模糊化处理。若 **LBL** 中的一个足迹为始于 t_s 时刻、终于 t_e 时刻的一段时间,则需要对 t_s 和 t_e 进行模糊化处理。记模糊时段长度为 p ,单位为 s ,则若算法结果中的足迹始于 $[t_s - p/2, t_s + p/2]$ 中任一时刻、终于 $[t_e - p/2, t_e + p/2]$ 中任一时刻,均认为算法检测的足迹是正确的。由于人类的反应时间通常介于 $0.1 \sim 0.4 s$,于是分别令 p 的值为 $0.1, 0.2, 0.3, 0.4$ 进行实验。模糊时段 p 在运动数据中表现为连续帧区间,其不同运动数据中所包含的帧数即为运动采样频率与 p 的乘积。最后,将算法计算所得标签与模糊化后的 **LBL** 比较,计算准确率。

3.2 样本方差阈值和模糊时段对准确率的影响

在窗口尺度参数 $w = 1$ 时,将样本方差阈值 r 在 $0.0002 \sim 0.0044$ 范围内按照步长 0.0002 依次取值,对应的本文算法的足迹检测准确率如图6所示。可以看出当模糊时段 p 取不同值时, r 关于准确率的变化趋势基本一致,均存在一个显著的高峰期以及两端的衰减期,且当 $r = 0.001$ 时,准确率达到峰值。并且随着 p 的增大,算法的准确率虽然也在逐渐增加,但增幅越来越小,并且两端的衰退趋势也在逐渐减弱。以上说明了 r 和 p 均存在合适的取值

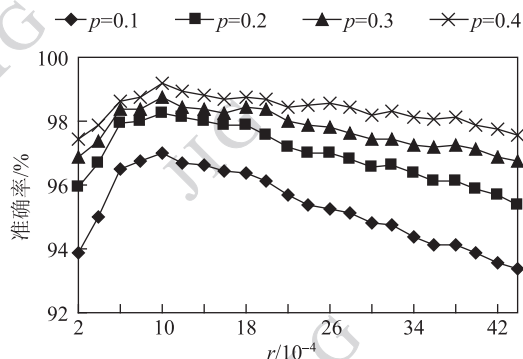


图6 样本方差阈值 r 和模糊时段 p 对整体准确率的影响

Fig. 6 Effect of sample variance threshold r and fuzzy period parameter p on global accuracy rates

范围,不合适的 r 和取值过小的 p 会导致算法整体准确率的降低,过大的 p 又将导致 r 取不同值时的算法准确率之间的差异性减弱。综合考虑, r 的取值在 $0.0006 \sim 0.002$ 、 p 的取值在 $0.2 \sim 0.3 s$ 时,算法的准确率可以达到 98% 左右。

3.3 窗口尺度参数对准确率的影响

改变窗口尺度参数 w 和模糊时段 p 的取值,并将样本方差阈值 r 在 $0.0002 \sim 0.0044$ 范围内按照步长 0.0002 依次取值,则每种 w 和 p 的组合情况下实验获得的最大整体准确率如图7所示。可以看出 w 的取值并非越大则准确率越高,这是由于 w 过大可能导致运动标签在足迹边界处被过度平滑,使得足迹边界处的足迹帧被平滑为非足迹帧,呈现为检测足迹均过短的现象,从而降低了算法准确率。由于目前窗口尺度参数 w 是以帧为单位,故该情况在采样频率较低的运动数据中更为明显,将来可通过结合采样频率进行改进。另一方面, w 取值越大,单个样本向量的长度越长,算法的运算时间越长,增加的时间主要消耗于聚类样本集的距离矩阵的计算。综合考虑, w 取值为 1 或 2 为较为理想的参数。

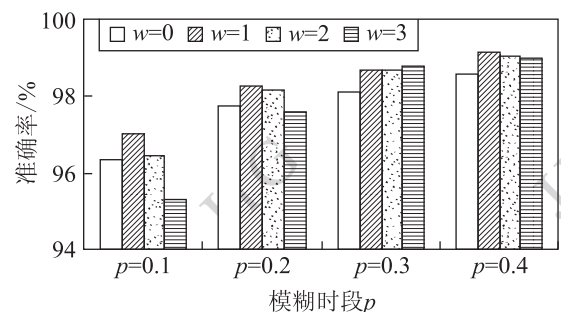


图7 窗口尺度参数 w 对最大整体准确率的影响

Fig. 7 Effect of window scale parameter w on maximal global accuracy rates

值得一提的是,介于实验环境所采用的 Matlab 软件的局限性,算法的运行性能在移植到适当平台后仍有很大的提升空间。同时,样本集的距离矩阵可通过设定距离阈值转换为稀疏矩阵,对稀疏矩阵的特征值和特征向量求解存在高效的解决方案。

3.4 高斯核函数尺度参数对准确率的影响

在窗口尺度参数 $w = 1$ 时,对高斯核函数的尺度参数 σ 统一赋值,令 σ 在 $0.07 \sim 0.47$ 范围内按照步长 0.05 依次取值,对应的本文算法的足迹检测

准确率如图 8 所示。可以看出,当 $\sigma = 0.17$ 时,算法应用于实验数据集的准确率达到峰值。

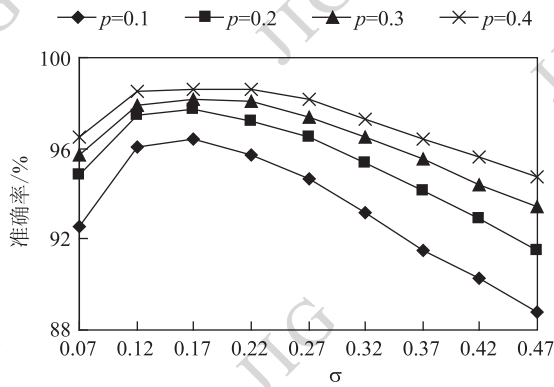


图 8 高斯核函数尺度参数 σ 对整体准确率的影响
Fig. 8 Effect of scaling parameter σ of Gaussian kernel function on global accuracy rates

为了证明在不同运动数据中采用样本方差阈值 r 自适应计算高斯核函数尺度参数 σ 的效果优于直接采用相同 σ 的效果,下面以 $w = 1$ 时图 6 和图 8 中准确率达到峰值时的参数取值,即 $r = 0.001$ 和 $\sigma = 0.17$ 为例,分析足迹检测算法的效果。

图 9 展示了当 p 取不同值时,分别令 $r = 0.001$ 和 $\sigma = 0.17$ 的算法整体准确率,可见前者的准确率略高,且与后者相差不多。然而,这并不足以证明 $r = 0.001$ 时的算法效果更好,因为实验数据集中的

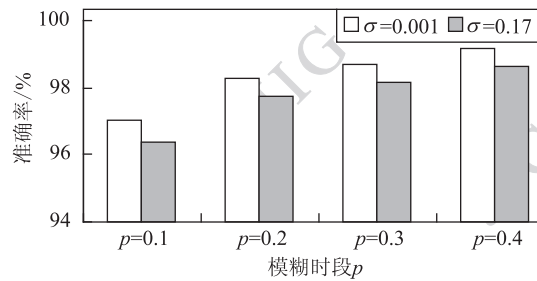


图 9 $r = 0.001$ 和 $\sigma = 0.17$ 时的整体准确率比较
Fig. 9 Comparison of global accuracy rates when $r = 0.001$ and $\sigma = 0.17$

35 个文件的长度、频率各异,每个文件的准确率对整体准确率的贡献并不相同,因此仍需考虑每个文件准确率的分散度。表 1 列举了 p 取不同值时两种条件下单个文件的准确率的最大值和最小值,以及所有文件准确率的标准差。显然, $r = 0.001$ 时所有文件准确率的标准差较 $\sigma = 0.17$ 时更小,即 $r = 0.001$ 时文件的准确率之间分散程度较小,并且随着 p 的增加,这一优势逐渐明显。从而说明采用样本方差阈值 r 自适应计算 σ 的方法对实验运动数据集的整体作用效果较好,对不同运动数据的适应性要明显优于采用统一的 σ 值的方法,因此,本文提出的通过设定样本方差阈值以自适应计算高斯核函数的尺度参数 σ 的方法对于增强算法的普适性是有显著效果的。

表 1 $r = 0.001$ 和 $\sigma = 0.17$ 时的文件准确率
Table 1 Comparison of file accuracy rates when $r = 0.001$ and $\sigma = 0.17$

p	r = 0.001			sigma = 0.17		
	单个文件准确率/%		全部文件准确率的标准差	单个文件准确率/%		全部文件准确率的标准差
	最小值	最大值		最小值	最大值	
0.1	85.1	100.00	0.033 783	84.81	100.00	0.038 27
0.2	90.33	100.00	0.022 245	86.08	100.00	0.033 396
0.3	92.67	100.00	0.018 853	86.08	100.00	0.032 607
0.4	94.33	100.00	0.014 467	87.82	100.00	0.028 453

3.5 其他实验结果

对实验数据集执行本文方法的部分具体实验结果如表 2 所示。可以看出,本文方法对于任意角色的具有清晰足迹的常规运动,如走路、跑步、跳跃等,效果非常好,有效克服了分类器方法不能无条件处理任意角色运动的局限性。

在参数为 $r = 0.001$ 、 $p = 0.3$ 、 $w = 1$ 时,对实验数据集采用本文方法后获得的足迹检测整体准确率为

98.72%。同时,在参数 $p = 0.3$ 时对实验数据集也分别采用高度基准线、速度基准线以及综合高度和速度的双基准线方法进行实验。当对所有文件采用统一基准线时,上述 3 种基准线方法的整体准确率为 86.64%、81.72% 和 90.96%;当人工为每个文件分别指定合适的高度或速度基准线时,3 种方法的整体准确率为 92.32%、96.33% 和 96.54%。对表 2 中的文件采用 3 种基准线方法的具体实验结果如表 3

表2 采用本文方法时部分运动数据的文件准确率
Table 2 Details of file accuracy rates by our method

文件编号	帧数	采样频率/Hz	运动类型	文件准确率/%		
				$r=0.0006,$ $p=0.2, w=1$	$r=0.001,$ $p=0.3, w=1$	$r=0.002,$ $p=0.3, w=1$
1_01	500	30	连续蛙跳	100.00	100.00	100.00
2_01	245	120	弧线走路	100.00	100.00	100.00
2_02	370	120	不规则跑跳	100.00	100.00	100.00
3_02	791	24	醉酒踉跄	97.85	99.30	99.37
3_03	361	24	直线慢跑	100.00	100.00	100.00
4_02	1 000	120	单脚拖步行走	93.49	94.59	97.90
5_01	300	60	跳现代舞	96.12	96.63	92.82
6_02	316	30	打保龄球	98.42	99.05	99.05
7_02	298	24	跳芭蕾舞	96.81	98.15	95.13
7_04	744	24	跛行到跛跑	98.86	99.46	98.99
7_07	336	24	单脚连跳	100.00	100.00	100.00
7_09	462	24	蹦蹦跳跳前进	99.68	100.00	99.89
8_03	171	30	左盾右剑格斗	95.61	94.74	95.03

表3 采用基准线方法时部分运动数据的文件准确率
Table 3 Details of file accuracy rates by baseline methods

文件编号	运动类型	全部文件采用统一基准线,高度=1.4、速度=1.2			单个文件手工指定基准线				
		高度基准线法的准确率/%	速度基准线法的准确率/%	双基准线法的准确率/%	高度	速度	高度基准线法的准确率/%	速度基准线法的准确率/%	双基准线法的准确率/%
1_01	连续蛙跳	87.5	82.1	87.5	1.81	0.16	97.1	100	100
2_01	弧线走路	100	87.6	100	0.68	0.2	100	100	100
2_02	不规则跑跳	100	97.3	100	1.31	0.46	97.8	100	100
3_02	醉酒踉跄	92.7	99.3	99.4	0.78	0.77	95.2	99.8	97.3
3_03	直线慢跑	100	98.2	98.3	0.45	1.85	100	99.9	100
4_02	单脚拖步行走	81.9	79.1	87.8	0.62	0.42	76.2	97.3	88.4
5_01	跳现代舞	85.3	69.5	87	1.11	0.6	85.2	82	91
6_02	打保龄球	42.1	97	46.2	3.15	0.67	77.5	97.9	86.6
7_02	跳芭蕾舞	83.7	97.8	96.8	0.4	0.72	90.1	99.5	96.8
7_04	跛行到跛跑	95.8	95	95.8	1.75	0.56	96.9	98.5	98.8
7_07	单脚连跳	98.8	84.7	100	0.64	0.89	99.6	91.7	100
7_09	蹦蹦跳跳前进	100	99.8	100	0.72	1.23	99.7	99.5	100
8_03	左盾右剑格斗	89.5	98.5	100	1.45	0.72	87.1	100	100

所示。显然,由于实验数据的来源多样,数据量纲差异较大,统一基准线方法很难良好适用于所有数据文件。而采用高度或速度的单一特征的基准线方法则难以有效处理一些特殊运动,如单脚粘连地面的拖步行走、单脚持续腾空的单脚连跳等。对于同时考虑高度和速度的双基准线方法,虽然可以改善使用单一特征对特殊运动造成的误判情况,但是基准线方法从根本上未能避免人工参与的过程,当应用于大数据集时难免导致效率低。而本文的方法可以有效克服以上不足,当应用于混合数据集时表现出的稳定性良好。

在算法的时间效率方面,一个 100 帧、200 帧和 300 帧的运动片段处理过程分别需时约 0.41 s、1.94 s 和 5.29 s。随着运动片段长度的增加,样本集的距离矩阵的维数呈平方倍数增长,故时间消耗

亦以平方倍数增加。不过,时间效率仍存在提升的空间,一方面 Matlab 平台本身的运算机制导致其效率有限,另一方面谱聚类算法本身仍存在优化的可能。在今后的工作中,将通过移植算法到更高效的平台,以及对距离矩阵通过设定距离阈值转换为稀疏矩阵,而后采用现有的针对稀疏矩阵的特征值和特征向量求解的成熟高效方案以优化谱聚类算法的时间效率,有望大幅提高算法的整体执行效率。

图 10 则展示了 3 个运动捕捉数据的部分片段的足迹检测结果,图 10 中的散点为表演者的两个脚尖的高度运动轨迹,从上至下的运动类型依次为不规则跑跳、单脚连跳和跳芭蕾舞。通过更多的实验结果还发现,惯性运动捕捉系统采集的运动数据比其他数据的执行结果好,这是由于其他数据的脚部运动大多存在噪声作用下的严重抖动。

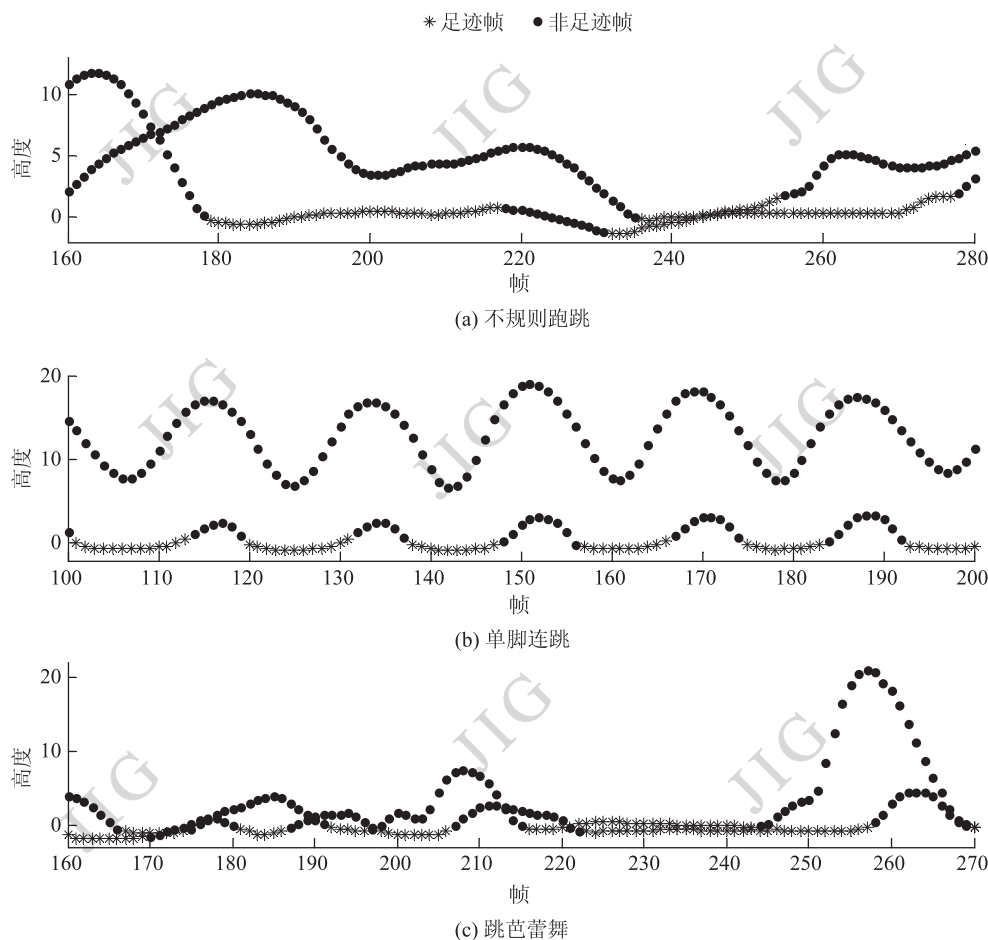


图 10 双脚尖高度轨迹上展示的足迹检测结果

Fig. 10 Foot plant detection results showed on foot end-effectors' height trajectories

实验效果不理想的情况大多产生于脚部的近地轻微运动状态。此时脚部的高度与速度均较小,使

得聚类样本的间距与运动数据的噪声位于相同数量级别,因此无法将该类小动作从噪声中区分出来。

对于这类问题,通过采用高频采样数据、对运动数据进行去噪预处理或者录制一小段相同角色的静止站立的运动片段参与聚类等方法,增加样本间距和数据噪声之间的数值差,可以适当改善。

4 结 论

由于几乎所有的运动编辑方法都会破坏原始运动捕捉数据中的足迹约束,所以足迹检测是运动编辑的重要前端处理,而运动捕捉数据中的采样噪声干扰使得直接检测足迹非常困难。现有的足迹检测方法在运动类型、采样对象等方面均存在不同程度的局限性,难以应对当前大数据背景下的来源和属性多样化的混合运动数据集。

本文提出一种面向运动捕捉数据的基于谱聚类的足迹自动检测方法,可应用于大部分运动表演者的平地运动数据,克服了现有方法的不足。通过逐帧提取脚部的运动特征进而表示为样本,再对样本集使用谱聚类方法划分出属于足迹部分的运动帧,以实现运动数据的足迹检测,并提出了一种高斯核函数尺度参数的自适应选择方法,以及通过实验确定了其他参数的最优取值,使得谱聚类过程避免了需要通过人工交互确定参数的传统步骤。实验结果表明了当方法应用于混合运动组成的实验数据集时,对于各式各样的运动数据均可获得较高的足迹检测准确率,可见其对不同类型运动的普适性良好。其中,对于步行、跑步等常规运动的足迹检测准确率可达100%;同时对于体育运动、舞蹈等复杂运动和单脚拖步行走等特殊运动的足迹检测准确率也可达95%以上。下一步工作将聚焦于窗口尺度参数、高斯核函数尺度参数的自适应选择方法的改进以及谱聚类算法效率的优化上。

参考文献 (References)

- [1] Kovar L, Schreiner J, Gleicher M. Footskate cleanup for motion capture editing[C] //Eurographics Symposium on Computer Animation. New York: ACM Press, 2002: 97-104. [DOI: 10.1145/545261.545277]
- [2] Ikemoto L, Arikian O, Forsyth D. Knowing when to put your foot down[C] // Symposium on Interactive 3D Graphics and Games. New York: ACM Press, 2006: 49-53. [DOI: 10.1145/1111411.1111420]
- [3] Wu J, Popović Z. Terrain-adaptive bipedal locomotion control[J]. ACM Transactions on Graphics, 2010, 29(4): #72. [DOI: 10.1145/1778765.1778809]
- [4] Ménardais S, Kulpa R, Multon F, et al. Synchronization for dynamic blending of motions[C] //Eurographics Symposium on Computer Animation. Aire-la-Ville: Eurographics Association Press, 2004: 325-335. [DOI: 10.1145/1028523.1028567]
- [5] Glardon P, Boulic R, Thalmann D. Robust on-line adaptive foot-plant detection and enforcement for locomotion[J]. The Visual Computer, 2006, 22(3): 194-209. [DOI: 10.1007/s00371-006-0376-9]
- [6] Glardon P, Boulic R, Thalmann D. On-line adapted transition between locomotion and jump[C] //Proceedings of Computer Graphics International. Washington DC: IEEE Computer Society Press, 2005: 44-50. [DOI: 10.1109/CGI.2005.1500367]
- [7] Zou X L, Chen W F, Feng G C. Fast image segmentations of Deut[J]. Journal of Image and Graphics, 2012, 17(2): 222-228. [邹小林, 陈伟福, 冯国灿. 判别割(Deut)的图像分割及其快速分割算法[J]. 中国图象图形学报, 2012, 17(2): 222-228.] [DOI:10.11834/jig.20120210]
- [8] Zelnik-Manor L, Perona P. Self-tuning spectral clustering[C] // Proceedings of Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 2005: 1601-1608.
- [9] Deng X Z, Jiao L C, Lu S. Spectral clustering ensemble applied to SAR image segmentation using nonnegative matrix factorization[J]. Acta Electronica Sinica, 2011, 39(12): 2905-2909. [邓晓政, 焦李成, 卢山. 基于非负矩阵分解的谱聚类集成 SAR 图像分割[J]. 电子学报, 2011, 39(12): 2905-2909.]
- [10] Ng A, Jordan M, Weiss Y. On spectral clustering: analysis and an algorithm[C] //Proceedings of Advances in Neural Information Processing Systems 14. Cambridge, MA: MIT Press, 2002: 849-856.
- [11] Zhao X. Kinematic control of human postures for task simulation[D]. Philadelphia, PA: University of Pennsylvania, 1996.
- [12] Jain A K. Data clustering: 50 years beyond K-means[J]. Pattern Recognition Letters, 2010, 31(8): 651-666. [DOI: 10.1016/j.patrec.2009.09.011]
- [13] Novacheck T F. The biomechanics of running[J]. Gait & Posture, 1998, 7(1): 77-95. [DOI: 10.1016/S0966-6362(97)00038-6]