

文献引用格式:李红竹.舞蹈视频图像中动作识别方法研究[J].电视技术,2018,42(7):34-37.

LI H Z. Research on motion recognition method in dance video images[J]. Video engineering, 2018, 42(7):34-37.

中图分类号:TN957.52

文献标志码:B

DOI:10.16280/j.ideoe.2018.07.008

舞蹈视频图像中动作识别方法研究

李红竹

(阿坝师范学院,四川 汶川 623002)

摘要:运用计算机及“看”视频是当前人机交互领域的一个热门话题。以计算机视觉对舞蹈视频图像中舞者的动作进行识别意义重大。以舞蹈视频图像动作识别为切入点,提出对舞蹈视频图像动作进行灰度转换、阈值化处理、背景减除、降噪等预处理,对舞蹈视频图像特征提取及关节点建模进行分析,形成了舞蹈视频图像动作识别的系统思路。

关键词:人机交互;舞蹈;视频图像;预处理;关节点;识别

Research on Motion Recognition Method in Dance Video Images

LI Hongzhu

(Aba Teachers University, Wenchuan 623002, China)

Abstract: Using computers and "looking" videos is a hot topic in the field of human-computer interaction. It is significant to identify the dancer's movements in dance video images with computer vision. Taking dance video image motion recognition as the entry point, the dance video image motion is processed by grayscale transformation, thresholding, background subtraction, noise reduction, etc., and the dance video image feature extraction and joint point modeling are analyzed. The systematic idea of motion recognition of dance video images.

Key words: human-computer interaction; dance; video image; preprocessing; articulation; recognition

人类视觉系统对信息的获取存在很多局限,不利于对信息进行理解和处理。计算机视觉是一种可以帮助人类准确获取信息的辅助手段,它通过让机器“看”,实现对人类智能的模拟、拓展甚至延伸,可以解决需要大规模的、复杂的视觉任务^[1]。目前,计算机视觉被广泛应用于图像生成、车辆监控、目标跟踪等,在医疗、农业和交通等各个领域都有着广阔的应用前景。近几年,计算机图像处理能力不断提升,计算机视觉的发展也日趋成熟。将计算机视觉应用到舞蹈视频图像的动作识别中,有多重意义。如可以作为一种辅助训练方式应用到舞蹈学习或教学中,它可以准确全方位识别视频中舞蹈者的动作姿势,便于对其进行客观评价,从而为舞蹈者提出改善建议,还可以作为智能体感游戏的设计手段,通过捕捉视频中舞蹈者的动作将其生成游戏图像,给玩家提供非凡的交互及体验效果。

1 舞蹈视频图像动作识别难点

动作识别经历了漫长的发展历程,手势识别、肢体动作识别等都属于低级阶段的人体运动分析。舞蹈视频图像中的动作识别本质是运动识别,属于高级阶段的人体运动分析,目的在于对视频中的运动动作进行分析、评判或提取,可创建个性化的教学系统、训练系统和虚拟仿真系统等。相较于手势识别、肢体动作识别,舞蹈视频图像中的动作识别更为复杂,运用简单的肢体定位算法并不能准确反映舞蹈视频图像中的运动动作。其难点主要表现在几个方面。

1.1 舞蹈动作运动系统复杂

舞蹈动作自由度较大,相较于人体的自然运动,舞蹈动作变化更多。常规的有俯冲拧扭踢、仰蹲屈躺伸等动作,比较复杂的则有摇摆步、滑步、云步、蹉步、飞脚、旋转、串翻身等动作。这些动作包含较多

自由度,简单的建模方法和动作识别方法难以准确表达舞蹈者的人体动作。

1.2 舞蹈动作遮挡问题突出

舞蹈动作中存在遮挡、自遮挡问题。当舞蹈视频中,只有一个舞者时,存在自遮挡问题,当舞者数量多于一个时,又会产生互相遮挡问题。由于拍摄视角的原因,会给舞者运动分析造成很大障碍,需要更好的方法解决舞者的人体遮挡问题。

1.3 舞蹈动作检测难度较大

舞者动作目标检测面临干扰因素多、难以准确检测的问题。对舞蹈视频图像中舞者动作的检测,视频中视角、光照条件、背景、舞者自身影子等多种因素干扰,很难保证检测的准确性、全面性,导致方法对运动目标进行准确分割,增加动作识别难度。同时,舞蹈动作具有极强的连续性,肢体的各个动作之间几乎是一气呵成,每个动作之间的衔接非常流畅,舞蹈视频中捕捉到的动作数据通常是一连串帧,给检测识别各个动作的边界增添了难度。

2 舞蹈视频图像动作预处理技术

舞蹈视频图像动作识别主要是通过计算机视觉系统进行的,但考虑到计算机运算量的问题,应在计算机识别前先对其进行预处理,以提取有效的图像信息,减轻计算机的运算负担。

2.1 将视频图像变换为灰度图像

舞蹈视频中的图像多是彩色的,如果不转换直接将其输入计算机视觉系统,会增加图像输入的信息量,造成后续运算量的加大。对视频中的彩色图像进行灰度处理则可以减少后续运算量,提高计算机视觉系统动作识别效率。彩色图像灰度化处理原理即对色彩空间 RGB 的三维通道进行降维,使之转换为一维^[2]。彩色图像灰度化过程如图 1 所示。

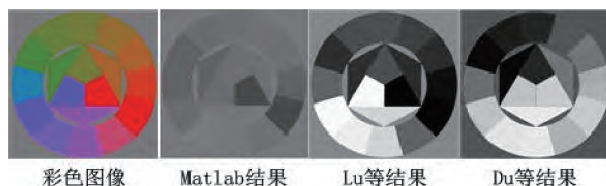


图 1 彩色图像灰度化过程

图像灰度化变换的方法很多,如单分量法、最大值法、全局映射法等,不同方法适用处理对象不同,采用不适用的灰度变换方法,极有可能导致不理想的灰度化效果。结合舞蹈视频图像的动作特征,本

文采用一种图像融合基础上的混合法灰度化处理方法。采用普拉斯加权计算 R,G,B 各个通道输入局部的变换总和,采用全局对比加权映射法通过灰度值、颜色距离差等要素来建立目标函数,然后通过最小化目标函数得出灰度化结果^[3]。采用这种混合型方法,可以最大程度保持图像的全局结构信息及局部对比度信息。通过此方法对舞蹈视频中的图像进行灰度化处理,效果如图 2 所示。



图 2 舞者图像灰度化处理

2.2 对视频图像进行阈值化处理

图像阈值化处理的本质是对舞蹈图像进行分割,目的在于压缩彩色图像的数据量,简化图像分析步骤。它按照灰度集划分像素集合,这个集合的每个子集对应各个舞蹈动作。在对舞蹈视频图像进行阈值化处理时,确定合适的阈值极为关键,因为它关系到图像中的像素所属位置。只有阈值合理,才能产生比较准确的二值图像。阈值一般用如下形式表示。

$$T = T[x, y, f(x, y), p(x, y)] \quad (1)$$

在这个式子中, $f(x, y)$ 是在像素点 (x, y) 处的灰度值; $p(x, y)$ 为该点的灰度梯度函数,通过这个式子就可以得出二值化后的图像。

2.3 对视频图像进行分割处理

对舞蹈视频图像进行分割是为了改变舞蹈图像表示形式,让图像更便于被计算机视觉系统理解和分析。图像分割法通过定位舞蹈动作的边界对其进行分割,使得图像中每个像素都有自己独有的标签,最终形成一个多像素集合。在对舞蹈动作图像进行阈值化处理得到二值图像之后,需要将舞者的运动区域从整个场景中进行分离,这就需要对舞者的运动区域进行分割,确定合适的阈值,以便找到舞者人

体轮廓边缘。如,设定当前时刻为 t ,视频图像在 t 时刻的帧检测值为 $p(x,y,t)$,得到的尺寸是 $M \times N$,那么此时,得到的二值图像就为 $A(x,y,t)$ 。接着可以用几种不同的方法对舞蹈视频图像进行分割,如幅度分割法,根据各个动作幅度来分割各个动作图像区域,区域分割法,根据动作形状来分割区域。

2.4 对视频图像进行背景减除及降噪

舞蹈视频图像中,除了舞者这个前景目标之外,还有许多其他背景,如场地器材、环境、灯光等,这些多要素背景的存在会增加计算机视觉系统对舞者动作识别的复杂度,通过背景减除可以获得前景目标,即完整的人体区域,可以降低动作识别的难度。目前背景减除的研究成果较多,本文基于舞蹈视频图像的特殊性,选择高斯混合模型来实现背景减除。其原理是使用 K 个(一般为5个左右)高斯模型表征舞蹈视频图像中的每一个像素点特征。用舞蹈视频图像中各个不同的像素点去匹配高斯模型,如果匹配成功,就可以判断其为背景点,如果匹配失败,就判定其为前景点,以此来实现背景与前景的分离。

(1) 建立模型

需要先假设不同时刻像素点的值,如假设在 t 时刻,那么像素点的值就为 X_t ,它发生的概率公式则如下:

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \cdot \eta(X_t, \mu_{i,t}, \delta_{i,t}) \quad (2)$$

(2)中, $\omega_{i,t}$ 表示在时刻 t 时,第 i 个高斯分布的权重, $\eta(X_t, \mu_{i,t}, \delta_{i,t})$ 则表示其对应的概率密度函数, $\mu_{i,t}$ 表示对应均值, $\delta_{i,t}$ 则为方差。

(2) 更新与匹配模型

对新输入的一帧图像,假设其像素点值为 X_t ,那么判断像素与高斯模型是否匹配的公式如下:

$$|X_t - \mu_{i,t-1}| \leq 2.5\delta_{i,t-1} \quad (3)$$

如果能够满足这个公式条件,那么就可以判断像素点与模型匹配,是背景点,如果不能满足,就可以判断其与模型不匹配,是前景点。

通过高斯模型获得背景减除效果图如图3所示。

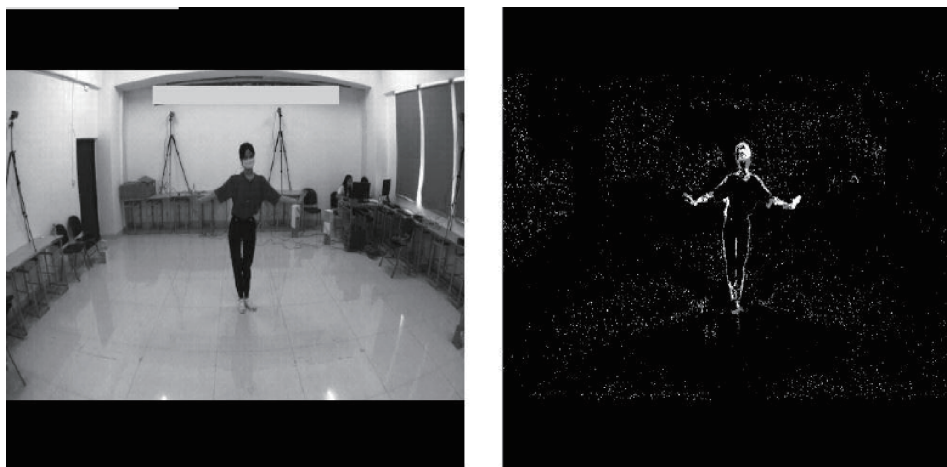


图3 高斯模型背景减除效果

(3) 降噪

从高斯模型背景消除效果图中可以看出,前景周围存在噪声,会对计算机视觉系统识别目标图像边缘造成干扰脉冲,因此,可以通过降噪操作来达到保护舞者人体图像边缘的目的。考虑到便捷性、实用性等原则,本文采用中值滤波法来降噪。具体过程如下:第一步,在舞蹈视频图像中移动模板,直到模板中心在移动过程中与舞蹈视频中的中心像素点重合,这个像素点就可以作为窗口中心,以这个中心来构建方形、圆形等不同形状的窗口。第二步,读取模板下窗口中所有像素的灰度值,然后对灰度值进

行排序,一般按照从小到大顺序排列。第三步,计算像素排列后灰度值的中值,并将其作为窗口中心点像素灰度值。中值滤波降噪后输出像素灰度值计算公式如下。

$$g(x,y) = \text{median}\{f(x-i,y-j) \mid (i,j) \in W\} \quad (4)$$

这个公式中, W 代表模板窗口, $g(x,y)$ 指的是输出的像素灰度值, $f(x-i,y-j)$ 指的是输入的像素灰度值。

3 舞蹈视频图像动作姿态提取及关节建模

3.1 采用姿态特征提取法识别动作特征

(1) 舞者动作识别特征分类

舞者的舞蹈动作和普通人的日常动作之间存在很大区别,很多动作都需要舞者手腿并用才能完成,在选择目标区域进行背景识别时,必须要掌握舞者的全身运动信息方能准确识别其动作。舞者的动作识别可分为几类:静态特征,主要表现为舞者人体目标的大小、颜色、身体轮廓、深度等,它能够传递舞者动作的整体信息,如通过舞者的轮廓特征可以得出其当前的基本形状;动态特征,主要表现为舞者

的运动速度、方向和轨迹,这些能够反映舞者的移动路径。对这些特征识别能够计算出舞者的运动方向特征,为建模创造条件;时空特征主要表现为时空形状、兴趣点等;描述性特征,包括舞者所处场景、周围物体、姿态等^[4]。

(2) 舞者姿态特征提取

采用姿态特征提取法时可以使用姿态估计传感器,它在动作跟踪领域和机器人视觉领域的运用十分普遍,能够确定舞者动作的方位指向^[5]。可以利用姿态估计器中的光流数值对图像中背景信息进行过滤,获得舞者关节坐标的区域,并能消除舞者服装等因素对舞者动作的遮蔽及影响,舞者姿态特征图如图4所示。

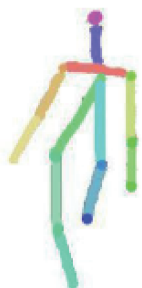


图4 舞者姿态特征

3.2 采用 Kinect 进行舞者关节识别建模

采用 Kinece 法是将人体看作是 25 个关节点坐标组成的一个坐标轴,以这些关节点来建立舞者的

人体骨骼结构,得到舞者人体骨骼模型^[6]。如图5所示。

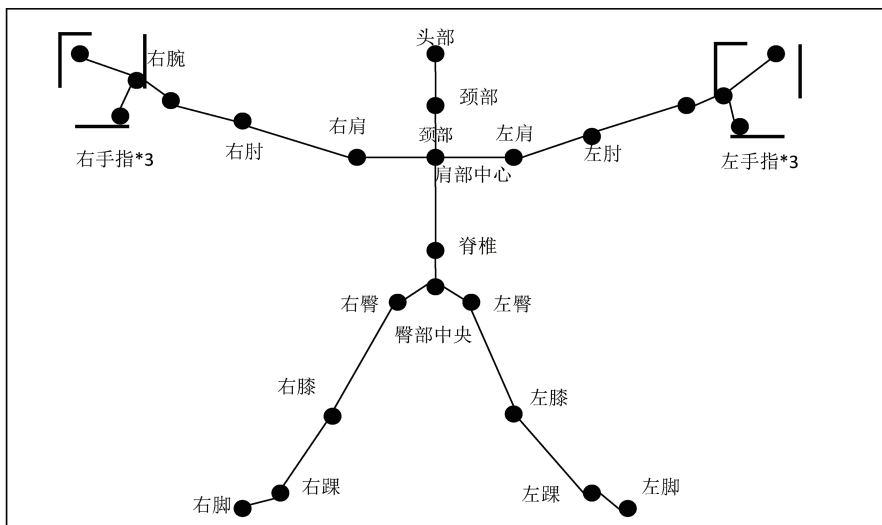


图5 Kinect 法舞者关节识别模型

(下转第52页)

习者对德育知识的理解更加深刻,充分激发他们在德育实践活动中的主动性和积极性。设计沉浸式的VR场馆,给学生提供感受历史、感受国家民族情怀、体验红色文化的虚拟实践空间,既能满足他们的探索欲望,又能促使他们理论联系实际,将德育知识与德育实践有机融合。将这种崭新的资源形式应用到德育教学中,既能突破德育实践教学时空限制,还能让学生进行沉浸式学习,可实现智慧德育的升级。设计这款实践学习系统,需要学校投入一定的物力、人力和财力购置相应的设备,到各个实体场馆进行模型采集,这是确保系统真实性的关键,还需要所有德育课程教师与设计人员配合,将VR场馆系统与德育实践活动紧密结合起来,设计更加丰富的体验活动。

参考文献:

- [1] 张志祯. 虚拟现实教育应用:追求身心一体的教育——从北京师范大学“智慧学习与VR教育应用学术周”说起[J]. 中国远程教育, 2016(6):5-15+79.

- [2] 刘宇. 大学生德育实践模式创新研究——以河源职业技术学院为例[J]. 内蒙古师范大学学报(教育科学版), 2014(9):56-58.
- [3] 郭云鹏,张弓. 虚拟现实技术的应用研究及发展趋势[J]. 电视技术, 2017(23):129-134.
- [4] 李勋祥,游立雪. VR时代开展实践教学的机遇、挑战及对策[J]. 现代教育技术, 2017(7):116-120.
- [5] 张必胜. 三维激光扫描与增强现实技术在虚拟场馆建设的应用[J]. 湖南科技大学学报(自然科学版), 2017(2):40-46.
- [6] 周忠,周颐,肖江剑. 虚拟现实增强技术综述[J]. 中国科学:信息科学, 2015(2):157-180.



作者简介:

夏振鹏(1983—),博士在读,主要研究方向:中国特色社会主义理论与实践;

任彦(1982—),博士在读,讲师,主要研究方向:思想政治教育教学研究、马克思主义中国化研究。

责任编辑:胡鑫

收稿日期:2018-04-13

(上接第37页)

从模型中可见,舞者的关节节点主要分布在四肢。头部、颈部、脊柱、肩部中心各有一个关节节点,关节节点分布最集中的地方则位于上肢,左上肢有左肩、左肘、左腕、左手指等关节节点;右上肢有右肩、右肘、右腕、右手指等关节节点,左下肢有左臀、左膝、左踝、左脚等关节节点;右下肢则有右臀、右膝、右踝、右脚等关节节点。以这些关节节点来建构模型,其原理是在舞者做各类动作的过程中,准确记录各个关节节点的运动,准确识别舞者的各个动作,从而输出正确的舞者运动骨架,有了舞者运动骨架模型,就能显著提高计算机视觉系统对舞者动作的识别准确率和效率,整个识别过程如图6所示。

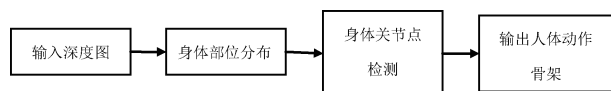


图6 舞者关节节点识别过程

4 结语

舞蹈视频图像识别要考虑到舞蹈背景、服装等对动作识别的影响,还要考虑舞者自身动作中的遮挡、自遮挡问题,要采用一种能够准确完整记录和反映舞者动作信息的动作识别技术,以便获得舞者的身体静态信息和动作信息。本文在已有研究

的基础上,提出了具体的识别方法,这些方法,贴合舞者动作特征,能够准确检测、识别舞蹈视频图像中舞者的动作,可提高识别的准确率和效率,对舞蹈视频动作分析及其他方面的应用有重要价值。

参考文献:

- [1] 陈利峰. 舞蹈视频图像中人体动作识别技术的研究[J]. 现代电子技术, 2017(3):51-53+57.
- [2] 罗会兰,冯宇杰,孔繁胜. 融合多姿势估计特征的动作识别[J]. 中国图象图形学报, 2015(11):1462-1472.
- [3] 尹建芹,刘小丽,田国会,等. 基于关键点序列的人体动作识别[J]. 机器人, 2016(2):200-207.
- [4] 段立娟,郭亚楠,乔元华,等. 基于判别性区域提取的视频人体动作识别方法[J]. 北京工业大学学报, 2017(10):1480-1487.
- [5] 覃剑,王美华. 采用在线高斯模型的行人检测候选框快速生成方法[J]. 光学学报, 2016(11):169-179.
- [6] 刘博,安建成. 基于关键姿势的人体动作识别[J]. 电视技术, 2014(5):38-41.



作者简介:

李红竹(1982—),讲师,硕士,研究方向:音乐学。

责任编辑:徐弘涛

收稿日期:2018-06-27