

系统仿真学报
Journal of System Simulation
ISSN 1004-731X, CN 11-3092/V

《系统仿真学报》网络首发论文

题目：基于 Kinect 的皮影动作实时匹配研究
作者：唐传谦，刘志强，苏毅俊，刘晓静
收稿日期：2020-10-10
网络首发日期：2021-01-05
引用格式：唐传谦，刘志强，苏毅俊，刘晓静. 基于 Kinect 的皮影动作实时匹配研究. 系统仿真学报. <https://kns.cnki.net/kcms/detail/11.3092.v.20210104.1909.022.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于 Kinect 的皮影动作实时匹配研究

唐传谦, 刘志强, 苏毅俊, 刘晓静

(青海大学 计算机技术与应用系, 青海省西宁市, 810016)

摘要: 皮影文化传承中由于受众老龄化和传承断代化使得皮影文化逐渐面临没落, 基于 Kinect 实现皮影动作实时匹配能够为传统皮影文化注入新的活力。根据皮影特点构建关节点皮影模型, 实现了皮影的静态数字化。基于 Kinect 获取人体深度图像, 通过分割遮罩和机器学习获取人体骨骼点坐标, 生成人体骨架。对关节点皮影模型进行骨骼绑定与权重设置, 将 Kinect 采集到的人体数据绑定到皮影上。通过骨骼树的遍历和动作匹配算法完成坐标映射, 实现人体与皮影动作的实时匹配, 从而完成皮影的动态表演数字化。

关键词: 关节点皮影模型; Kinect; 深度数据采集; 动作捕捉; 动作实时匹配

中图分类号: TP391.9

文献标识码: A

Research on Real-time Matching of Shadow Play Based on Kinect

Tang Chuanqian, Liu Zhiqiang, Su Yijun, Liu Xiaojing

(Department of Computer Technology and Application, Qinghai University, Xining, Qinghai Province, 810016, China;)

Abstract: In the inheritance of shadow play culture, due to the aging of the audience and the dying of inheritance, the shadow play culture is gradually facing decline. Real-time matching of shadow play movements based on Kinect can inject new vitality into traditional shadow play culture. According to the characteristics of shadow play, a joint point shadow play model was constructed, and the static digitization of shadow play was realized. The human body depth image is obtained based on Kinect, and the human skeleton point coordinates are obtained through segmentation mask and machine learning to generate the human skeleton. Bone binding and weight setting are performed on the joint point shadow play model, and the human body data collected by Kinect is bound to the shadow play. Through the traversal of the skeleton tree and the action matching algorithm, the coordinate mapping is completed to realize the real-time matching of the human body and the shadow play action, thereby completing the digitalization of the dynamic performance of the shadow play.

Keywords: joint point shadow model; Kinect; depth data collection; motion capture; real-time motion matching

引言

皮影戏是一种以人物剪影为表演形式的中国

传统文化和表演艺术, 蕴含着深厚的文化和艺术特征^[1]。随着时代的发展, 人们对皮影的关注越来越少, 皮影面临着文化和传承断代的困境^[2], 亟需通过现代化的手段对皮影进行传承和保护。

对于皮影艺术的传承和保护, 可从以下两个方面进行: 首先, 皮影的原材料一般为动物毛皮, 制作技艺较为复杂, 且极易损坏。因此, 针对皮影实体, 可根据皮影特点构建数字化皮影形态, 实现皮



收稿日期: 2020-10-10 修回日期: 2020-11-11
基金项目: 国家自然科学基金: 基于文化特征的唐卡图像建模绘制及艺术评价体系研究 (编号: 61862053)
作者简介: 唐传谦(1993-), 女, 重庆梁平, 硕士, 助理实验师, 研究方向为信息可视化与媒体计算; 刘志强(1981-), 男, 重庆, 讲师(高工), 硕士, 研究方向为虚拟现实应用; 苏毅俊(1998-), 男, 青海格尔木, 本科学生。

影的静态数字化。其次,皮影的表演艺术随着皮影社戏的衰落逐渐衰落,传承人断代化使得皮影文化的传承受到一定的影响,基于表演艺术现状,应考虑皮影的动态表演数字化。皮影的静态数字化和动态表演数字化,旨在通过多种模式的数字化使得皮影艺术的长远发展。

本文根据皮影关节的运动特征和艺人表演方式,提出了皮影实体三维数字化和基于 Kinect 的多用户动作捕捉与实时匹配的交互表演数字化。此数字化方式为传统皮影文化注入了新的活力,创新了艺术的传播方式。

1 关节点皮影模型

皮影制作的原材料为动物毛皮,制作工艺复杂,且容易损坏,总结皮影实体的特点就是易损坏难制作。国家大力倡导非物质文化遗产数字化保护以来,数字化技术也越来越先进,开始针对不同非物质文化遗产的特点采取不同的数字化保护手段。本文针对皮影实体的特点,构建关节点皮影模型^[3]实现皮影实体的静态数字化。

关节点皮影模型的构建分为两个阶段,第一个阶段根据皮影厚度薄的特点,采用基于轮廓路径建模方法构建三维关节点皮影模型。轮廓提取算法为数字化提供了轮廓信息,以此轮廓为路径生成皮影各肢体的基础体,对各肢体进行整体拼接。具体实现方法如图 1 所示。

由于皮影模型在表演中是运动的,因此为保证关节点皮影模型的真实感和肢体灵活性,第二个阶段对模型进行骨骼绑定及蒙皮^[4]。根据皮影的表演方法:三杆控制表演、五杆控制表演方法和人体骨骼结构,对皮影模型进行骨骼绑定。本文通过形体模式和 Bipe 骨骼调整^[5]进行骨骼绑定。图 2 所示为皮影骨骼绑定。

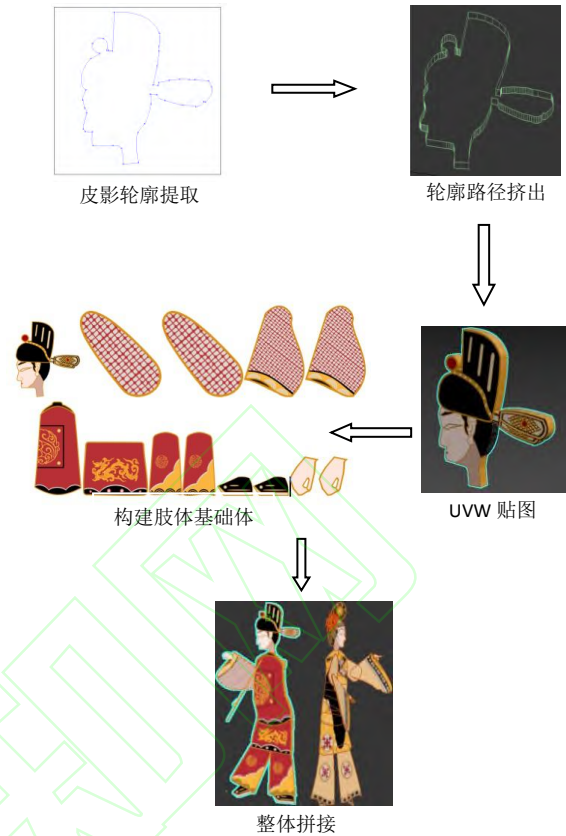


图 1 基于轮廓路径构建关节点皮影模型
Fig1. Construction of joint point shadow play model based on contour path

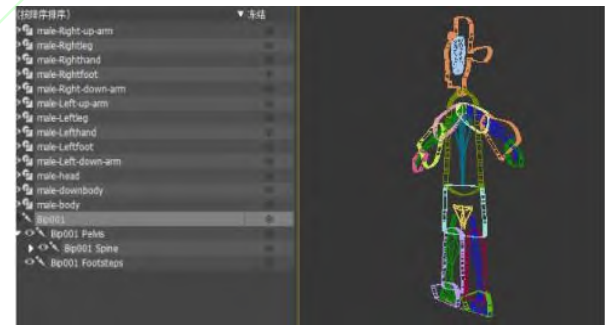


图 2 关节点皮影模型骨骼绑定
Fig2. Joint point shadow play model bone binding

调整完骨骼与模型的位置之后,为使得关节点皮影模型可随着骨骼的移动、旋转和缩放产生形变,需对皮影进行蒙皮。通过骨骼权重设置和对模型的不同部位进行封套^[6],使得每一节骨骼在运动时所带动的部位正常活动。当部分骨骼相邻过近,骨骼的权重分布重合,不调整权重会导致部分关节不符合运动特征,出现穿模现象。当对骨骼关节点

进行权重设置和封套后,这些顶点就能正确地依附在骨骼上。如图 3 所示为进行了权重设置和封套效果和未进行权重设置与封套效果的对比。经过上述两个阶段,完成了关节点皮影模型的构建。

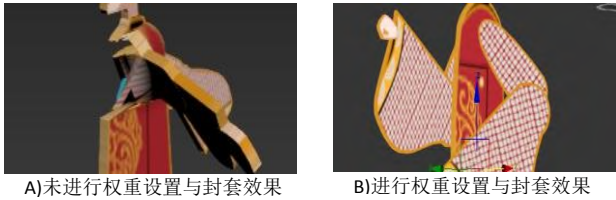


图 3 效果对比
Fig3. Effect comparison

2 基于 Kinect 的数据采集

在实现静态三维关节点皮影模型数字化的基础上,本章开始探讨皮影的动态数字化,即依托关节点皮影模型和动作采集设备实现实时人体-皮影动作匹配的动态表演数字化。首先应获取到人体数据,本文采用 Kinect 作为数据采集设备,主要原因是 Kinect 除了具有能够获取彩色图像的摄像头外,还有深度红外摄像头和红外投影机,通过区域的叠加^[7],获取深度图像^[8]。

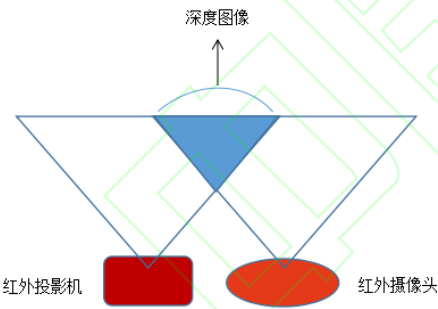


图 4 深度图像的获取
Fig4. Depth image acquisition

深度红外摄像头基于结构光技术采集深度信息,将采集到的数据转化为高中低像素 3 种分辨率,每种分辨率都包含 16 位深度数据,其格式如表 1 所示。

表 1 深度数据格式
Tab.1 Depth data format

Depth data	Bits of Depth	Bits of index
Bit	D15~D3	D2~D0

Kinect 采用 Laser Speckle 技术获取到随机的衍射斑点,当光源标定后,空间中的散斑被记录,当物体放进该空间后,获取物体表面的散斑,进而获取该场景的深度图像,获取到的光斑如图 5 所示。



图 5 人体光斑信息
Fig5. Human body spot information

3 人体骨骼数据获取

Kinect 能够获取深度图像,从某种程度上来说,可以作为动作捕捉^[9]设备捕捉人体的动作。Kinect 在获取的人体深度图像上剔除背景,通过机器学习对深度图像进行评估,以此对人体部位进行分类,计算出骨骼点坐标,最后生成骨骼。

3.1 分割遮罩

通过 Kinect 采集人体深度图像,为获取到人体骨骼,需对深度图像进行图像分割,从而确定人体的图像区域。在深度图像中剔除背景获取前景的准确性决定了动态皮影数字化的人体训练数据集的准确性。因此,在获取前景过程中应去除图像噪声,降低计算的复杂度,通过人体部位的坐标信息和深度信息进行部位区域的分割。

图像分割算法如式 1 所示。其中 z_{palm} 表示图像分割的临界值,分割点的坐标 $D(x,y)$ 的取值取决于是否位于临界值区间 $Z_{palm}-200 < Z(x,y) < Z_{palm}+150$ 之内。

$$D(x,y)=\begin{cases} 0xff, Z_{palm}-200 < Z(x,y) < Z_{palm}+150 \\ 0x00, 其他区域 \end{cases} \quad (1)$$

如图 6 所示为分割遮罩提取人体深度前景图像。

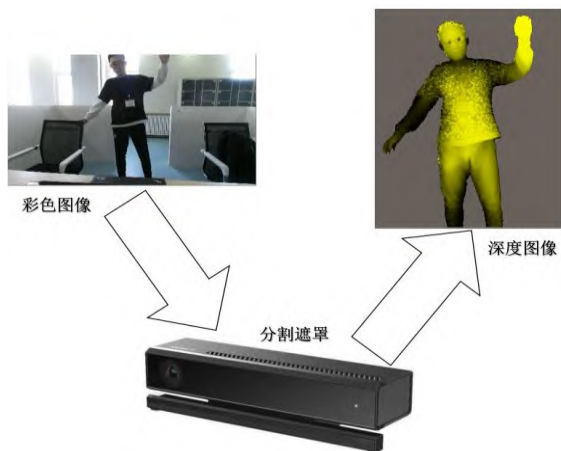


图 6 分割遮罩提取深度前景图像

Fig6. Segmentation mask to extract deep foreground image

3.2 建立学习集

将人体深度数据的每个像素通过机器学习进行特征提取与分类，以此来捕获各种人体动作，构成一个大型训练库，提高对人体部位及姿态识别的准确度。kinect 采集到的数据集如图 7 所示，对人体图像进行特征值提取，并按照人体的姿态^[10]分类计算出骨骼点坐标，从而获取人体骨骼点坐标。

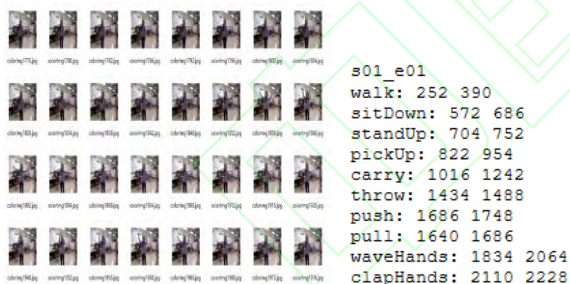


图 7 Kinect 采集到的数据集与姿态分类

Fig7. Data set collected by Kinect and pose classification

在获取精确的深度图像后，由于人体的姿态复杂^[11]，Kinect 无法直接准确地识别人体的关节点，当人体出现在镜头中时，如果未建立学习集，人体的骨骼是乱的，当建立学习集之后，可以观察到人体在坐着有遮挡时其骨骼依然能够正确获取，因此需建立一个大的数据集。建立学习集前和建立学习集后的效果对比如图 8 所示。

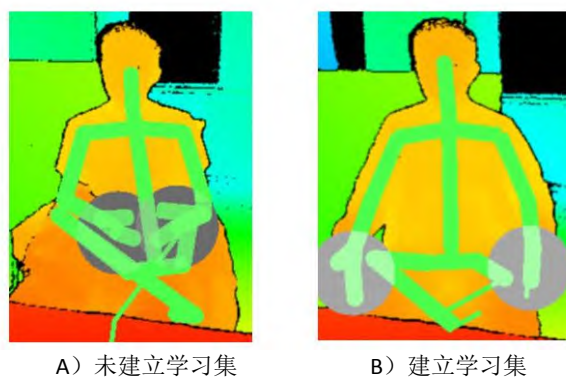


图 8 效果对比

Fig8. Effect comparison

分割遮罩提取出人体目标，然后通过机器学习对深度图像进行评估，识别人体的所有部位。Kinect 数据采集时会追踪接近人体的大小比例的人型的物体，如图 9 所示，该思路可应用于多个人体-皮影实时动作匹配中。

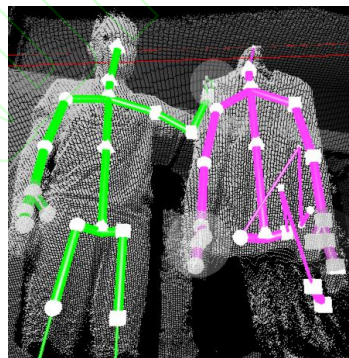


图 9 追踪人形物体

Fig9. Tracking humanoid objects

3.3 生成骨骼

人体深度数据使得各部位关节点坐标信息呈现方式为三维坐标，即多了一维深度信息。在获取了人体关节点坐标 (x_i, y_i, z_i) 后，要生成骨骼，需完成关节点坐标到现实空间坐标的转换，其转换方式如式 2 所示。

$$\begin{aligned} x_r &= (x_i - \frac{w}{2})(z_r - 10)(\frac{w}{h})p \\ y_r &= (y_i - \frac{h}{2})(z_r - 10)p \\ z_r &= 11.48 * \tan(Hz_i + 1.18) - O \end{aligned} \quad (2)$$

其中 (x_r, y_r, z_r) 为现实空间坐标， (x_i, y_i, z_i) 为深度坐标， H_{z_i} 为获取的骨骼角度， $O=3.7$ ， $p=0.004$ ，

$h=0.003$ 。基于 Kinect 跟踪的骨骼点坐标生成完整的骨骼系统,需判断是否是关节点,以及它是关节点的哪个部分。通过坐标转换,准确地评估和分析人体各部位的实际位置,从而实现实时人体数据的骨骼生成,效果如图 10 所示。



图 10 骨骼生成效果
Fig10. Bone generation effect

4 动作实时匹配

本文要实现基于 Kinect 的人体-皮影动作实时匹配,即皮影动态数字化^[12],应从传统皮影的表演形式入手。传统皮影的表演形式为幕布皮影投影,如图 11 所示。

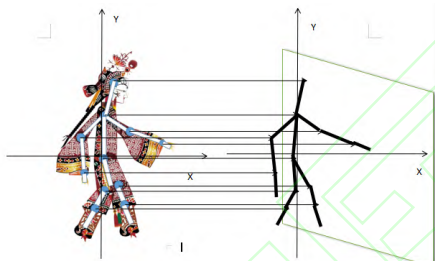


图 11 传统皮影的幕布投影
Fig11. Screen projection of traditional shadow play

借助传统皮影幕布投影的原理,将人体骨骼关节点与皮影关节点绑定,本文增加了深度信息,通过深度驱使实现转动动作。本文通过 Kinect 获取的人体骨骼点坐标与皮影骨骼绑定,遍历骨骼树,以式 3 计算当前关节点的位置 (x_i, y_i) 。

$$[x_i, y_i, l] = [x_j, y_j - L_{ij}, l] \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ f_x & f_y & 1 \end{bmatrix} \quad (3)$$

其中 L_{ij} 为当前关节点 i 和父节点 j 之间的长度, (x_j, y_j) 为父节点坐标, f_x 和 f_y 为骨骼与原来位置的偏移量, θ 为转动夹角。通过动态帧替换算法实时更新人体骨骼数据,对骨骼数据进行平滑处理。按照指数递增的形式计算每帧数据的控制因子

x_i 的大小,式 4 为 x_i 的计算方式。

$$x_i = x(1-x)^i \quad (4)$$

其中 x 为权重参数,将 $x_1 \sim x_n$ 带入式 5 计算出目前数据的平滑值 \hat{x} ,计算公式如式 5:

$$\hat{x} = \sum_{i=1}^n (\alpha_i x_i / \sum_{i=1}^n \alpha_i) \quad (5)$$

基于动态帧替换算法和数据平滑实现人体-皮影实时动作匹配^[13],关节点皮影模型能够实时跟随人体骨骼运动,效果如图 12 所示。



图 12 人体-皮影动作匹配
Fig12. Human body-shadow play action matching

在 Kinect 采集到的人体数据序列传输时,采用动态帧替换实时更新数据,利用数据平滑算法实现数据的实时处理,在数据采集、生成骨骼和人体-皮影动作匹配方面支撑动作实时性。结合非物质文化遗产皮影的特点和关节点皮影模型的构建,人体实景实时骨骼生成和匹配的实时效果呈现如图 13 所示。

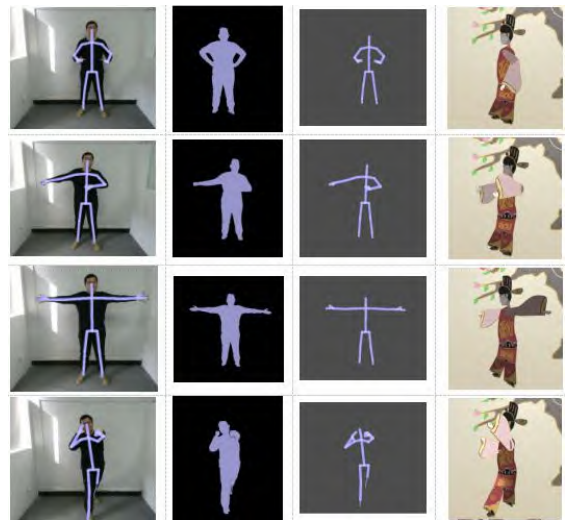


图 13 人体骨骼与皮影实时动作匹配
Fig13. Real-time action matching of human bones and shadow play

传统的皮影表演模拟主要采用的方式为 K 帧动画,即将关节点皮影模型进行动作的设定,每帧给模型摆一个动作,使用连续帧动画的形式完成皮影的动态表演。这种形式在制作过程中需要较多的时间,并需要对运动规则有充分了解,否则制作出的动画不符合人体力学运动规律,表演出来的效果死板且卡顿。如图 14 所示为基于 K 帧动画的皮影动态表演,主要呈现的是模型的每帧动画的制作[14]。



图 14 基于 K 帧动画的皮影动态表演

Fig14. Dynamic performance of shadow play based on K-frame animation

基于 K 帧动画动态表演技术,本文提出的基于 Kinect 的皮影动作实时匹配能够实现人体与皮影动作的准确匹配,相较于三维动画制作中的基于 K 帧动画[15]的动作匹配,能够实时接入人体动作,动作表演更为多样且具有灵活性,用户能够实时控制皮影模型运动。同时本文提出的基于 Kinect 的皮影动作实时匹配能够实时显示人体与关节点皮影模型的匹配表演效果,如图 15 所示,为皮影提供更为多样的表演和数字化形式。考虑皮影表演的实际场景:多个人物的互动表演形式,本文提出多人体-皮影实时匹配,能够同时追踪多位用户的深度数据和骨骼信息,实现多用户同时进行实时交互,实现效果如图 16 所示。



图 15 单用户人体-皮影实时骨骼匹配

Fig15. Single user human body-shadow play real-time bone matching



图 16 多用户人体-皮影实时骨骼匹配

Fig16. Multi-user human body-shadow play real-time bone matching

通过实验发现人体动作和关节点皮影模型动作匹配良好,未出现模型穿插现象和骨骼破损现象。能够很好的实现人体与皮影动作的实时匹配,从而完成皮影的动态表演数字化,使传统皮影焕发新生[16]。

5 结论

本文基于非物质文化遗产——皮影的传承与保护工作现状,提出皮影实体的静态数字化和皮影动态交互表演数字化两种数字化形式。但皮影动态交互表演数字化主要应用范围为人形皮影,对于其他类型的皮影支持度存在一定的问题。本文通过构建关节点皮影模型实现皮影实体的三维数字化保存,使得皮影实体的保存不再受时间和空间限制。考虑皮影表演的多实体性,本文提出多用户实时体感动作交互,基于 Kinect 实现人体深度数据的采集,对数据进行获取、存储、处理,获取到骨骼点坐标,从而生成人体骨骼。实现人体骨骼与关节点皮影模型的骨骼绑定与映射,以此完成基于 Kinect 的数据采集与绑定动作实时匹配,实现皮影的动态表演数字化。以数字多元化的形式,使得人们能够更加方便、直观、沉浸的接触到非遗皮影文化,便于中国文化的国际传播与交流,助力了国家文化软实力的发展。

参考文献:

- [1] 孟宪宁. 数字动画语境下民间皮影艺术的传承与创新[J]. 大众文艺, 2018, (6):170.
Meng Xianning. Inheritance and innovation of folk

- shadow art in the context of digital animation [J]. Popular Literature and Art, 2018, (6): 170.
- [2] 杨军燕. 对皮影戏衰落的思考[J]. 艺术教育, 2006, (5):14-15.
Yang Junyan. Thoughts on the decline of shadow play[J]. Art Education, 2006, (5): 14-15.
- [3] 李保强. 面向皮影戏动画制作的运动捕捉技术研究[D]. 华侨大学, 2016.
Li Baoqiang. Motion capture technology research for shadow play animation[D]. Huaqiao University, 2016.
- [4] 金波. 基于 3ds max 的角色动画制作研究[J]. 三峡大学学报(人文社会科学版), 2017, 39(S2):162-163.
Jin Bo. Research on character animation production based on 3ds max [J]. Journal of China Three Gorges University (Humanities and Social Sciences Edition), 2017, 39(S2): 162-163.
- [5] 单超杰. 皮影人物造型与三维建模技术结合的创新研究[D]. 东华大学, 2013.
Shan Chaojie. Innovative research on the combination of shadow figure modeling and 3D modeling technology [D]. Donghua University, 2013.
- [6] Vogel D , Lubos P , Steinicke F . AnimationVR - Interactive Controller-Based Animating in Virtual Reality[C]// 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, 2018.
- [7] Zhipeng Chen. Kinect-based Behavior Measurement in Group Discussion[C]. SCIENCE and Engineering Institute (SCIEI), Wuhan University, Proceedings of 2019 the World Symposium on Software Engineering (WSSE 2019). 2019:126-130.
- [8] 刘俊毅. 彩色图像引导的深度图像增强[D]. 浙江大学, 2014.
Liu Junyi. Depth map enhancement Under the guidance of color image [D]. Zhejiang University, 2014.
- [9] 邱望标, 李超. 基于运动捕捉技术的中国少数民族舞蹈艺术保护方法研究[J]. 电子科技大学学报(社科版), 2009, 11(4):101-104.
Qiu Wangbiao, Li Chao. Research of Protection Methods about the Dances Artistic of National Minorities Based on Motion Capture Technology[J]. Journal of University of Electronic Science and Technology of China (Social Science Edition), 2009, 11(4): 101-104.
- [10] 王亮, 胡卫明, 谭铁牛. 人运动的视觉分析综述[J]. 计算机学报, 2002(3):225-237.
Wang Liang, Hu Weiming, Tan Tieniu. A survey of visual analysis of human motion[J]. Chinese Journal of Computers, 2002(3):225-237.
- [11] 朱洪堃, 殷佳炜, 冯文字, 华亮, 费敏锐, 张堃. 一种轻量化实时人体姿势检测模型研究与应用[J/OL]. 系统仿真学报:1-14[2020-10-11]. <http://kns.cnki.net/kcms/detail/11.3092.V.20200727.1101.002.html>.
Zhu Hongkun, Yin Jiawei, Feng Wenyu, Hua Liang, Fei Minrui, Zhang Kun. Research and application of a lightweight real-time human posture detection model [J/OL]. Journal of System Simulation: 1-14 [2020-10-11] .<http://kns.cnki.net/kcms/detail/11.3092.V.20200727.1101.002.html>.
- [12] 陈璇, 张明敏, 潘志庚. 数字皮影的交互仿真和动画绘制[J]. 中国图象图形学报, 2014, 19(10):1490-1499.
Chen Xuan, Zhang Mingmin, Pan Zhigeng. Interaction and animation simulation of digital shadow play[J]. Journal of Image and Graphics, 2014, 19(10): 1490-1499.
- [13] 赵瑛, 王冬晖, 李琦, 于爱萍, 谷宇. 仿真假体视觉下基于深度图像的手势识别研究[J]. 现代电子技术, 2019, 42(16):131-135+139.
Zhao Ying, Wang Donghui, Li Qi, Yu Aiping, Gu Yu. Research on gesture recognition based on depth image under simulated prosthetic vision[J]. Modern Electronic Technology, 2019, 42(16): 131-135+139.
- [14] Pan Y , Mitchell K . PoseMMR: A Collaborative Mixed Reality Authoring Tool for Character Animation[C]// 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). IEEE, 2020.
- [15] 朱大诚. 皮影风格动画研究[D]. 山东大学, 2014.
Zhu Dacheng. Research of shadow graph-style animation [D]. Shandong University, 2014.
- [16] 冀南. 涉足三维领域 创新民族动画 古老皮影幻化新生[J]. 国家人文历史, 2018(2):12-13.
Jinan. Involved in the three-dimensional field, innovating national animation, ancient shadow puppets turned into a new life[J]. National Humanities History, 2018(2): 12-13.