

杨惠

电话: 15182982150 | 邮箱: 2410383183@qq.com

教育经历

上海科技大学【芒果TV联培，项目型硕士】

2020年09月 - 2023年06月

计算机科学与技术 硕士 信息科学与技术学院

上海

成都信息工程大学

2016年09月 - 2020年06月

软件工程 本科 软件工程学院

成都

荣誉: 二等奖学金(2016-2017), 一等奖学金(2017-2018), 一等奖学金(2018-2019)

虚拟人相关经历

中国科学院香港创新研究院

2023年08月 - 至今

科研助理

• Music2Dance

收集整合Music2Dance数据集(Phantom dance, AIST++, GDdance, DanceWithMelody), 基于开源算法实现音乐生成舞蹈动作。

• 动作重建【论文HardMo++】

为了解决目前单张RGB人体重建技术在手腕、脚踝、侧身估计存在的瓶颈, 我们制作了一个包含100万张图像的, 带有精确标注的HMR数据集。数据集包含两个子集: FreeMan数据集和Hardmo数据集。通过ScoreHMR的multiview guidance对Freeman的多视角图像处理, 校准侧身标注; 通过ScoreHMR的keypoints2d guidance校准脚踝标注; 通过HaMeR和PyMafx生成精确的手势, 并校准手腕标注。

• Text2motion和去脚滑【论文StableMoFusion】

基于diffusion模型实现了文本生成动作。为了推进text2motion在工业界的应用, 本文高效地实现了模型加速和动作去脚滑功能。其中模型加速采用了如下策略: 1)使用高效的采样器(SDE DPM-Solver++2M)减少降噪迭代次数 2)缓存文本特征, 减少冗余计算 3)在推理过程实现并行条件引导机制 4)采用半精度浮点计算。去脚滑模块使用了力学模型检测脚步受力, 从而判断脚触地情况, 据此约束脚步运动, 消除脚滑。

• 人与物体交互【论文OOD-HOI】

为了解决文本指导的人与物体交互方法在域外(Out-of-Domain, 简称OOD)生成的限制, 本文对两个相关域外生成进行了探索, 分别是Text-OOD和Object-OOD。1)通过设计新颖的网络结构, 对物体外观形变来进行数据增强等方法, 解决了Object-OOD, 实现了对未知物体的交互。2)通过大语言模型实现了文本的数据扩充, 解决了Text-OOD, 实现了在未知文本下的动作生成。3)还设计了优化策略, 解决生成手势与物体间存在的“漂浮”, “穿模”等情况, 以推进该技术在工业级的应用。

• 视频人物替换【专利CN202411064734.6】

结合关键点检测算法、图像分割算法、视频编辑算法消除视频中的人物, 得到视频背景; 使用3D姿态估计算法检测出视频中人物的动作; 使用SMPLX, Mixamo, Auto-Rig Pro等Blender插件实现将动作重定位到Mixamo虚拟角色上, 并对动作进行渲染; 将渲染后的3D虚拟角色动作和背景图像组合得到最终的替换结果。

• 手术室模拟系统【论文DigitSurge】

为了助力医学教育和手术规划, 本文提出了一种虚拟手术室系统, 用于实现手术模拟。该系统包含3个主要模块: 手术环境仿真、任务规划、动作生成。1)我们在Unreal Engine中搭建了虚拟手术室环境, 包含手术床, 手术灯, 托盘桌等不可移动的元素, 和可驱动的数字人。2)通过LLM对用户输入的指令实现任务分解。3)通过动作生成模块分别生成动作轨迹和原地的医学动作。为了消除domain gap, 我们构建了一个医学动作数据集, 通过爬取视频网站上外科职业技能培训视频, 整合相关HMR方法和VideoLLama2实现smplx动作序列估计及文本描述生成, 得到医学场景下特有的动作文本数据集。

上科大联培

2021年08月 - 2023年07月

• 基于语音素材的骨骼动作数据生成【硕士毕业设计】

目标是生成语音协同的骨骼动作数据, 以解决前两种驱动方式对动作演员有要求的问题, 从而提高虚拟人的可交互性。模型采用了Seq2Seq的结构, 实验尝试了使用Transformer、RNN、LSTM、CNN、TCN等网络作为基本结构。通过引入动作质量判别器和语音判别器分别提高了生成动作序列的质量和语义相关性。

出版物

- 《StableMoFusion: Towards Robust and Efficient Diffusion-based Motion Generation Framework》【第二作

者, ACM MM Oral, 接受率<3.97%】

- 《HardMo++: A Large-Scale Hardcase Dataset for Motion Capture》【已完稿, 待投稿】
- 《OOD-HOI: Text-Driven 3D Whole-Body Human-Object Interaction Generation Beyond Training Domains》【第二作者, ICCV2025在投】
- 《DigitSurge: A Digital Operating Room for Surgery Simulation》(第一作者, MICCAI2025在投)
- 《一种虚拟人驱动方法,装置,设备及可读存储介质》【专利号: 202310098261.0】
- 《一种将视频中的人物替换为3D虚拟角色的方法》【专利号: CN202411064734.6】
- 《个人博客系统》【软件著作权, 登记号: 2020SR1250753】

技能/证书及其他

- 技能: Matlab, Pytorch, C++, Unreal Engine, Blender
- Github: <https://github.com/svitass>
- 证书: 英语 (CET-6)

作品总结

