

朱昊

✉ haozhu96@gmail.com

☎ +86-13024050123 (同微信)

🏠 1996

🌐 <https://www.zhuhaozh.xyz> (google scholar)

📅 2.5 年全职 + 2 年实习



就业经历 Employment History

- 2023.8 – 研究工程师, 三维数字内容生成 (数字人) - 平台设施研究部 - 上海人工智能实验室
最多带领 2 人, 从事于 3D 数字人生成与驱动相关工作 (商汤同组转签)
- 2021.7 – 2023.8 算法研究员 (L6), 扩展现实实验室 - 研究院 - 商汤科技
最多带领 4 人, 从事于 AIGC 与 数字人 领域工作
- 2019.7 – 2020.12 见习研究员, 智能视频 - MIG - 商汤科技
共研发/支持 3 个有关 *Human-centric Generation* 与 数字人 方向项目

教育背景 Education

- 2018.9 – 2021.7 联培硕士 - 中国科学院, 自动化研究所
模式识别国家重点实验室; 导师: 赫然; 研究方向: 音视频生成与理解
- 学术型硕士 - 安徽大学, 计算机科学与技术
国家奖学金 / 省优秀硕士毕业生 / 省双优生 / 省优秀硕士论文
- 2014.9 – 2018.7 本科 - 安徽工程大学, 计算机科学与技术

工作经验 Work Experience

- 2023 场景中的数字人 (Avatar in the Scene) (2023.8~)
职责 & 项目描述: 主要研发者, 与 2 名同事共同完成。为实现行人在模拟器中的真实场景, 实现规模性人与场景交互数据的生产。项目要求高质量的渲染效果、行人在外观与行为上的多样性、运动与交互的自然性。该项目包含: 三维人体生成、动作迁移、轨迹预测等。
阶段成果: 1) 数据接口: 在相关模拟器中实现行人的感知功能、人体数据接口整合、运动重定向等; 2) 轨迹预测: 实现初步人体轨迹预测功能;
- 项目预研 - 多模态内容生成方向 (2023.2~2023.12)
职责 & 项目描述: 主要研发者, 带领 1 名同事共同完成。该项目计划实现数字人相关内容多模态内容生成。通过相关数据的生成, 加速美工人物设计、训练数据生成等工作。具体方向包含 text-to-3D 人物生成, 2D 人体模特图生成。
项目成果: 1) 3D 人体纹理生成: 实现 SOTA 人体纹理生成效果; 相关工作已在后续项目中作为资源支持; 2) 2D 人体模特图生成: 设计新的模特图生成方案, 在保证衣物的高保真的同时, 实现人台图与真人图的模特生成。
- 2022 数字人项目 (2022.3~2023.3)
职责 & 项目描述: 主要研发者。数字人项目需支持 2D 与超写实 3D 方向。目标为根据输入的音频/文本等信息, 驱动数字人的运动。运动的内容包括嘴唇运动, 以及如眨眼等。项目所需技术: 语音处理、口型与表情的预测; 3D 人脸等相关技术。
项目成果: 1) 模型效果: 进一步提高了嘴唇运动的自然度和准确性; 2) 泛化支持: 增加多种驱动模型的支持; 3) 项目落地: 基于该项目开发的一体机已经服务于银行等客户。

工作经验 Work Experience (continued)

项目预研 - 多模态 AIGC 数据生产方向 (2022.1~2022.12)

职责 & 项目描述: 主要研发者, 带领 3 位同事共同完成。1) 该项目计划实现数字人相关伪数据的生产, 实现相关算法的伪造数据生成, 以更快地支持下一代应用方向的研究; 2) 该项目包含 3 个数据集搭建以及 1 个新的模型设计。**项目所需技术:** Video Diffusion、Stable Diffusion、LoRA、场景图生成以及文本生成等技术。

项目成果: 1) CelebV 系列数据集: 首个人工标注的、以及含文本标注的面部视频数据集, 共计 500+ GitHub star。2) 百万级高质量 text-to-image 数据集: 半自动化的标注方案, 包括物体关系在内的细粒度描述; 3) 新的 Video Diffusion 模型: 适用于多模态驱动和长序列视频的生成; 4) 项目支持: 该项目已实现内部相关项目的数据支持工作, 加速了相关算法迭代。

2021

活照片项目 - SenseMARS Platform. (2021.9~2021.12)

职责 & 项目描述: 主要研发者, 带领 1 名同事完成。该项目为面部动作驱动任务, 实现实时返回生成视频与实时预览。**项目所需技术:** 面部驱动、视频生成、GAN 网络、人脸超分等技术。

项目成果: 1) 效果优化: 完成 6 种运动模板的支持, 在保证鲁棒性的条件下, 生成质量满足产品侧需求; 2) 模型加速: 在较低性能的服务器上实现实时输出; 3) 功能支持: 在原有的基础上实现预览功能; 4) 项目落地: 计划上线自研 APP, 多平台用户数达 10w+。

妆发迁移项目 - SenseMARS Platform. (2021.7~2021.12)

职责 & 项目描述: 主要研发者之一。与 2 位同事共同研发, 并带领 1 名实习生。该产品实现妆容、发型等快速迁移。给定参照图片, 实现妆容与发型的风格迁移。**项目所需技术:** 风格迁移、人脸生成、GAN 网络、人脸超分等技术。

项目成果: 1) 数据集设计与搭建: 完成项目所需数据的设计并发布采集与标注任务, 实现数据资产的建设; 2) 效果优化: 完成发型发色妆容等迁移要求, 还原度达到 90%; 实现 30+ 个模板的支持工作; 3) 项目落地: 该项目服务于某公司, 部分风格模板上线自研 APP。

项目预研 - SenseMARS Platform (2021.7~2021.12)

职责 & 项目描述: 主要研发者, 带领 2 位同事完成。该项目旨在更好且更平滑地支持后续项目开展, 在确定后续需求中可能存在的难点, 对相关方向进行预研。方向包括自由度更高的面部重现 (face reenactment) 和年龄范围更广的年龄编辑。

项目成果: 1) 3D 先验的人脸重演模型: 支持细节运动如: 挑眉、眨眼、眼球运动, 以及高自由度的头部运动等动作; 2) 终生年龄编辑模型: 支持 0 ~ 120 岁的编辑范围的终生年龄编辑, 以及一个带有年龄和分割标注的合成脸部图像数据集。

实习期项目经历 (2019-2020): 视频换脸项目 & 2D 数字人项目 & 人体运动迁移项目

部分论文发表 Selected Publications

CVPR/NeurIPS/ECCV/AAAI/IJCAI/...

- 1 J. Yu, H. Zhu, L. Jiang, C. C. Loy, W. Cai, and W. Wu, "PaintHuman: Towards high-fidelity Text-to-3D human texturing via denoised score distillation," in *AAAI*, 2024.
- 2 J. Yu, H. Zhu, L. Jiang, *et al.*, "CelebV-Text: A large-scale facial text-video dataset," in *CVPR*, 2023.
- 3 H. Zhu, W. Wu, W. Zhu, *et al.*, "CelebV-HQ: A large-scale video facial attributes dataset," in *ECCV*, 2022.
- 4 H. Zhu, C. Fu, Q. Wu, W. Wu, C. Qian, and R. He, "AOT: Appearance optimal transport based identity swapping for forgery detection," in *NeurIPS*, 2020.
- 5 H. Zhu, H. Huang, Y. Li, A. Zheng, and R. He, "Arbitrary talking face generation via attentional audio-visual coherence learning," in *IJCAI*, 2020.