

郑泽荣

+86 17888842659 • zrzhen1995@foxmail.com
<https://zhengzerong.github.io>

教育经历

清华大学 自动化系
博士

中国 北京
2018/08 - 2023/06

- 导师: 刘焯斌教授
- 实验室: 清华大学成像与智能实验室
- 荣誉: 国家奖学金、未来学者奖学金、清华之友合肥英才一等奖学金等

清华大学 自动化系
学士

中国 北京
2014/08 - 2018/07

- GPA: 90/100, 排名 10%
- 实验室: 清华大学成像与智能实验室
- 荣誉: 学业优秀奖学金、科创优秀奖学金、优秀毕业论文等

工作经历

字节跳动 - Seed
算法科学家

中国 深圳
2025/12 至今

- 部门: Seed-Application-Vision
- 职责: 主要负责 Seedance 2.0→2.5→... 系列模型的多模态参考能力的研发, 包括人像参考、音色参考、场景参考等子能力。作为能力项纵向 POC, 负责数据准备、预训练、后训练、推理优化等全链条环节。

字节跳动 - 智能创作
算法科学家

中国 深圳
2024/07 - 2025/12

- 部门: Data-智能创作-数字人 & AR
- 职责: 主要负责多模态数字人大模型 (OmniHuman 系列) 的从 0 到 1 的建设和研发, 是 OmniHuman 1.0→1.1→1.5→2.0 的核心算法贡献者。

杭州新畅元科技有限公司
首席算法科学家

中国 杭州
2023/07 - 2024/07

- 公司简介: 杭州新畅元科技是清华大学成果转化公司, 由清华大学刘焯斌教授联合其他创始人创办, 主营业务为数字人相关业务, 目标是将学术研究成果进行落地转化, 用数字人技术赋能电商、教育、文旅等行业。
- 部门: 技术中心——算法
- 职责: 领导团队成员, 主导公司在 2D/3D 数字人方向的技术研发和落地。

实习经历

Facebook
研究实习生

美国 旧金山
2019/06 - 2019/09

- 部门: Facebook Reality Lab at Sausalito
- 主管: Dr. Tony Tung

南加州大学
本科访问学者

美国 洛杉矶
2017/06 - 2017/08

- 学院: Vision and Graphics Lab, USC Institute for Creative Technologies
- 导师: Prof. Hao Li

项目经历

长期聚焦于多模态视频生成、数字人、三维重建和渲染的相关研究。截至到目前, 已在各大顶会顶刊发表论文 50+ 篇, Google Scholar 引用量 5000+。近 3 年负责的项目如下:

多模态视频生成模型:

2025 - 至今

代表性成果:

- Seedance 2.0: 全球领先的多模态视频生成大模型, 同时支持文生视频、图生视频、以及多种参考、编辑、延长任务, 镜头调度灵活性和画面表现力显著优于同期模型, 在 2026 年上半年保持领先状态。
- Seedance 2.5: 2.0 模型的进一步升级, 通过数据、模型的进一步 scale-up, 提升模型的画质、美感、生成时长和复杂指令响应能力, 更好满足专业创作需求。

多模态数字人大模型:

2024 - 2025

代表性成果:

- 第一代数字人大模型 OmniHuman 1.0, 首次将 DiT 视频模型引入数字人, 表现力、灵活性、泛化性显著领先同期模型, 获 Forbes、CNBC 等国内外媒体的关注和报道, 发表ICCV 2025论文一篇 →  Demo
- 内部更新版本 OmniHuman 1.1, 相比 1.0 版本, GSB 提升 23%, 口型可用率 MOS 打分 3.44 vs 3.07。
- 第二代数字人大模型 OmniHuman 1.5, , 增强数字人大模型对语音输入的语义适配和情感动作表现力。相比 1.1 版本, GSB 提升 28.9%, 非真人形象场景下 GSB 提升 63.6%, 发表ICLR 2026论文一篇% →  Demo
- 对标 Sora Cameo 功能的 OmniHuman 2.0, 实现带形象参考和音色参考的人物演绎音视频联合生成, 春节前上线即梦 APP。

数字人三维重建和渲染:

2023 - 2024

代表性成果:

- Animatable Gaussians: 基于投影高斯图的可驱动数字人表示和学习框架, 实现高清人体 avatar 建模、驱动和渲染, 发表CVPR 2024论文一篇 →  Demo
- Gaussian Head Avatar: 将高斯 Splatting 引入到人头化身学习任务中, 实现高保真度、高灵活性的人头 avatar 建模, 发表CVPR 2024论文一篇 →  Demo
- AvatarRex: 全球首个单视角下的真人 NeRF 全身数字人实时驱动系统力, 发表SIGGRAPH 2023 Journal Track论文一篇, 并收录到ACM Trans. Graph. →  Demo

代表论文 (完整论文列表见Google Scholar)

2026:

Jianwen Jiang*, Weihong Zeng*, **Zerong Zheng***, Jiaqi Yang*, Chao Liang*, Wang Liao*, Han Liang*, Yuan Zhang, Mingyuan Gao.

OmniHuman-1.5: Instilling an Active Mind in Avatars via Cognitive Simulation.

International Conference on Learning Representations (ICLR) 2026 (oral).

Lizhen Wang, Yongming Zhu, Zhipeng Ge, Youwei Zheng, Longhao Zhang, Tianshu Hu, Shiyang Qin, Mingshuang Luo, Jiaxu Zhang, Xin Chen, Yulong Wang, **Zerong Zheng**, Jianwen Jiang, Chao Liang, Weifeng Chen, Xing Wang, Yuan Zhang, Mingyuan Gao.

FlowAct-R1: Towards Interactive Humanoid Video Generation.

arXiv.

2025:

Tongchun Zuo, Zaiyu Huang, Shuliang Ning, Ente Lin, Chao Liang, **Zerong Zheng**, Jianwen Jiang, Yuan Zhang, Mingyuan Gao, Xin Dong.

DreamVVT: Mastering Realistic Video Virtual Try-On in the Wild via a Stage-Wise Diffusion Transformer Framework.

arXiv.

Chao Liang, Jianwen Jiang, Wang Liao, Jiaqi Yang, **Zerong zheng**, Weihong Zeng, Han Liang.

AlignHuman: Improving Motion and Fidelity via Timestep-Segment Preference Optimization for Audio-Driven Human Animation.

arXiv.

Lizhen Wang, Zhurong Xia, Tianshu Hu, Pengrui Wang, Pengfei Wei, **Zerong Zheng**, Ming Zhou, Yuan Zhang, Mingyuan Gao.

DreamActor-H1: High-Fidelity Human-Product Demonstration Video Generation via Motion-designed Diffusion Transformers.
arXiv.

Zhenzhi Wang, Jiaqi Yang, Jianwen Jiang, Chao Liang, Gaojie Lin, **Zerong Zheng**, Ceyuan Yang, Dahua Lin.
InterActHuman: Multi-Concept Human Animation with Layout-Aligned Audio Conditions.
arXiv.

Gaojie Lin*, Jianwen Jiang*, Jiaqi Yang*, **Zerong Zheng***, Chao Liang.
OmniHuman-1: Rethinking the Scaling-Up of One-Stage Conditioned Human Animation Models.
IEEE International Conference on Computer Vision (ICCV) 2025 (**Spotlight**).

2024:

Yushuo Chen, **Zerong Zheng**, Zhe Li, Chao Xu, Yebin Liu.
MeshAvatar: Learning High-quality Triangular Human Avatars from Multi-view Videos.
European Conference on Computer Vision (ECCV) 2024.

Ruizhi Shao, Youxin Pang, **Zerong Zheng**, Jingxiang Sun, Yebin Liu.
Human4DiT: 360-degree Human Video Generation with 4D Diffusion Transformer.
ACM SIGGRAPH Asia 2024.

Siyu Lin, Zhe Li, Zhaoqi Su, **Zerong Zheng**, Hongwen Zhang, Yebin Liu.
LayGA: Layered Gaussian Avatars for Animatable Clothing Transfer.
ACM SIGGRAPH 2024.

Zhe Li, **Zerong Zheng**, Lizhen Wang, Yebin Liu.
Animatable Gaussians: Learning Pose-dependent Gaussian Maps for High-fidelity Human Avatar Modeling.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2024.

Yuelang Xu, Benwang Chen, Zhe Li, Hongwen Zhang, Lizhen Wang, **Zerong Zheng**, Yebin Liu.
Gaussian Head Avatar: Ultra High-fidelity Head Avatar via Dynamic Gaussians.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2024.

Chenyang Wang, **Zerong Zheng**, Tao Yu, Xiaoqian Lv, Bineng Zhong, Shengping Zhang, Liqiang Nie.
DiffPerformer: Iterative Learning of Consistent Latent Guidance for Diffusion-based Human Video Generation.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2024.

Xiang Deng, **Zerong Zheng**, Yuxiang Zhang, Jingxiang Sun, Chao Xu, XiaoDong Yang, Lizhen Wang, Yebin Liu..
RAM-Avatar: Real-time Photo-Realistic Avatar from Monocular Videos with Full-body Control.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2024.

Ruizhi Shao, Jingxiang Sun, Cheng Peng, **Zerong Zheng**, Boyao Zhou, Hongwen Zhang, Yebin Liu.
Control4D: Efficient 4D Portrait Editing with Text.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2024.

Yuemei Zhou, Tao Yu, **Zerong Zheng**, Gaochang Wu, Guihua Zhao, Wenbo Jiang, Ying Fu, Yebin Liu.
ProbIBR: Fast Image-Based Rendering with Learned Probability-Guided Sampling.
IEEE Transactions on Visualization and Computer Graphics (TVCG) 2024.

2023:

Zerong Zheng, Xiaochen Zhao, Hongwen Zhang, Boning Liu, Yebin Liu.
AvatarReX: Real-time Expressive Full-body Avatars.
ACM Transactions on Graphics (Proceedings of SIGGRAPH) 2023.

Zhe Li, **Zerong Zheng**, Yuxiao Liu, Boyao Zhou, Yebin Liu.
PoseVocab: Learning Joint-structured Pose Embeddings for Human Avatar Modeling.
ACM SIGGRAPH 2023.

Siyu Lin, Boyao Zhou, **Zerong Zheng**, Hongwen Zhang, Yebin Liu.
Leveraging Intrinsic Properties for Non-Rigid Garment Alignment.
IEEE International Conference on Computer Vision (ICCV) 2023.

Ruizhi Shao, **Zerong Zheng**, Hanzhang Tu, Boning Liu, Hongwen Zhang, Yebin Liu.
Tensor4D : Efficient Neural 4D Decomposition for High-fidelity Dynamic Reconstruction and Rendering.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2023.

Hongwen Zhang, Siyou Lin, Ruizhi Shao, Yuxiang Zhang, **Zerong Zheng**, Han Huang, Yandong Guo, Yebin Liu.
CloSET: Modeling Clothed Humans on Continuous Surface with Explicit Template Decomposition.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2023.

2022:

Zerong Zheng, Han Huang, Tao Yu, Hongwen Zhang, Yandong Guo, Yebin Liu.
Structured Local Radiance Fields for Human Avatar Modeling.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2022.

Hao Zhao, Jinsong Zhang, Yu-Kun Lai, **Zerong Zheng**, Yingdi Xie, Yebin Liu, Kun Li.
High-Fidelity Human Avatars from a Single RGB Camera.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2022.

Ruizhi Shao, **Zerong Zheng**, Hongwen Zhang, Jingxiang Sun, Yebin Liu.
High Quality Human Reconstruction via Diffusion-based Stereo Using Sparse Cameras.
European Conference on Computer Vision (ECCV) 2022 (**oral**).

Zhe Li, **Zerong Zheng**, Hongwen Zhang, Chaonan Ji, Yebin Liu.
AvatarCap: Animatable Avatar Conditioned Monocular Human Volumetric Capture.
European Conference on Computer Vision (ECCV) 2022.

Siyu Lin, Hongwen Zhang, **Zerong Zheng**, Ruizhi Shao, Yebin Liu.
Learning Implicit Templates for Point-Based Clothed Human Modeling.
European Conference on Computer Vision (ECCV) 2022.

Ruizhi Shao, Liliang Chen, **Zerong Zheng**, Hongwen Zhang, Yuxiang Zhang, Han Huang, Yandong Guo, Yebin Liu.
FloRen: Real-time High-quality Human Performance Rendering via Appearance Flow Using Sparse RGB Cameras.
ACM SIGGRAPH Asia 2022.

2021:

Zerong Zheng, Tao Yu, Qionghai Dai, Yebin Liu.
Deep Implicit Templates for 3D Shape Representation.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2021 (**oral**).

Tao Yu, **Zerong Zheng**, Kaiwen Guo, Pengpeng Liu, Qionghai Dai, Yebin Liu.
Function4D: Real-time Human Volumetric Capture from Very Sparse Consumer RGBD Sensors.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2021 (**oral**).

Zhe Li, Tao Yu, **Zerong Zheng**, Kaiwen Guo, Yebin Liu.
POSEFusion: Pose-guided Selective Fusion for Single-view Human Volumetric Capture.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2021 (**oral**).

Yang Zheng*, Ruizhi Shao*, Yuxiang Zhang, Tao Yu, **Zerong Zheng**, Qionghai Dai, Yebin Liu.
DeepMultiCap: Performance Capture of Multiple Characters Using Sparse Multiview Cameras.
IEEE International Conference on Computer Vision (ICCV) 2021.

2020:

Zerong Zheng, Tao Yu, Yebin Liu, Qionghai Dai.
PaMIR: Parametric Model-Conditioned Implicit Representation for Image-based Human Reconstruction.
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) .

Zhuo Su, Lan Xu, **Zerong Zheng**, Tao Yu, Yebin Liu, Lu Fang.
RobustFusion: Human Volumetric Capture with Data-driven Visual Cues using a RGBD Camera.
European Conference on Computer Vision (ECCV) 2020 (spotlight).

Zhe Li, Tao Yu, **Zerong Zheng**, Yebin Liu.
Robust 3D Self-portraits in Seconds.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2020 (**oral**).

2019:

Zerong Zheng, Tao Yu, Yixuan Wei, Qionghai Dai, Yebin Liu.
DeepHuman: 3D Human Reconstruction from a Single Image.
IEEE/CVF International Conference on Computer Vision (ICCV) 2019 (**oral**).

Tao Yu, **Zerong Zheng**, Yuan Zhong, Jianhui Zhao, Qionghai Dai, Gerard Pons-Moll, Yebin Liu.
SimulCap : Single-View Human Performance Capture with Cloth Simulation.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2019.

2018:

Zerong Zheng, Tao Yu, Hao Li, Kaiwen Guo, Qionghai Dai, Lu Fang, Yebin Liu.
HybridFusion: Real-Time Performance Capture Using a Single Depth Sensor and Sparse IMUs.
European Conference on Computer Vision (ECCV) 2018.

Tao Yu, **Zerong Zheng**, Kaiwen Guo, Jianhui Zhao, Qionghai Dai, Hao Li, Gerard Pons-Moll, Yebin Liu.
DoubleFusion: Real-time Capture of Human Performances with Inner Body Shapes from a Single Depth Sensor.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2018 (**oral**).

学术服务

独立审稿人:

CVPR 2019~, ICCV 2019~, ECCV 2020~, NeurIPS 2020~, ICML 2021~, SIGGRAPH 2022~, EuroGraphics 2023~, TVCG 2023~, TIP 2023~