

张舰

✉ zjrandomyeah@gmail.com · 🌐 Jian-Zhang-3DV · 📄 Google Scholar · 🏠 Homepage

🎓 教育背景

厦门大学, 厦门
硕士研究生

2023 年 9 月 – 至今

南昌大学, 南昌
学士学位, 人工智能专业

2019 年 9 月 – 2023 年 6 月

📄 学术成果

VLM-3R: Vision-Language Models Augmented with Instruction-Aligned 3D Reconstruction (ArXiv) 2025

Authors: **Jian Zhang**^{*}, Zhiwen Fan^{*}, Renjie Li, Junge Zhang, Runjin Chen, Hezhen Hu, Kevin Wang, Huaizhi Qu, Dilin Wang, Zhicheng Yan, Hongyu Xu, Justin Theiss, Tianlong Chen, Jiachen Li, Zhengzhong Tu, Zhangyang Wang, Rakesh Ranjan (*Equal Contribution)

VLM-3R: 提出一个通过 3D 重建指令微调增强视觉-语言模型 (VLM) 的统一框架, 可直接从单目视频处理 3D 信息。结合几何、相机视图 token 及超 20 万 3D 重建指令问答对, 显著提升 VLM 空间理解和时空推理能力。

📄 Paper · </> Code (★ 133, 开源于 2025-05-26) · 🌐 Project Page

Large Spatial Model: End-to-end Unposed Images to Semantic 3D (NeurIPS) 2024

Authors: **Jian Zhang**^{*}, Zhiwen Fan^{*}, Wenyan Cong, Peihao Wang, Renjie Li, Kairun Wen, Shijie Zhou, Achuta Kadambi, Zhangyang Wang, Danfei Xu, Boris Ivanovic, Marco Pavone, Yue Wang (*Equal Contribution)

LSM: 提出一个端到端框架, 从无姿态 RGB 图像实时重建含几何、外观和语义的 3D 辐射场。通过 Transformer、跨视角注意力和多模态融合, 有效整合 2D 图像特征与语义至 3D 空间, 首次实现实时语义 3D 重建, 在新视角合成和开放词汇 3D 分割上表现优越。

📄 Paper · </> Code (★ 186, 开源于 2024-12-22, Citations: 13) · 🌐 Project Page

InstantSplat: Sparse-view Gaussian Splatting in Seconds (ArXiv) 2024

Authors: Zhiwen Fan^{*}, Kairun Wen^{*}, Wenyan Cong^{*}, Kevin Wang, **Jian Zhang**, Xinghao Ding, Danfei Xu, Boris Ivanovic, Marco Pavone, Georgios Pavlakos, Zhangyang Wang, Yue Wang (*Equal Contribution)

InstantSplat: 提出一种基于自监督学习和神经渲染的快速 3D 场景重建方法。通过高效优化 3D 场景表示 (高斯溅射) 和相机位姿, 从稀疏视角 2D 图像快速生成高质量 3D 模型。

📄 Paper · </> Code (★ 1.4k, 开源于 2024-08-01, Citations: 92) · 🌐 Project Page

DynamicVerse: Physically-Aware Multimodal Modeling for Dynamic 4D Worlds (Preprint) 2025

Authors: Kairun Wen^{*}, Yuzhi Huang^{*}, Runyu Chen, Hui Zheng, Yunlong Lin, Panwang Pan, Chenxin Li, Wenyan Cong, **Jian Zhang**, Junbin Lu, Chenguo Lin, Dilin Wang, Zhicheng Yan, Hongyu Xu, Justin Theiss, Yue Huang, Xinghao Ding, Rakesh Ranjan, Zhiwen Fan (*Equal Contribution)

DynamicVerse: 提出一个物理感知的多模态 4D 建模框架, 理解动态世界的 3D 结构、物理运动及文本语义。通过大模型处理信息, 将长视频转为 4D 多模态格式, 并提供超 10 万视频和 80 万标注掩码的大规模数据集, 赋能具身智能体。

🌐 Project Page (论文与代码即将发布)

🏆 获奖经历

南昌大学特等奖学金
南昌大学优秀毕业生

2023 年
2023 年