



□姓名: 杜壹华 □联系电话: +86 12345678910

□个人主页: <https://hit-perfect.github.io/> □电子邮件: ydu967@connect.hkust-gz.edu.cn/duyihua0130@gmail.com

教育经历

| | | | |
|--|-----------|-------|------------------|
| ➤ 哈尔滨工业大学 | 计算机科学与技术 | 本科 | 2021/09--2025/07 |
| 学习成绩: 学分绩:90.58/100; GPA:3.787/4; Rank:16/117(13.7%); WES:3.81/4; IELTS:7; | | | |
| 在校荣誉: 2022、2023 年度校优秀学生; 2023 校优秀团干部; 人民奖学金; | | | |
| ➤ 香港科技大学(广州) | 信息枢纽-人工智能 | 全奖研究生 | 2025/08--至今 |

实习经历

| | | | |
|--------------|--------------------------|-------|------------------|
| ➤ 腾讯(光子工作室) | 光子技术发展部 | 算法研究员 | 2026/01—至今 |
| ➤ 香港科技大学(广州) | AI-Thrust (ENVISION LAB) | 研究助理 | 2025/02--2025/08 |
| ➤ 美团(北京三块技术) | 基础安全部门 | 算法策略 | 2024/03--2024/08 |

科研项目

- **StereoPilot: Learning Unified and Efficient Stereo Conversion via Generative Prior** Arxiv 2025
Guibao Shen*, **Yihua Du***, Wenhang Ge*, Jing He, Chirui CHANG, Donghao Zhou, Zhen Yang, Luozhou Wang, Xin Tao, Ying-Cong Chen[†]
 - 共同第一作者(第二位), 我们系统性指出单目转双目视频主流 DWI (Depth-Warp-Inpaint) 管线的关键瓶颈: 误差级联、反射/透明等场景的深度歧义, 以及 Parallel 和 Converged 立体格式不一致带来的训练/评测错配与不公平对比; 据此构建并发布 UniStereo-首个同时覆盖两种立体格式的统一大规模基准(约 103K 立体视频对, 含文本标注与标准化测试集), 并提出 StereoPilot, 基于预训练视频扩散 Transformer 的端到端前馈新视角合成框架, 无需显式深度与迭代采样即可直接生成目标视角, 结合可学习域切换器以及循环一致性约束实现跨格式自适应与更稳定的视差一致性; 在 UniStereo 基准上相较现有方法在视觉保真度/一致性与推理效率方面均取得显著提升。
- **VideoMemory: Toward Consistent Video Generation via Memory Integration** Arxiv 2026
Jinsong Zhou*, **Yihua Du***, Xinli Xu*, Luozhou Wang, Zijie Zhuang, Yehang Zhang, Shuaibo Li, Xiaojun Hu, Bolan Su, Ying-Cong Chen[†]
 - 共同第一作者(第二位), 提出 VideoMemory, 一个用于多镜头视频生成的实体中心框架。核心创新在于将长程一致性问题重新定义为实体资产管理问题, 为角色、道具与场景维护可持续复用且可更新的显式实体状态; 据此设计可读写的 Dynamic Memory Bank 与多智能体生成流程, 采用“基于实体属性检索—未命中则生成新实体—写回更新入库”的闭环机制, 实现已有实体的跨镜头复用与新实体的增量构建, 从机制上抑制跨镜头身份漂移并支持剧情驱动的受控演化。与此同时, 构建首个包含 54 个案例的多镜头一致性评测基准, 覆盖角色、道具与背景持久化场景, 为长程实体一致性提供标准化评测, 并在实验中验证了方法在一致性与整体观感上的提升。
- **A Mechanistic View on Video Generation as World Models: State and Dynamics** Arxiv 2026
Luozhou Wang*, Zhifei Chen*, **Yihua Du**, Dongyu Yan, Wenhang Ge, Guibao Shen, Xinli Xu, Leyi Wu, Man Chen, Tianshuo Xu, Peiran Ren, Xin Tao, Pengfei Wan, Ying-Cong Chen[†]
 - 我们深入分析了视频生成模型如何逐步演化为世界模型模拟器。首先引入“状态(state)”这一核心概念, 随后围绕 State Construction 与 Dynamics Modeling 两条主线, 对现有方法进行了系统梳理; State Construction 上, 我们总结了隐式状态范式中基于上下文管理的记忆机制, 以及显式状态范式中通过紧凑潜变量实现长期状态建模的代表性路径, 并分析了不同状态构建方式在长期一致性、历史压缩和可扩展建模上的关键权衡。在此基础上, 我们进一步回顾了视频生成模型在因果动态建模、时空一致性建模以及未来演化生成方面的能力, 讨论了其迈向更强世界模型模拟器的潜在方向, 并对未来研究进行了展望。

科创竞赛

- 基于多角度取证的活体检测方法 哈工大本科综合设计-卓越创新综合设计奖(2025)
- 基于扩散模型的虚拟创作交流社区 国家级大创(2024)
- 第六届全球校园人工智能算法精英大赛 全国二等奖(二等奖排名第一、准确率全国第二)(2024)
- 第十七届全国大学生软件创新大赛 东北区域赛二等奖(东北赛区前 20 名)(2024)