



AI for VR: 虚拟现实中的人工智能

王莉莉¹, 许威威², 刘烨斌³, 汪淼¹, 王贝贝⁴, 杨旭波⁵, 许岚⁶, 谭张尧¹, 樊润泽¹,
王梓竣¹, 王驰², 张鸿文⁸, 温怡健⁸, 杨浩中¹, 吴健^{1*}, 樊家辉⁹, 王晖⁵, 张启煊⁶,
汪国平⁷, 王涌天¹⁰, 赵沁平¹

1. 北京航空航天大学虚拟现实技术与系统全国重点实验室, 北京 100191

2. 浙江大学计算机辅助设计与图形系统全国重点实验室, 杭州 310058

3. 清华大学自动化系, 北京 100084

4. 南京大学智能科学与技术学院, 南京 215163

5. 上海交通大学软件学院, 上海 200240

6. 上海科技大学信息科学与技术学院, 上海 201210

7. 北京大学计算机学院, 北京 100871

8. 北京师范大学人工智能学院, 北京 100875

9. 南京理工大学计算机科学与工程学院, 南京 210094

10. 北京理工大学光电学院, 北京 100081

* 通信作者. E-mail: lanayawj@buaa.edu.cn

随着硬件与信息技术的持续进步, 虚拟现实在医学^[1]、教育^[2,3]和工业^[4]等多个领域的应用日益广泛, 且其应用形态也愈加复杂. 虚拟现实系统的特征已从最初的“3I”——沉浸感 (immersion)、交互性 (interaction) 和构想性 (imagination), 逐步演化为“6I”, 新增了智能化 (intelligentization)、互联性 (interconnection) 和演化性 (iteration). 其中, 智能化正成为虚拟现实技术不可逆转的发展趋势, 并受到越来越多的关注. 特别是近年来, 随着神经辐射场 (neural radiance fields)、三维高斯泼溅 (3D Gaussian splatting)、神经渲染 (neural rendering)、生成对抗网络 (GANs)、扩散模型 (diffusion models) 以及大型语言模型 (LLMs) 等人工智能技术的迅速发展, 虚拟现实的核心技术正在迎来深刻变革. 这些前沿人工智能模型正在显著推动虚拟现实的内容生成渲染、交互方式与系统性能的提升. 因此, 有必要对这些关键的虚拟现实技术进行系统性的梳理与评估, 结合最新的人工智能进展, 旨在提供新的研究视角, 并帮助新进入该领域的研究人员迅速掌握当前的突破性进展.

SCIENCE CHINA Information Sciences 在第 69 卷

第 1 期出版了 “Artificial intelligence for virtual reality: a review”, 系统回顾并分析了精心筛选的 485 篇相关文献, 旨在全面梳理人工智能方法在虚拟现实核心技术中的融合应用, 呈现该领域的最新研究进展与发展趋势. 我们将虚拟现实研究的核心对象划分为三大类: 虚拟场景、虚拟角色以及二者之间的交互. 在此基础上, 对虚拟场景相关内容进行了更为精细的拆解与归纳, 最终提出了 6 大人工智能与虚拟现实深度融合的核心研究方向, 分别为: 人工智能驱动的内容表征、内容渲染、内容生成、物理仿真、虚拟角色构建以及人机交互. 这一分类框架不仅理清了当前研究的技术脉络, 也展示了每一方向在突破性成果方面的代表性进展, 全面体现出人工智能在推动虚拟现实技术革新与应用变革方面的强大驱动力.

● 人工智能驱动的内容表征: 神经辐射场 (NeRF) 与 3D 高斯点云 (3DGS) 成为两大主流 AI 驱动表征方式, 前者以隐式神经函数实现高质量视角合成, 后者通过显式高斯椭球实现实时渲染与交互, 二者在重建精度、渲

英文原文: Wang L L, Xu W W, Liu Y B, et al. Artificial intelligence for virtual reality: a review. Sci China Inf Sci, 2026, 69: 111101, doi: 10.1007/s11432-024-4541-9

染速度与可编辑性上形成互补;

- **内容渲染:** 聚焦可微渲染、NeRF 渲染与 3DGS 渲染三大管线,探讨了如何通过神经网络优化光照、材质与视角合成,并引入人眼视觉模型实现注视点渲染,显著降低计算负载;

- **内容生成:** 梳理了从 2D 生成模型“升维”到 3D、引入多视角先验、再到原生 3D 生成模型的发展脉络,指出扩散模型与 Transformer 架构在文本/图像驱动 3D 内容创作中展现强大潜力,但仍面临几何一致性、生成速度与可编辑性挑战;

- **物理仿真:** AI 对传统仿真流程的重塑.神经网络不仅可替代或加速流体、软体、刚体等复杂物理求解器,还能通过可微仿真实现反向优化与实时控制,推动 VR 从“可视”走向“可感”;

- **虚拟角色构建:** 涵盖动作生成、avatar 重建与智能体行为建模,指出文本/音频驱动的动作合成、基于 NeRF 或 3DGS 的高保真 avatar、以及大模型驱动的自主智能体正成为构建“数字人”生态的三大支柱;

- **人机交互:** 从行为识别、交互优化到感知增强层层递进,强调多模态感知(眼动、手势、生理信号等)与 AI 模型(CNN、LSTM、强化学习等)在提升交互自然性、个性化与沉浸感中的关键作用.

该文对每个主题的研究发展脉络进行梳理,介绍代表性方法和技术,分别分析其挑战与机会,为该领域的研究人员提供新的视角.

根据对 AI 技术在虚拟现实各个子领域主题的调研,我们发现 AI 已全面渗透 VR 内容生产与体验闭环,使虚拟环境在视觉真实、物理可信、角色智能与交互自然 4 个方面实现质的飞跃;然而,这一融合仍面临多重深层挑战.

内容表征方面, NeRF 与 3DGS 的内存占用高昂、少样本重建困难、编辑接口匮乏,限制了大规模动态场景的实时应用;**内容渲染方面,** 动态信息不可预获取导致时序一致性难以保证,神经网络对高频光照与复杂几何的表达能力不足,去噪与超分在毛发、透明等极端材质上仍显吃力;**内容生成方面,** 扩散模型迭代式去噪带来的延迟阻碍了“所想即所得”的即时创作,且海量 3D 资产缺乏统一语义与检索机制,难以与具身 AI 系统无缝衔接;**物理仿真方面,** AI 黑盒特性与物理可解释性的矛盾尚未解决,实时高精度多物理场耦合仿真依旧计算

密集;**虚拟角色方面,** 高保真服装动态与微表情尚无法兼顾,情感认知模型缺乏对人类心理与文化背景的深度建模,智能体决策过程不可视、不可调,削弱了用户对决策的信任度;**人机交互方面,** 高速、高精度、多模态行为识别在遮挡、运动模糊与信号混叠下鲁棒性不足,自然交互范式尚未跳出“手势+语音”的窠臼,跨用户、跨场景的个性化适配仍缺少统一框架.此外,数据隐私与伦理风险、终端算力与能耗瓶颈、以及元宇宙级大规模并发场景下的负载均衡,都是横亘于 AI+VR 愿景前的现实壁垒.

面向未来,亟需研究构建更轻量、可解释、可编辑的 3D 表征,使 NeRF 与 3DGS 在保持视觉质量的同时降低显存、支持语义级修改;发展面向动态、无先验信息的实时渲染新管线,让神经网络学会“想象”运动与光影;探索基于蒸馏、量化或潜空间初始化的快速 3D 生成框架,并建立可检索、可组合、可语义查询的 3D 资产库;推动可微仿真与神经求解器的模块化、可验证化,实现物理规律与数据驱动的高效协同;在虚拟角色方向,融合 3DGS、可微仿真与大模型,打造具备肌肉-服装-表情联动、情感记忆与可解释决策的数字生命体;交互层面则迈向低延迟、多模态、个性化行为意图理解,让系统在人眼转动、心率波动、语音停顿间捕捉需求,并以自然语言、触觉反馈、环境动态调整等方式主动回应.最终,AI 将不仅是 VR 内容的生产工具,更将成为虚拟世界的“内在规律”与“智能居民”,与人类共建持续演化、情感共鸣、物理可信、体验极致的元宇宙.

参考文献

- 1 Wu Y, Hu K, Chen D Z, et al. AI-enhanced virtual reality in medicine: a comprehensive survey. 2024. ArXiv:2402.03093
- 2 Bosman K, Bosse T, Formolo D. Virtual agents for professional social skills training: an overview of the state-of-the-art. In: Proceedings of the 10th EAI International Conference on Intelligent Technologies for Interactive Entertainment, 2019. 75-84
- 3 Gandedkar N H, Wong M T, Darendeliler M A. Role of virtual reality (VR), augmented reality (AR) and artificial intelligence (AI) in tertiary education and research of orthodontics: an insight. Semin Orthod, 2021, 27: 69-77
- 4 Devagiri J S, Paheding S, Niyaz Q, et al. Augmented reality and artificial intelligence in industry: trends, tools, and future challenges. Expert Syst Appl, 2022, 207: 118002