

# 三维视觉与元宇宙专题论文点评

曹明伟<sup>1</sup> 李成龙<sup>2</sup> 高浩<sup>3</sup> 朱虎<sup>4</sup> 吕智涵<sup>5</sup>

1 安徽大学计算机科学与技术学院 合肥 230601

2 安徽大学人工智能学院 合肥 230601

3 南京邮电大学自动化学院 南京 210023

4 南京邮电大学通信与信息工程学院 南京 210003

5 乌普萨拉大学 乌普萨拉 瑞典 SE-75105

2021年 Mark Zuckerberg 将 Facebook 更名为 Meta,并向外界介绍了元宇宙技术。自此,元宇宙受到了工业界和学术界的广泛关注,并成为热点研究问题。元宇宙是一个虚拟的数字宇宙,是真实世界的虚拟再现。三维视觉作为元宇宙的一个重要基础技术组成部分,能为构造数字宇宙提供创建虚拟场景的解决方案。然而,针对不同的元宇宙应用场景,需要不同的三维视觉技术为其创建虚拟场景。此外,随着人工智能技术的快速发展,元宇宙的基础理论架构有待进一步探索和完善,且元宇宙技术的潜在应用场景也需要进一步挖掘。

为了促进三维视觉与元宇宙技术的研究和应用,及时、集中、全面地报道其相关技术、理论和算法等方面的最新研究成果和进展,《计算机科学》组织了一期专题,面向科研工作者征集了相关研究论文,经过严格的评审程序,最终录用了 11 篇优秀专题论文,包括 2 篇综述论文和 9 篇基础算法论文。希望这些最新的前沿科研成果能为相关领域的专家学者提供参考,促进学术界和工业界的深度融合,推动三维视觉与元宇宙技术的发展。

## 一、综述论文

1) 论文《元宇宙技术发展与应用综述》全面深入地阐述了元宇宙技术的起源、发展过程和应用,探讨了当前面临的问题,同时对元宇宙的未来发展方向进行了展望。此综述论文能够为读者提供关于元宇宙技术的全面的理论框架,为未来进一步对其探索和应用提供参考。

2) 论文《元宇宙中三维场景重建技术综述》对元宇宙中基于场景生成的三维重建技术进行了总结和归纳。首先回顾了元宇宙的发展历程;然后梳理了基于三维高斯和神经辐射场表示的三维重建技术,重点分析了三维重建技术与触觉信号以及大语言模型的创新融合方法;最后探讨了元宇宙中基于场景生成的三维重建技术面临的挑战和未来研究方向。此综述论文全面概述了元宇宙中的三维重建技术,有助于元宇宙中三维重建技术的发展和應用。

## 二、基础算法

1) 深度图估计是三维重建领域的基础研究问题之一,论文《LpDepth: 基于拉普拉斯金字塔的自监督单目深度估计》提出了一种基于拉普拉斯金字塔的自监督单目深度估计方法。该方法的核心思想是:首先,使用拉普拉斯残差图丰富编码特征,以弥补在下采样过程中丢失的特征信息;然后,在下采样过程中使用最大池化层突显和放大特征信息,使编码器在特征提取过程中更容易提取到训练模型所需要的特征信息;最后,使用残差模块解决过拟合问题,提高解码器对特征的利用效率。在 KITTI 等多个公开数据集上的测试结果均证明了该方法的有效性。

2) CT 图像分割是医学图像处理领域的热点研究问题,论文《基于边缘增强的选择性特征融合肾癌三维 CT 图像分割》针对肾癌三维 CT 图像存在病变区域多尺度、边缘像素稀疏、对比度低以及肿瘤形状复杂且不规则等问题,提出了一种基于边缘增强的选择性特征融合肾癌三维 CT 图像分割网络,所提方法在保证病变区域整体定位准确的同时,强化了对小目标特征信息的挖掘利用,从而提高了分割的准确性和鲁棒性。其在 KiTS19 公开数据集上取得了较好的结果。

3) 虚拟化身在元宇宙场景中具有广泛的应用前景,论文《基于区域编码的可驱动头部虚拟化身重建算法》针对基于 3D 可变形模型的头部化身方法存在的问题,提出了一种以单目视频为原始数据,将一个数量动态增长的点云用来构建可驱动的头部虚拟化身的方法。此方法能够计算出高精度的纹理信息,且在多个开源数据集上均取得了优越的结果。

4)人像合成能够为元宇宙场景构造人物数字内容,论文《基于区域显著性与空间特征提取的说话人像合成方法》针对基于神经辐射场(NeRF)的多个说话人像合成方法存在语音-嘴唇同步欠佳、躯干抖动和合成视频清晰度较低等问题,提出了一种基于区域显著特征和空间体积特征的高保真说话人像合成方法。在多个人物主体的测试中,所提方法的结果优于当前最优的基线模型。

5)三维人体姿态估计是三维视觉领域的热点研究问题,论文《基于中心点注意力的多视角多人三维人体姿态估计》融合了空间体素方法与基于回归的姿态估计方法的优点,提出了一种基于中心点注意力回归的多视角多人三维人体姿态估计方法。该方法使用一个小规模的体素网络粗略估计人体中心点位置,并以此构建初始姿态,随后在人体中心点的范围内进行回归预测,得到更精确的人体姿态。此方法具有较好的泛化性和通用性,使得基于回归的模型可以在新场景下通过小数据量的训练快速部署。

6)单视图草图三维重建是三维重建领域的热点研究问题,也是为元宇宙空间创造数字内容的技术之一。论文《基于注意力机制与对比损失的单视图草图三维重建》针对手绘草图难以避免的前景和背景模糊、绘制风格差异和视角偏差问题,提出了一种基于注意力机制与对比损失的单视图草图三维重建方法,重建过程无需交互,实现了自动化重建过程。多个实验结果均证明了该方法的有效性。

7)地理定位在元宇宙中具有广泛的应用前景,论文《跨视角地理定位中的三维交互机制》针对现有跨视角地理定位方法存在的问题,提出了一种跨视角地理定位三维交互机制,所提出的三维交互机制在3个通道中使用不同的注意力,使模型对跨视角图像的平移、缩放、旋转具有鲁棒性。实验结果证明了该方法在无人机视角定位和无人机导航任务中的有效性。

8)隐私保护是元宇宙空间构建亟需研究和解决的问题之一,论文《基于拆分联邦学习的元宇宙视线交互中的隐私主动保护方法研究》针对目前元宇宙中许多视线交互功能往往需要利用特定个体隐私以提供更好的使用体验,从而导致视线交互在效用与隐私之间存在显著的矛盾冲突问题,创新性地提出了一种融合联邦学习与拆分学习的视线隐私保护方法,通过构建联邦化的视线估计分布式架构实现隐私保护的多个体客户端联合训练,引入信息瓶颈技术剥离个体隐私并于本地进行保护。此方法在视线数据的隐私保护和交互性能方面展现出了双重优势,为元宇宙环境中视线交互的效用与隐私保护提供了新颖的解决方案和技术路径。

9)三维目标检测是三维视觉领域的热点研究问题之一,也是自动驾驶中最关键的技术之一。论文《融合动态加权图卷积的三维目标检测》针对基于激光雷达的三维目标检测方法不能充分利用点云的结构信息,导致目标物体的误检和漏检问题,提出了基于动态加权图卷积的 DEG R-CNN 方法。在 KITTI 等数据集上的测试结果均证明了此方法的有效性。



**曹明伟** 博士、安徽大学计算机科学与技术学院副教授、硕士生导师,2017年博士毕业于合肥工业大学。主要研究领域为三维重建和计算机视觉。发表论文50余篇,授权国家发明专利20项。IEEE会员,中国计算机学会会员,中国图象图形学学会会员,中国人工智能学会会员,中国图象图形学学会三维视觉专委会委员。



**李成龙** 安徽大学人工智能学院教授、博士生导师、元宇宙工程系系主任,博士毕业于安徽大学,曾在中国科学院自动化研究所从事博士后研究。主要研究领域为计算机视觉和深度学习。在 IEEE TPAMI, IEEE TIP, IEEE TNNLS 等国际权威期刊和 CVPR, ECCV, AAAI 等顶级会议上发表学术论文120余篇。授权国家发明专利15项,其中1项实现重要成果转化。主持国家自然科学基金(2项)、安徽省杰出青年基金等科研项目。获得安徽省科技进步二等奖、安徽省教学成果一等奖(2项)等奖项。



**高 浩** 博士、南京邮电大学自动化学院教授、博士生导师。研究方向为人工智能和三维重建。拥有境外3年以上研究工作经历,发表SCI论文50余篇,授权专利20余项。中国人工智能学会脑科学与人工智能专委会委员,中国电子学会VR/AR技术与产业专委会委员,中国图象图形学学会三维视觉专委会委员。



**朱 虎** 南京邮电大学通信与信息工程学院教授,博士生导师。主要从事人工智能与图像处理等方面的研究。在国际期刊及国际会议上发表学术论文100余篇,申请发明专利60余项;主持完成了多项国家自然科学基金、国防项目以及其他各类课题20余项。担任IEEE、计算机学会、中国人工智能学会、中国图象图形学学会会员。



**吕智涵** 博士,乌普萨拉大学副教授,博士生导师。IEEE高级会员,英国计算机学会会士(BCS Fellow),美国计算机协会杰出演讲者(ACM Distinguished Speaker),4次入选斯坦福全球前2%顶尖科学家终身科学影响力排行榜,中国图象图形学学会虚拟现实专委会委员。获得中国海洋大学和法国巴黎大学联合培养博士学位,曾担任法国国家科学研究中心研发工程师,瑞典于默

奥大学博士后,西班牙FIVAN基金会资深研究员,英国伦敦大学学院博士后,西班牙巴塞罗那大学博士后,中科院深圳先进院助理研究员,获得欧盟玛丽居里学者称号。发表高质量论文300余篇,获瑞典UMINOVA学术商业竞赛“Best Idea”奖、中国“挑战杯”创业计划竞赛特等奖、IEEE Access杰出编辑奖等20多个国内外奖项,担任30多个高水平期刊的编辑。