



王立军，国家特聘专家，俄罗斯工程院外籍院士，北京市战略科学家，目前担任西安电子科技大学杭州研究院先进视觉研究所所长、领军教授，SID 国际信息显示学会（中国）AR/VR 专业技术委员会主席，曾担任美国西北大学博士后研究员，世界 500 强美国康宁公司高级研究科学家。主要从事超高速光纤通信与光互联，光电集成，新型光电显示，XR 与元宇宙技术及应用研发。

卷首语

Foreword

元宇宙造物主：人工智能 3D 场景和交互内容生成

王立军

元宇宙是下一代沉浸式互联网，是未来数字经济的集大成者。在元宇宙中，人工智能扮演着至关重要的角色，特别是在 3D 场景和交互内容生成方面，AI 不仅可以作为“造物主”创造出丰富多彩的 3D 虚拟世界，还能提供自然的人机交互体验。

人工智能在 3D 场景生成中的应用，主要依赖于深度学习和计算机视觉技术。通过训练大量的 3D 模型数据，AI 可以学习到物体的三维结构、材质、光照效果等信息，从而自动生成复杂的 3D 场景。例如，使用生成对抗网络（GANs）可以创建出逼真的 3D 模型和虚拟环境。神经辐射场（neural radiance fields, NeRF）用于从稀疏视图创建高质量的 3D 场景。NeRF 通过建模光线在场景中的行为，能够处理复杂的光照和遮挡关系，生成新的视角图像，适用于高质量的虚拟现实内容制作和影视特效。从复原消失了的古老城堡到创造未来城市，AI 已经能够根据给定的条件和参数创造出具有高度真实感的局部场景。这些技术的发展，使得元宇宙中的世界构建变得更加快速和高效。

在元宇宙的虚拟世界中，用户交互体验的自然性和流畅性至关重要。人工智能在此领域的应用包括自然语言处理（natural language processing, NLP）、语音识别和机器视觉等技术，它们共同作用于提高用户与虚拟环境之间的交互质量。AI 可以实现虚拟角色的自然语言对话，模拟真实人类的行为和反应，甚至能够根据用户的行为和偏好进行个性化的交互内容生成。这些交互不仅限于文本和语音，还包括肢体语言和表情识别，让虚拟交流变得更加丰富和真实。

尽管人工智能在元宇宙的 3D 场景和交互内容生成方面展现出巨大的潜力，但仍然面临巨大的技术挑战。首先，生成高质量、大规模连续 3D 场景和交互内容需要在模型算法方面取得突破，比如英伟达提出的状态空间模型（state space model, SSM），需要比文本和图像生成高出数量级的计算资源。其次，保持虚拟环境的真实感和交互的自然性，需要 AI 技术在理解复杂环境和用户意图方面达到更高的准确度。随着生成式人工智能在文本、图像和视频（Sora）生成领域接连取得突破，我们相信 3D 人工智能生成（3D AIGC）技术的成熟将指日可待。