

元宇宙与现实双空间交互系统研究

陈 杰¹
CHEN Jie

摘 要

现阶段线下观展存在拍照不清晰和游览不全面的问题,线上云观展存在体验不真实的问题,针对以上问题,文章提出一种“云用户—现实用户”双向实时交互系统。通过整合 GPS 精确室内定位技术与 Bursa-Wolf 转换模型构建出厘米级空间映射网络,同时结合三维重建技术将现实展馆 1:1 复刻到元宇宙展馆中,最后创新性引入“位置触发+用户意图”双重匹配机制,将云用户与现实用户的交互时延限制在 700 ms 内。为提升合成图像的自然度,提出了基于 NeRF 与 HDR 的虚实同框合成算法。实践证明,该算法显著提升了线下观展用户的拍照分享意愿与云观展用户的留存率,同时提升了用户对于展馆知识的吸收率。

关键词

元宇宙; Bursa-Wolf 模型; 实时交互; 合成图像; 文化展览

doi: 10.3969/j.issn.1672-9528.2026.07.028

0 引言

随着数字技术的飞速进步,人们的沟通方式也产生了巨大变化。5G 通信、云端实时渲染,还有 GPS 定位等技术的不断精进,让元宇宙^[1]理念逐渐渗透到人们日常工作生活场景中。元宇宙技术作为虚实融合的新型数字生态体系,构建了现实与线上虚拟空间的实时交互桥梁,但在真正使用时还是面临着许多技术问题。尤其在复杂的跨时空交互领域场景下,现在的技术方案普遍存在着通信慢、位置记录误差大等问题。

以博物馆展览场景为例,传统线下观展模式存在场馆光线较暗导致拍摄照片质量差、游客位置随机导致路人容易入镜、展品静态陈列且有的位置较远无法观其全貌等问题。游客在游览时可能会因为没有拍到满意的照片、某个展品前人多无法了解展品信息影响其观展心情。单纯的线上云展览虽不受游客地理位置的限制,却因游客无法实时感知导致达不到线下观展的真实体验感,这种虚实割裂的现状同时影响着云游客与现实间的观展体验。本文巧妙结合元宇宙与人工智能技术,彻底打通元宇宙-现实双空间的社交壁垒,创造出的一套智能交互系统。该系统主要从三个方面对现有技术进行改造:一是持续优化改进的 Bursa-Wolf 坐标转换模型,将现实空间与虚拟空间的精准映射,有效解决了游客位置坐标误差大的问题;二是构建多形态交互模块,实现语音、文字、图像等多形态的实时双向通信,有效解决了交互场景单一的

问题;三是本系统的元宇宙-现实双空间影像合成模块,通过智能图像融合算法将现实人物与元宇宙空间对应位置的背景无缝融合,有效解决了线下游客的分享欲望低迷的问题。该方案具备良好的场景迁移性,可广泛应用于各类文旅场所,未来也许会改变各大景区的游览模式,激发游客参与文化旅游活动的积极性。

1 相关研究现状与技术基础

1.1 虚拟现实交互技术发展现状

虚拟现实交互技术近年来已从 VR 设备等应用,扩展到日常的工作生活等场景。在硬件层面,头戴式显示设备让用户有了沉浸式的视觉体验,AI 手套等设备则进一步丰富了用户的触觉体验。然而,这些设备大多沉浸于虚拟空间的沉浸式交互,在元宇宙云空间与现实用户之间的交互产品反而很少。

现阶段的虚拟交互技术由于 GPS 的定位技术^[2]在室内定位不准确、跨网络场景通信时延较高、用户交互方式单一难以满足现在年轻化用户的多样化互动需求。针对上述问题国内外专家提出了多种解决方案,比如在获取用户实时位置时,加入了超宽带(ultra-wideband, UWB)和视觉-惯性里程计(visual-inertial odometry, VIO)实现厘米级精确定位和低能耗通信。在通信时延大的问题上,通过结合 5G 技术和边缘计算将时延降低至 700 ms,为后续降低交互时延提供了有效帮助。

从实用性方面看,如果使用上述方案,方案部署的成本会大幅增加,而且现阶段交互协议的不统一,导致不同系统

1. 中通服咨询设计研究院有限公司 江苏南京 210019

间的兼容性较差。以上这些因素都是阻碍元宇宙世界与现实世界交互产品产生的主要原因，需要尽快革新技术让这类产品问世。

1.2 元宇宙 AI 技术关键算法

元宇宙的出现为跨空间实时互动提供了新的思路，其核心在于展示现实世界进行建模生成虚拟的平行世界的思路。典型的元宇宙 AI 技术算法包含感知层、网络层、计算层和应用层四个层级：感知层通过多源传感器采集现实世界数据；网络层依托 5G/6G 和边缘计算提供低时延传输；计算层利用云渲染和 AI 算法处理复杂任务；应用层则支持各类交互场景的快速部署。

在转换现实游客坐标方面，通过 Bursa-Wolf 模型^[3]引入平移、旋转参数和尺度因子，实现了从真实世界到元宇宙空间的 1:1 复刻。该模型使用坐标转换公式将现实游客的真实位置转换到元宇宙空间的对应位置，并且位置误差控制在既定范围内，比传统方式的转换准确率大幅提升。具体来说，AI 元宇宙 - 现实双空间互动算法先计算出转换参数，随后利用三维大地坐标^[4]与空间直角坐标的微分关系对现实游客的位置进行实时映射，并且将游客在元宇宙世界的数字人实时移动到对应的位置上。

现阶段最先进的图像虚实融合技术，是由神经辐射场（neural radiance fields, NeRF）^[5]与高动态范围（high dynamic range, HDR）融合技术^[6]结合实现的。NeRF 利用隐式神经表示场景建模时的光照特性，同时利用 HDR 技术保留了真实环境的光照动态范围，确保图像融合的真实性。实践证明，使用这种方式合成的图像比传统方法合成的图像更加自然真实，同时大大降低其渲染时长。从技术视角看，该图像合成方案为构建可扩展的元宇宙交互系统做出了很大贡献，但在渲染效率和能耗优化方面仍有改进的空间。

2 系统设计与实现方法

2.1 系统架构设计

本系统采用“云端元宇宙—现实展馆”双层架构（如图 1 所示），实现线下用户与云用户在文化展览场景中的深度交互。

（1）物理感知层

多源定位模块：使用智能手机的 GPS 定位功能，并结合系统的 UWB 信标高精度定位技术获取现实游客在展馆中的实时位置，并且通过 Bursa - Wolf 模型将实时位置转换成元宇宙空间的对应位置，驱动现实游客在元宇宙中数字人形象移动到该位置。下面详细介绍 Bursa - Wolf 模型是如何将两个不同空间直角坐标系的坐标进行转换。

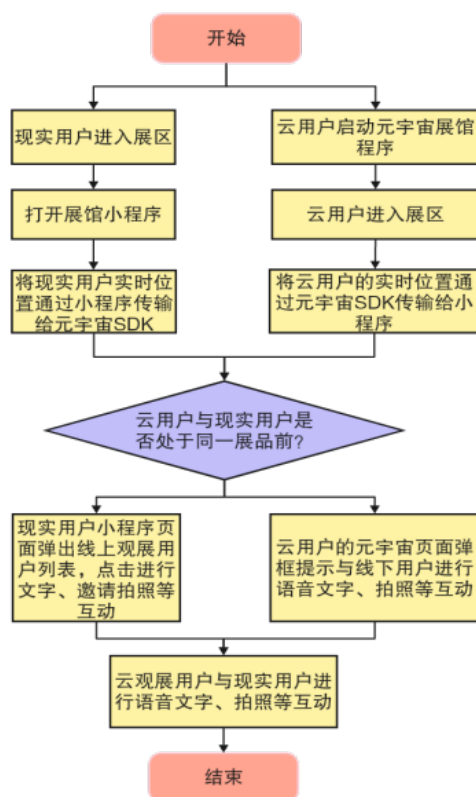


图 1 云端元宇宙 - 现实展馆架构图

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_D = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_G + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + \begin{bmatrix} 0 & -Z & Y & X \\ Z & 0 & -X & Y \\ -Y & X & 0 & Z \end{bmatrix} \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \\ k \end{bmatrix} \quad (1)$$

式（1）为 Bursa - Wolf 模型的简化形式，整个公式包含 7 个转换参数。 Δx 、 Δy 、 Δz 是三个平移参数， ε_x 、 ε_y 、 ε_z 是三个旋转参数， k 是尺度因子。通过 Bursa - Wolf 模型可以直接转换两个不同空间的坐标，当公共点数为 3 个以上时，便可以通过平方差的方法求得转换参数。求得转换参数后，再利用该模型进行坐标之间的转换。

$$\begin{cases} X = (N + H) \cos B \cos L \\ Y = (N + H) \cos B \sin L \\ Z = [(N(1 - e^2) + H)] \sin B \end{cases} \quad (2)$$

$$N = a / \sqrt{1 - e^2 \sin^2 B} \quad (3)$$

式（2）~（3）在 Bursa - Wolf 模型基础上进行优化，将 GPS 坐标和三维大地坐标相互转换， N 为椭球曲率半径， a 为椭球长半轴， e 为椭球第一偏心率， H 为大地高。

设高程异常为 s ，正常高为 h ，用公式表示为：

$$H = h + \xi \quad (4)$$

通过公式（1）~（4）可以得到三维坐标系转换误差方程：

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}_D = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_G + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + \begin{bmatrix} 0 & -Z & Y & X \\ Z & 0 & -X & Y \\ -Y & X & 0 & Z \end{bmatrix} \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \\ k \end{bmatrix} - \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_D - \begin{bmatrix} \cos B \cos L \\ \cos B \sin L \\ \sin B \end{bmatrix}_G \xi \quad (5)$$

在式(5)中的 $\begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \\ k \end{bmatrix}$ 部分是由 h 替换成 H 代入式(2)中

得出,将计算出的元宇宙空间位置和游客的实际位置之间的误差控制在既定范围内。

环境感知模块: 基于激光雷达点云数据(LiDAR point cloud data, LiDAR)与RGB-D相比,实时捕捉展品几何特征,根据捕捉到的数据,例如:展品位置、光照参数等元数据1:1构建出元宇宙展区。

(2) 数据处理层

用户行为预测模块: 采用长短期记忆(long short-term memory, LSTM)网络^[7]分析游客在元宇宙空间的运动轨迹、通信时长等行为,提前预测游客未来行为,为系统的优化做好铺垫。

分布式消息队列: 基于分布式流式日志(apache kafka cluster, Kafka)集群应用高峰时刻的消息通信,通过负载均衡算法将消息通信时延控制在预定目标值内。

(3) 交互应用层

多形态通信模块: 使用消息队列遥测传输(message queuing telemetry transport, MQTT)协议实现元宇宙世界的云游客与现实游客之间实时文字通信,使用Web实时通信(web real-time communication protocol, WebRTC)协议将游客间的语音消息准确、准时传输。

系统增加图像融合模块,使用ARCore渲染管线技术,将元宇宙3D模型背景与人物抠像自然融合,并通过通信模块将融合图片传回到现实游客手机,游客可以自由选择保存图片以及分享给云端游客。

2.2 核心功能实现

2.2.1 虚实空间动态匹配

设计“精确定位+位置同步”双重触发机制:

精确定位: 当用户进入展品交互区域时,通过系统的定位模块获取现实游客的精确位置,通过式(5)计算出坐标偏差:

$$\Delta = \sqrt{[(xr - xv)^2 + (yr - yv)^2]} \quad (6)$$

若 Δ 小于既定目标值且停留时间超过10s,触发交互。

位置同步: 将提取到现实游客的实时位置同步给元宇宙系统,实现及时更新现实用户在元宇宙中位置的功能,让现实游客的位置误差保持在既定范围内。

2.2.2 元宇宙与现实实时协同拍照系统

实现流程如图2所示。

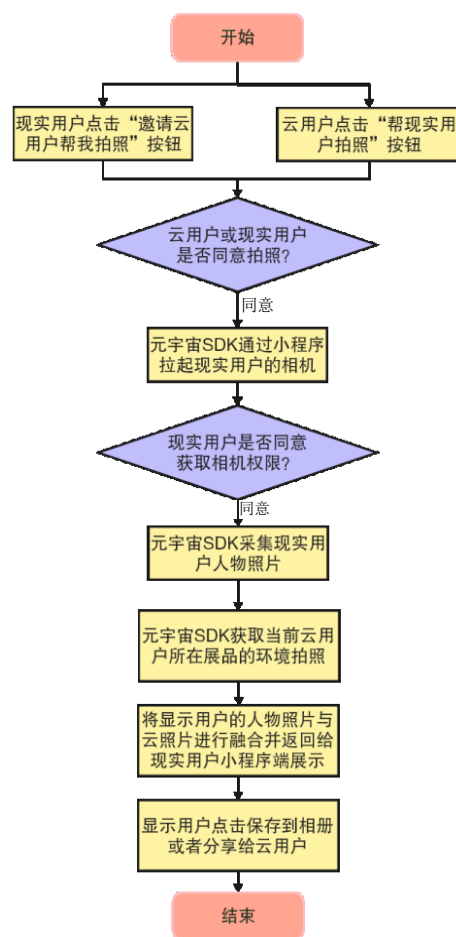


图2 元宇宙与现实实时协同系统拍照流程图

图像采集: 通过智能手机摄像头获取线下游客的图像,将其传入AI元宇宙-现实双空间互动系统的图像融合模块。

前景分割: 系统的图像融合模块借助卷积神经网络(convolutional neural network, CNN)^[8]抠除图像背景,保留人物抠像。然后根据传入的图像数据调整人物抠像的曝光度等光线参数,让其不会太暗或者太亮。

神经渲染合成: 图像融合模块通过NeRF模型获取元宇宙空间对应位置的背景,与上面的人物抠像进行融合生成新图像,并返回给现实游客的手机,现实游客可以保存到手机,也可以分享给云游客。

2.2.3 多形式交互优化

本系统在语音通道方面,采用降噪算法(RNNoise模型)^[9]对干扰噪声进行过滤,优化游客间语音的清晰度,使得语音交互更加流畅自然。

本系统在文字通道方面,通过加入AI技术,对相关知识进行文字解读,拓宽了相关的知识面,使游客的体验感更加舒适。

2.3 关键技术突破

(1) 高精度坐标转换算法

本研究对Bursa-Wolf模型进行优化创新:

①通过原有技术的基础上,引入高程异常修正项 Δh ,可以减少由于不重合导致的椭球面偏差,如式(4)所示,提高系统定位的精确性;

②通过引入自适应粒子群算法智能分析优化相关参数,大幅提升相关效率;

③通过引入动态误差补偿矩阵,将坐标偏移精确控制在3 mm内,提升了空间坐标的整体精确性。

(2) 实时交互引擎

系统将 Unity 虚拟场景与 ARCore 实景数据进行时空对齐,实现实时交互引擎,保证时间戳同步误差 <10 ms。

通过 Kafka 分布式架构^[10]实现全链路的负载均衡,面对数十万级的客户在同时访问时,保持35帧/s渲染帧率。

(3) 智能拍照合成

系统可以同步处理前后台数据,用户在获取预览图的同时,在后台执行图像的优化。

系统可以采用 GPU 加速处理相关任务,突破传统只用 CPU 处理的瓶颈,大幅度缩短等待时间。

2.4 系统性能评估

本系统在展馆测试结果为:室内定位比较精确,平均误差为4.3 cm(标准差 ± 0.8 cm)系统交互比较迅速,端到端传输延迟中位数为182 ms(其中语音210 ms,文字150 ms);系统较为稳定,连续运行3天无故障,CPU/GPU 利用率稳定在65%~78%。使用该系统生成的自拍图对比如图3~4所示。



图3 线下自拍



图4 线上用户帮拍

3 结论与展望

本文元宇宙与现实双空间交互系统,首先对 Bursa-Wolf 模型进行改造使其能准确转换位置坐标,其次对 GPS 定位技术进行改造使其更精确地获取现实游客在展馆中的位置,最后加入神经网络让系统融合的图像更加真实美观,有效解决了现阶段元宇宙云观展用户和现实观展用户结合互动少的难题。实践表明,本系统的坐标转换模块极大降低了坐标转换的误差率,交互模块也降低了消息发送和接收时延,最后生成的图像更加真实、美观,让游客有分享欲望。同时让云观展游客有了更加真实的观展体验,提高了游客在元宇宙的留存率。

未来本系统将从两个方面进行优化:首先,使用更先进的定位系统获取更精确的游客位置,不断优化图像融合算法使其生成的图片更加自然、美观,提升游客的满意度;其次,不断扩展系统的模块,比如加入更多的交互模式,不局限于文字、语音,让其可以应用到教育等更多领域,为元宇宙的商业化发展添砖加瓦。

参考文献:

- [1] 潘宇轩.基于区块链和深度学习的元宇宙虚拟空间构建关键技术研究[D].北京:北京邮电大学,2025.
- [2] 井文杰.GPS定位技术在复杂地形环境中的误差分析与精度提升策略[J].科技风,2025(13):51-53.
- [3] 宋超,原菲菲,尚姣,等.基于 Bursa-Wolf 模型的坐标转换工具设计[J].地理空间信息,2014,12(2):139-140.
- [4] 张莉君,高云涵,邹小凡,等.融合多源多视角数据的真彩三维重建技术[J].测绘通报,2025(6):62-67.
- [5] 雷仲夏.基于V-系统的NeRF改进算法及其在商品重建中的研究与应用[D].北京:北方工业大学,2025.
- [6] 祝新力,张雅声,方宇强,等.多曝光图像融合方法综述[J].激光与光电子学进展,2023,60(22):23-40.
- [7] 印佳丽,韩津,陈斌,等.基于一致性感知特征融合的高动态范围成像方法[J].计算机学报,2024,47(10):2352-2367.
- [8] Gu Jiuxiang, Wang Zhenhua, Kuen J, et al. Recent advances in convolutional neural networks[J]. Pattern Recognition, 2018, 77: 354-377.
- [9] 肖纯鑫,陈雨.基于循环神经网络的实时语音增强算法[J].计算机工程与设计,2021,42(7):1989-1994.
- [10] 周兴龙,李永健,曾纪钧,等.基于CSP的Kafka数据流并发与容错机制的形式化验证[J/OL].计算机科学,1-11[2026-06-29].<https://link.cnki.net/urlid/50.1075.tp.20260527.0905.006>.

【作者简介】

陈杰(1993—),男,山西吕梁人,硕士,研究方向:云计算、人工智能、网络安全等。

(收稿日期:2025-11-27 修回日期:2026-07-10)