

一种遥操作场景下的混合现实 视觉增强交互系统

施乐洋¹, 李晓玲¹, 王龙¹, 陈漳沂¹, 郑帅²

(1. 西安交通大学机械工程学院, 710049, 西安; 2. 西安交通大学电子与信息学部, 710049, 西安)

摘要: 针对遥操作机器人系统在高风险任务中近端交互临场感差、操作者难以获得真实体验感的问题,提出了一种适用于遥操作场景的混合现实视觉增强交互系统。首先,针对远程场景传输和机械臂的远程控制这两个重要环节,通过分析遥操作交互任务中的目标需求,从人体视觉感知与认知特性的角度构建并设计了遥操作场景下的视觉增强交互系统框架;然后根据系统功能需求,研究了混合现实场景下的远程视景实时呈现技术,设计了混合现实场景下的机械臂交互控制策略,实现了混合现实全息对象与真实物理场景相映射的虚实融合效果,并以此搭建了遥操作场景下的混合现实视觉增强交互系统,实现了混合现实终端的可视化与系统端到端的网络通信。实验结果表明:与传统的二维系统相比,该交互系统所采用视觉增强形式的任务完成时间减少了17.3%,准确度提高了7%。证明该系统具有较好的应用前景。

关键词: 混合现实;遥操作;视觉增强;机器人系统;交互设计

中图分类号: TP391.7 **文献标志码:** A

DOI: 10.7652/xjtuxb202310003 **文章编号:** 0253-987X(2023)10-0020-10

A Mixed Reality Visual Augmentation and Interaction System in Teleoperation Scenarios

SHI Yueyang¹, LI Xiaoling¹, WANG Long¹, CHEN Zhangyi¹, ZHENG Shuai²

(1. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China;

2. Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: To address the challenges of poor situational awareness and lack of realistic experience for operators in high-risk teleoperation tasks faced by robotic systems, a mixed reality visual augmentation and interaction system applicable to teleoperation was introduced. Firstly, the system framework was designed from the perspective of human visual perception and cognitive characteristics and based on requirements for teleoperation interaction tasks, with a specific focus on remote scene transmission and robotic arm control in teleoperation scenarios. Secondly, according to the system's functional requirements, extensive research was conducted on real-time rendering techniques in mixed reality environments, and an interactive control strategy of the robotic arm was developed for mixed reality environments. This effort enabled a fusion effect between virtual and physical scenes by mapping holographic objects to the real-world environment. This supports the establishment of the mixed reality visual augmentation and interaction system that allows op-

erators to visualize the mixed reality terminal and facilitates seamless end-to-end network communication. Experimental results show that the interaction system has better performance than a traditional system, as manifested by a remarkable 17.3% reduction in task completion time and a notable 7% improvement in accuracy. These findings suggest the promising application prospects of the system.

Keywords: mixed reality; teleoperation; visual augmentation; robotic system; interaction design

随着自动化控制、自然人机交互、信息通信等技术的快速发展,面向遥操作环境的高度智能化与自主化的机器人系统在高风险任务中发挥着越来越重要的作用^[1]。同时,遥操作技术也被广泛应用于排雷排爆^[2]、搜索救援^[3]、高空作业^[4]、水下探索^[5]和太空探索^[6-7]等极端复杂环境中。在这些应用领域里,机器人往往要代替人们进入危险环境,因此需要重点研究的是该如何赋予机器人以近似人类的感官系统和认知能力,从而能够直观地在操作员端再现远端场景^[8]。

研究表明,在人类的 5 种感官系统中,视觉的神经元数量以及信息传输率要远高于其他感觉模式^[9]。从某些程度上来讲,人类大脑的一半以上直接或间接地用于处理视觉影像,视觉信息的处理又支配着来自其他模态信息的处理,由此可见视觉信息的正确反馈对于人类的重要性^[10]。然而,目前大多数遥操作系统都是通过 2D 显示器提供视觉模态的反馈,在这种方式的遥操作过程中,存在着大量的用户操作失误以及设备自身的人机交互问题^[11]。

混合现实技术是一种将视觉作为最重要的输入通道之一的交互技术,多年以来,混合现实中的人机交互研究一直在探索更加自然和高效的交互方式^[12-13],旨在通过其立体、多维的场景输入^[14]改善传统系统中的人机矛盾。目前,混合现实技术已经开始应用于远程医疗手术^[15]、太空舱外作业^[16]以及排雷排爆等较为特殊的使用场景。在这些应用场景中,用户均可以借助空间信息和直观的交互技术(如手势),来实现交互数据可视化或多模式遥操作,并可以实时获取真实操作状态信息,帮助自身保持高度的情景意识。因此,混合现实技术在提升操作系统的人机交互与协作效率方面具有巨大潜力。针对遥操作机器人这一场景,Su 等提出了一种基于混合现实视觉和触觉反馈方法的焊接方案^[17],并用于遥控机器人焊接。该方案由机器人工作区域进行混合现实视觉反馈,同时通过混合触觉约束,引导操作员的手部沿混合现实标注导向移动,并将焊接操作限制在无碰撞区域内。这种允许用户进行沉浸式交互的设计方案,具有以用户为中心、将观察视角和态势

感知相结合的特点,可以使操作系统快速理解人类指令,丰富了人机协作的应用场景及范围^[18-19]。除此以外,华南理工大学的研究团队利用 Leap Motion 体感控制器对操作者的手部动作进行检测^[20],并将其转化为机器人的控制指令,通过混合现实技术对机器人的作业场景进行实时的视觉反馈,以确保操作的安全性和效率,提高了人机交互体验^[21]。Livatino 等基于混合现实技术设计了一种组织信息呈现方法和一套视觉辅助工具,通过在用户界面中使用混合现实和视觉辅助工具,向用户呈现传感器数据,以促进远程操作控制面板中数据的视觉通信,不仅提高了操作员态势感知能力,还加快了决策速度^[22]。

上述研究现状表明,混合现实技术在遥操作交互领域具有很大的应用价值,但现有研究往往单独着眼于改进系统的某一个方向,如视觉反馈、交互反馈等;在进行系统整体设计时,仍没有脱离传统系统的设计思路;在进行交互任务时,由于各部分的混合现实程度不一,使用效率并没有得到较大的提升;而且,混合现实场景下的临场感往往比较低,交互性能不高。因此,本文基于以上研究内容与技术背景,面向遥操作应用场景,利用混合现实技术在人体视觉认知以及人机协作方面的优势,设计了一种视觉增强交互系统架构,以满足遥操作过程中操作员的交互需求;然后,采用摄像机到 HoloLens(一种混合现实显示设备)的端到端视频传输方法,实现了远程视景的实时呈现;通过研究混合现实场景下遥操作主端的交互流程与技术实现,建立了基于循环坐标下降逆运动学法的机械臂虚拟交互模型,通过虚实融合技术为遥操作机器人提供了一种可视化效果较好的交互控制方案,解决了目前主从式遥操作系统缺乏临场感、交互性能差的问题,对系统的安全性、稳定性起到了很大的促进作用。

1 混合现实视觉增强交互系统设计

1.1 混合现实遥操作交互需求分析

当前的遥操作系统,尤其是排雷机器人系统(本文主要探讨这一背景),操作员通常是通过显示屏或

传统屏幕(如移动设备、平板电脑或笔记本电脑等)观看机器人端摄像头返回的实时视频流。在这种情况下,操作员只能看到二维的图像信息,明显缺乏临场感知^[23]。另外,传统的排雷系统中使用的摇杆、手柄、键鼠等遥控设备限制了交互的空间,通常情况下,远端机器人的控制操作与视觉反馈两个模块呈现相互独立的态势。在实际作业与交互过程中,操作员不仅要实时关注机器人端摄像头画面,还需要实时获取机器人的运动姿态并进行远程操作,即操作员必须在视觉反馈区域和控制区域之间不断进行切换,这样不仅会导致注意力分散,而且是不自然和低效的^[24-25]。基于上述问题,本文重点研究了如何利用混合现实技术提供一种新型的沉浸式视觉反馈与交互系统,以解决现有系统视点跳跃、交互区域不统一的问题。

从遥操作作业存在的问题和实际需求出发,可以总结出本系统的功能需求如下。

(1)远程视景实时呈现:用户的首要需求是通过实时视频流信息来感知远端环境,因此该系统的主要功能就是采集与传输实时画面,且实时地呈现在用户面前。

(2)人机自然交互:在人机交互过程中,远程机械臂的交互控制需要纳入混合现实交互场景中,以实现视线交互区域与遥控操作交互区域相统一。

(3)虚实融合:在遥工作场景中,将混合现实全息对象和真实的环境元素结合在一起,从而增强交互场景的视觉体验。

非功能需求主要针对软、硬件环境与应用场景,具体包括如下内容。

(1)系统开发平台:Windows 平台,不需要大量的存储空间,系统尽量小巧。

(2)用户需求:用户需具备基本的计算机作业经验与能力,且在不熟悉计算机操作和混合现实系统的条件下,能做到快速学习并很好地适应。

(3)环境条件:应当尽量避免强光,以适用于设备采集到合适亮度的视频为准。混合现实显示设备也同样不适用于强光环境。

总体来说,通过混合现实技术改变传统的第三人称视角观察的多视频流传输,有助于理解操作的上下文。此外,交互系统能为用户提供直观、立体的混合现实机械臂交互控制方式,使其更加身临其境并得到自然的交互体验。

1.2 视觉增强交互系统总体架构

综合以上需求分析,本文提出了一种新的视觉

增强交互系统框架,采用“视频视角”的混合现实显示方案,将机器人端返回的视频信息视为真实世界,将混合现实场景中的全息对象用于远程机器人控制,并在数字化视频流上完成混合现实的工作。该方法将视觉反馈区域与控制区域集成一体,最终呈现出一套完整的基于混合现实的闭环遥操作交互系统。此系统从三维显示和自然交互的角度出发,在为操作员提供实时、准确、有效的作业场景信息的同时,实现了立体直观的交互控制指令反馈。所设计的系统结构框架如图 1 所示,主要分为主动端的操作员站和从动端的机器人作业站,二者通过无线自组网设备实现远程通信。

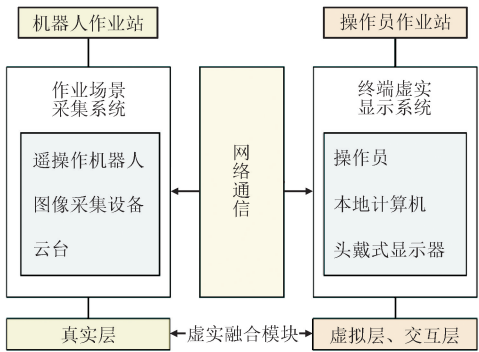


图 1 遥操作视觉增强交互系统框架
Fig. 1 Teleoperation vision enhancement system frame

主动端为操作员、头戴式显示器、本地计算机等组成的操作员站。在本地计算机上建立一个混合现实场景,该场景包括立体视觉显示和人机交互界面,可以直接给操作员提供立体可视化的远端场景和交互界面,操作员通过头戴式显示器,在安全环境中实时、准确地观察机器人姿态并完成交互控制任务。从动端为遥操作机器人、图像采集设备、云台等组成的机器人作业站。主动端和从动端通过自组网设备进行通信,自组网设备内嵌有视频流传输协议,能够把摄像头采集到的图像发送回主动端,在作业任务时提供辅助参考给操作员。交互系统按功能可分为真实层、虚拟层和交互层,其中也包括了作业场景采集系统、终端虚实显示系统、跟踪注册模块、虚实融合模块等,其工作步骤如下。

步骤 1 真实层的作业场景采集系统首先读取摄像机采集到的实时流协议(RTSP)视频流,再对其进行编解码,从而获得实时图像信息。

步骤 2 通过计算机终端的虚实显示系统,完成虚拟图像渲染与栅格化,生成虚拟层的虚拟信息。

步骤 3 跟踪注册模块完成对摄像机视野和视角的跟踪,生成相关的位置、坐标等信息,同时将虚拟层的数字信息与真实环境相匹配,实现虚拟层与真实层的融合。

步骤 4 将整合后的信息输出至头戴式显示器,为操作人员提供视觉反馈和机器人控制界面,然后操作人员与虚拟模型进行交互,完成遥操作任务并发出控制指令。

2 关键交互技术研究

2.1 远程视景实时呈现

混合现实交互系统首先需要实现的一个功能就是通过摄像机捕捉或记录近端任务环境,与此同时,远程用户在混合现实头盔(HMD)上获得近乎实时的视频流。摄像机-HoloLens 端到端视频传输实现方法如图 2 所示,首先设置流媒体通道,用于发布视频输出、编码视频和发布视频流,然后在全息应用程序中将这个视频流通道部署至 HoloLens。

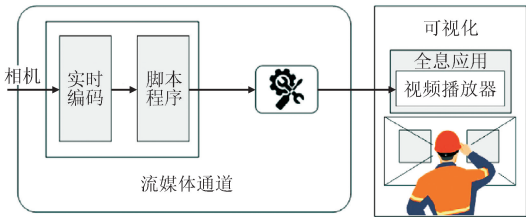


图 2 摄像机-HoloLens 端到端视频传输实现方法
Fig. 2 End-to-end video delivery implementation methods

图像传输的流程如图 3 所示,可通过自定义服务器端和远程客户端传输与接收的数据类型,提高网络通信过程中的稳定性。

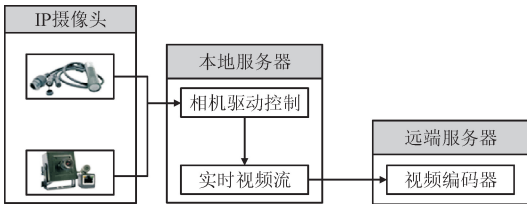


图 3 图像传输流程
Fig. 3 Image transfer flowchart

采用 2 个网络协议(IP)摄像头采集远端场景,分别获取远端整体抓取场景的全景图像以及手部抓取的特写图像。在图像数据的无线传输过程中,图像帧率为 30 Hz,分辨率为 1 920×1 080 像素。另外,为了减小数据传输量,在获取实时图像后将数据压缩为 JPEG 格式,并根据自定义的传输类型进行

打包发送。服务器端程序流程如图 4 所示,远程客户端负责接收图像数据,并将其解包、解码后在混合现实场景中实时显示。

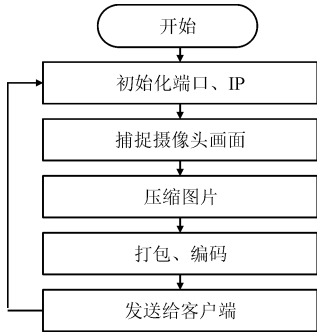


图 4 服务器端程序设计流程
Fig. 4 Server-side programming

本文基于 Unity3D 软件实现实时视频播放功能,但由于其原生视频播放器仅支持本地视频文件的播放,因此为了将外部世界的 IP 摄像头输入到开发环境中,使用 Universal Media Player 插件搭建了一个可以播放 RTSP 实时视频流的场景,用脚本控制 RTSP 视频流的播放,设计并开发了相应的程序接口与软件界面,如图 5 所示。

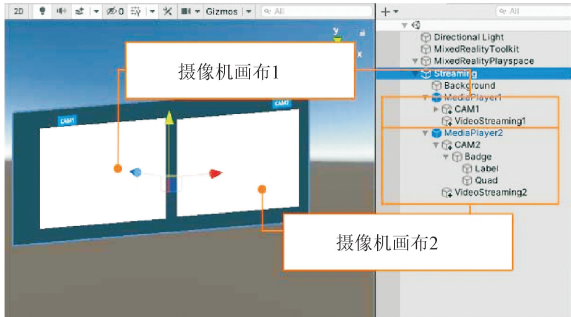


图 5 Unity3D 软件中的场景架构
Fig. 5 Scene architecture in Unity3D

该场景包括 2 个并列排布的 2D 画布(canvas),分别用于渲染显示主视角相机(CAM1)和手部视角相机(CAM2)的实时视频流,并通过背景面板有效地避免强光环境中的光线干扰。随后设置画布的层次结构,挂载相应的脚本文件,直接调用 RTSP 视频流,将画面渲染到对应的画布上,实现实时视频流的接收与播放。运行该场景程序时,会在本地计算机打开一个提供网络摄像头画面的新窗口,以输出视频编码器正在接收图像数据的状态信息。在混合现实显示器中,为了生成高质量且稳定的混合现实场景,通过无线串流在 Unity3D 软件中实现 HoloLens 与 PC 电脑在同一局域网内的无线连接,从而

在 HoloLens 中同步 Unity 中的输出结果,如图 6 所示。针对此类二维信息的注册,只需确定其相对于真实空间的坐标信息,就能较好地实现信息叠加。在这个过程中,PC 计算机作为后台服务器进行数据处理,HoloLens作为客户端向用户提供操作环境。



图 6 Unity3D 软件中的视频流显示

Fig. 6 Video stream display in Unity3D

2.2 Unity 场景下的机械臂自适应控制

本文所提出的混合现实交互系统面向遥操作作业环境,需要根据真实环境的物理数据创建遥操作平台虚拟交互模型,从而实现对遥操作机器人作业场景中各要素的构建与实际控制。以遥操作机械臂为例,由于需要通过对机械臂模型各关节角度的实时接收发控制远端的真实机械臂,因此首先构建机械臂模型,具体流程如图 7 所示。首先利用 SolidWorks 软件对机械臂进行几何重建,确定各零部件之间的装配关系与旋转中心,然后考虑到形状、材质、纹理等因素的影响,利用计算机图形学及相关渲染算法进行模型着色,以提高混合现实三维立体显示的质量。



图 7 机械臂虚拟模型构建流程

Fig. 7 Robotic arm virtual model building process

由于 Unity3D 软件所支持的模型文件为 FBX 格式,所以必须先将 SolidWorks 软件生成的模型导入 3DsMAX 软件进行转化,再将转化后的文件导入 Unity3D 软件,建立对应模型的材质球以刻画模型的细节特征,从而提高混合现实视觉质量。此外,通过降低虚拟模型的真实感,保证了系统的实时性需求,即在进行建模与实时渲染时,对模型的精确度和绘制速度进行了权衡,既能使画质得到保证又不会给用户带来任何视觉观感上的不适。模型生成后需要贴上纹理,使其更具有真实感。三维模型贴图的

核心是建立纹理贴图与模型表面的映射关系,其实现过程如图 8 所示。首先计算纹理在三维模型空间坐标系中的空间位置,确定纹理空间点 (u, v) 与模型空间坐标的映射关系,然后依据该映射关系,实现三维模型的着色。

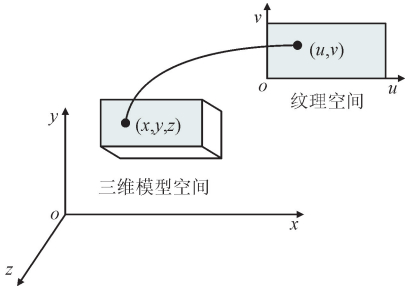


图 8 纹理映射过程示意图

Fig. 8 Texture mapping process

在场景实现过程中,模型建立后需要对机械臂模型进行控制与交互方法的设计。由于采用机械臂仿真 Unity 铰链关节后进行角度输出的传统方案存在无法精确控制角度的问题,以及当项目套用混合现实工具包(MRTK)组件在混合现实场景下运行时,本文所使用的机械臂需要进行 6 轴的关节计算,给系统运行造成了较大的负载,常常会出现系统崩溃、模型显示逻辑混乱的情况,因此,本文提出采用循环坐标下降逆运动学(cyclic coordinate descent inverse kinematics,CCDIK)方法实现混合现实场景下机械臂的交互。实现机械臂交互控制的框架如图 9 所示。

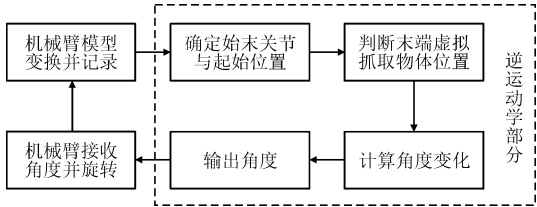


图 9 基于末端抓取物的机械臂交互控制框架

Fig. 9 Robotic arm interactive control frame based on end-grabbing objects

本文基于 Unity2020. 3. 41f1c1 版本,利用 CCDIK 算法在混合现实场景中实现 Unity 机械臂的自适应交互。CCDIK 算法是一种较为快速的逆运动学求解算法,在实时性要求较高的环境中运用比较广泛,且由于这种算法的计算量比较小,对系统的负载影响也较小,因此适合在混合现实设备算力较弱的情况下使用。该算法的具体过程为:首先在

Unity 中创建一个机械臂模型,包括机械臂的各个关节和连接器,并设置好每个关节的转动范围和初始角度;然后,搭建 CCDIK 脚本来构建机械臂的逆运动学模型,该脚本可以根据机械臂的目标位置和姿态,计算出每个关节需要旋转的角度,从而实现机械臂末端的精确定位;最后,将 CCDIK 脚本挂载到机械臂模型上,设置目标位置和姿态,并进行运动控制。本文通过编写脚本来控制机械臂的运动,使用手部姿势来操作目标抓取物体,通过在更新进程中进行不间断的坐标判断,向脚本不断输出目标位置,实现连续、实时的机械臂姿态变换。随后,进一步对机械臂模型进行优化和改进,添加了碰撞检测以提高模型的精度和稳定性。

2.3 三维注册与融合流程

混合现实系统的重要部分是将虚拟的全息对象与真实环境融合起来。本文所设计的混合现实场景中的全息对象包括两大类,如图 10 所示,一类是 2D 对象如字体、图片、UI 等,用于向用户展示当前环境的动态信息及相关属性;另一类是 3D 对象,主要是指三维全息对象如机械臂、抓取物、小车等虚拟模型,用于远程机器人系统的交互控制,以增强用户的沉浸感体验。

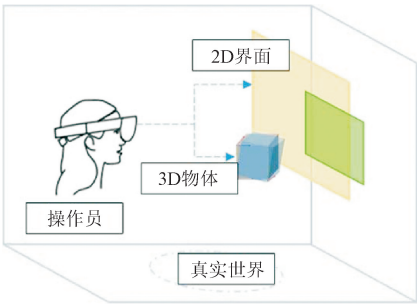


图 10 混合现实场景中的虚拟信息

Fig. 10 Virtual information in a mixed reality scenario

采用基于自然特征点的无标识三维注册方法,对这些全息对象进行空间三维注册,利用 Vuforia SDK 工具包和 Unity3D 软件开发对应的 Unity 脚本,实现对全息对象空间位置和姿态的实时定义。该三维注册系统由 4 个部分组成,分别是三维目标预设、图像实时处理、三维注册以及图像输出。实现的基本流程如图 11 所示。

通过选取多张标记图像建立目标样本数据库,并将数据库导入到 Unity 项目中,可实现增强现实标记的调用。通过绑定模型与目标物图像,根据 Unity3D 软件中空间参考坐标系与真实空间坐标系

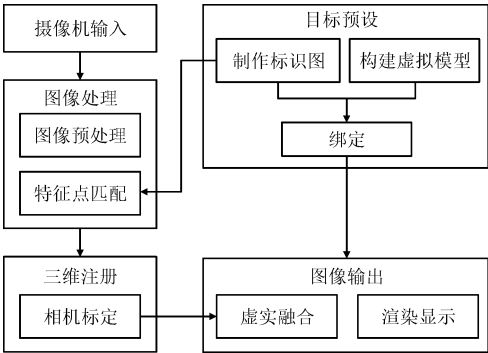


图 11 三维注册与融合流程

Fig. 11 3D registration and fusion flowchart

的关系,设定目标物与虚拟模型的相对位置关系,可实现虚实融合与显示。采用 HoloLens 设备进行虚实融合的过程如图 12 所示。由图可见,该三维注册方法能将全息对象和现实所观察到的遥操作场景相融合,使全息影像在相机识别到目标对象时,可以按照预先设置的相对关系在混合现实场景中显示,增强了全息图像的真实感,极大地改善了物理环境与虚拟环境的虚实融合效果。基于此方法,系统提供了可以进行机器人操作的全向感知平台,允许操作员在作业过程中与虚拟机器人系统进行交互,并能从各个方向观察虚拟机器人与虚拟抓取场景,有助于操作员直接感受远端环境并进行遥操作控制,以获得更好的真实感和沉浸感。

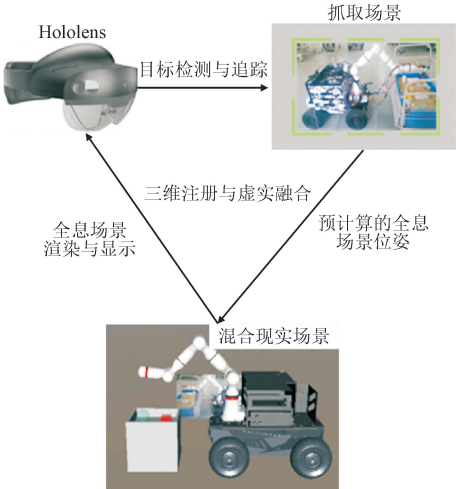


图 12 虚实融合实现过程

Fig. 12 Virtual and real fusion effect

3 交互系统构建

3.1 交互界面可视化

在 Unity3D 软件中进行系统交互界面的构建,作为遥操作视觉增强交互系统的重要组成部分,主

要包括利用远程摄像机在混合现实场景中获取传输的实时视频流,以及拖动三维可视化后的虚拟机械臂进行控制两部分内容,因此可将虚拟交互空间进行划分,并根据相应的交互功能进行布局。如图 13 所示,左、右两侧展示了遥操作实时视频流画面,分别用于显示遥操作端的手部实景和遥操作端的全局实景。AR 识别平台位于视野正前方,用于实时追踪并定位虚拟机械臂控制的交互平台。用户通过观察立体传输的视频流信息,变换机械臂的虚拟模型位姿以进行远端机械臂的同步控制。交互模式选择模块集合了性能分析器、手势交互方式选择、AR 平台关闭与显示、场景切换等功能按键,分别响应于对应的特征。该界面用户最终通过 HoloLens 设备进行显示,辅助遥操作系统实现了自然人机交互与沉浸式作业。

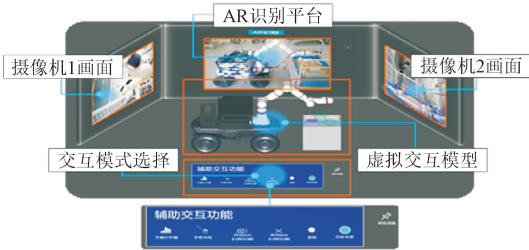


图 13 混合现实系统界面

Fig. 13 Hybrid reality system interface

交互系统中使用的界面标定方式为环境锁定,但为了适应不同应用场景以及用户的个性化需求,在实际应用过程中也提供了基于头部追踪技术的界面标定方式供操作员选择,此方式能有效地避免由于用户移动造成的画面消失或频繁的手动定位。

3.2 系统通信实现

本系统网络通信由两套独立的子系统实现,分别为相机端-主机端视频流的通信实现以及主动端-从动端交互指令的通信实现。通过不同的数据传输协议,从而实现主从端各子系统之间的信息、控制数据流转。

本系统通过传输控制协议(transmission control protocol,TCP)通信实现主从端点云数据流的传输,其方法为在从端的 Linux 主机上搭建服务端,在主端的 Unity 程序中搭建客户端并编写对应的脚本文件。TCP 是一种面向连接的、可靠的、基于字节流的传输层通信协议,其通过建立稳定连接、按顺序传输数据包以及提供错误检测和重传机制,确保了数据的完整性和可靠性。该协议在本系统的具体

应用如图 14 所示。

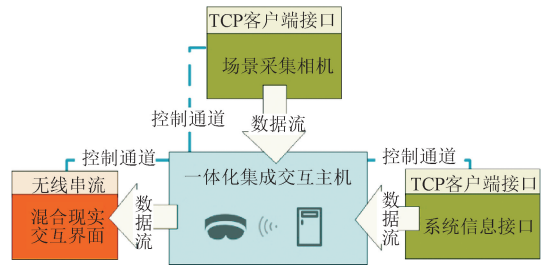


图 14 场景传输 TCP 通信流程

Fig. 14 Scenario transmission TCP communication process

基于 Unity 的自适应机械臂实现交互控制后,需要进一步搭建局域网传输机械臂的远程控制指令,具体的通信流程设计如图 15 所示。由于模型机械臂与真实机械臂的初始位姿不一定同步,因此在系统开启之前需要先进行初始化同步以保证两者的位姿相同,再通过用户数据报协议(UDP)通信使 Unity 工程与远端机器人进行远程数据交换并发送位置指令。

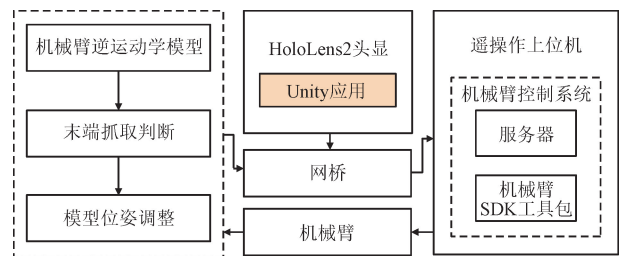


图 15 混合现实机械臂控制通信流程

Fig. 15 Mixed reality robotic arm control communication process

为了控制机械臂的运动,需要获取机械臂关节角度的变化信息。这些信息可以从 Unity 工程中提取,在进行位姿同步后,机械臂模型接收末端位姿的坐标并进行实时的姿态变换计算,随后将变换后的角度信息通过 UDP 通信传输给远端机器人。

在该 Unity 通信进程中,每一帧即 $1/12\text{ s}$ 都会产生一次角度数据,因此在短时间内产生了大量的、重复的角度数据,并被机械臂服务器端全部接收,这对机械臂的角度处理造成了较大的干扰,频繁出现单次姿态变换中移动卡顿的问题。基于此问题,本文通过物体抓取操作的监听进行数据传输优化,由于系统构建中采用 MRTK 的交互脚本实现了对虚拟抓取物的交互操作,因此使用 MRTK 的可交互组件进行物体抓取操作时,可以通过监听手势交互终止事件(manipulation ended)来实现每次抓取结束后输出

一次角度信息的功能。实现的最终效果为在每次抓取操作结束后,才将旋转信息发送到服务器,并输出角度信息,极大地优化了机械臂的交互控制效果。

4 系统验证

4.1 遥操作场景下传输延时实验

在混合现实视觉增强交互系统的实际应用中,遥操作主从端的传输延时是衡量系统整体性能的一个关键指标。一般情况下,系统延迟响应约为 70~110 ms,一旦超过 300 ms,就会直接影响到操作者的作业效率。设 Δt_1 是相机从新的位置采集图像并返回显示器所造成的延迟,主要受由网络传输性能的影响,并随着传输图像尺寸的增大而增大, Δt_2 是在 HMD 上渲染获取的图像所造成的延迟,因此系统总延迟 Δt 为 Δt_1 与 Δt_2 之和。实验设定视频以 1 920×1 080 像素高清分辨率进行传输,使用 MJPG 编码,帧率为 30 Hz,传输图像尺寸为 640×480 像素,服务器和客户端在同一局域网中。设置主从端距离为 15 m,在摄像机前执行固定动作直到该动作在交互系统中复现后,计算执行动作与复现动作之间的时间差即系统延时。

实验得到的结果如图 16 所示,统计结果表明,图像传输过程中的延时保持在 210 ms 到 300 ms 之间,平均值为 260 ms,方差为 29 ms,能够满足实时性要求。

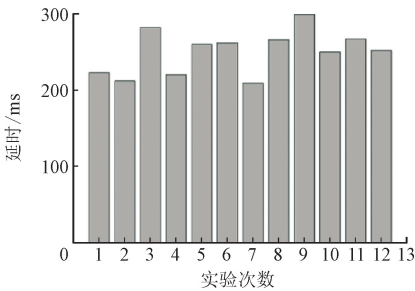


图 16 主从端距离为 15 m 时图像传输延时实验结果
Fig. 16 Experimental results of image transmission delay when the master-slave distance is 15 m

4.2 人机交互功能测试

对本文所设计的遥操作交互系统进行人机交互功能测试,测试结果主要用于评估系统的手势识别与交互性能。实验随机招募 14 名被试者(10 名男性,4 名女性),年龄范围为 21~26 岁(平均为 23.79 岁,标准差为 1.42),均无 VR、AR 使用经验。被试者进行基础学习与训练后,执行手势 30 次/命令。手势由 5 个命令组成,分别为选择虚拟对象、调整对

象位置、调整对象尺寸、机械臂姿态变换和多通道选择交互,要求被试者完成混合现实环境下的关于系统相关模块的基本交互任务。实验记录下检测到的手势数,并将统计结果列于表 1。

由表 1 可见,该系统对手势命令的检测准确率较高,交互时间随着交互任务的复杂度增加。实验过程中帧率总体波动范围为 55~75 Hz,最大差值保持在 10 Hz 以内。

表 1 人机交互功能测试任务及结果统计
Table 1 Test tasks and results statistics of human-machine interaction

测试任务	帧率/Hz	平均时间/s	交互成功率/%
选择虚拟对象	68~75	1.25	100
调整对象位置	63~65	3.20	100
调整对象尺寸	55~65	5.60	100
机械臂姿态变换	57~65	4.85	93
多通道选择交互	68~75	3.60	100

综上所述,所设计的混合现实视觉增强交互系统,能够满足基本的人机交互功能需求,并且具有良好的运行性能与流畅性。同时,通过与实验搭建的利用二维显示器进行场景传输的传统遥操作系统相比,在进行完整的遥操作抓取任务中,同样完成选择对象、机械臂变换以及抓取对象这 3 组动作,传统二维系统完成任务所需的平均时间为 11.73 s,准确度为 0.87,而本文所设计的遥操作交互系统在任务平均完成时间上减少了 17.3%,准确度提高了 7%。

5 结 论

本文针对遥操作机器人系统在高风险任务中近端交互临场感差、操作者难以获得真实体验感的问题,提出了一种适用于遥操作场景的混合现实视觉增强交互系统,并对该交互系统的交互性能进行分析和实验验证,探讨了混合现实交互技术的应用对于遥操作机器人系统人机交互性能的影响,得到以下主要结论。

- (1)以人机交互通道为研究对象,从多通道信息的整合入手,研究了 Unity3D 平台下视频流的采集与流式无线传输方法,完成了在硬件平台上的搭载并进行了传输测试,有效地完成了远程场景还原的系统需求。
- (2)基于机械臂的逆运动学构建了混合现实场景下的机械臂控制脚本,设计了一种立体可视化的遥操作控制流程,最大程度地保障了用户以自然的

方式与控制对象交互,降低了学习成本。

(3)基于视觉感知以及可视化交互理论构建系统交互界面,提出了二维信息与三维模型相结合的混合式系统信息构建方法。该方法将视觉反馈区域与控制区域集成一体,改善了遥操作机器人系统的复杂度和稳定性。

本文所提出的混合现实视觉增强方法遥操作交互系统,有效地解决了远端作业场景的同步再现问题,提高了操作员的作业效率和准确率,对于军事领域中的排雷排爆等遥操作作业具有一定的应用价值。

参考文献:

- [1] 张宇,李宇翔,张贺飞,等.撼地者:基于遥操作与自主导航的移动救援机器人[J].中国科学技术大学学报,2023,53(1):37-49,71-72.
ZHANG Yu, LI Yuxiang, ZHANG Hefei, et al. Earthshaker: a mobile rescue robot for emergencies and disasters through teleoperation and autonomous navigation [J]. Journal of University of Science and Technology of China, 2023, 53(1): 37-49, 71-72.
- [2] 曾建军.排爆机器人遥操作性能研究[D].上海:上海交通大学,2010.
- [3] SUN Zeyuan, YANG Hong, MA Yifan, et al. BIT-DMR: a humanoid dual-arm mobile robot for complex rescue operations [J]. IEEE Robotics and Automation Letters, 2022, 7(2): 802-809.
- [4] 龚朋朋,李笑.遥操作工程机器人双目视觉定位[J].机械设计与制造,2016(9):120-123.
GONG Pengpeng, LI Xiao. Binocular vision location of Tele-Operated construction robot [J]. Machinery Design & Manufacture, 2016(9): 120-123.
- [5] 张鸿健,韩捷,蒋维涛,等.柔性水下悬停机器人的液气相变驱动及制造[J].西安交通大学学报,2021,55(2):10-17.
ZHANG Hongjian, HAN Jie, JIANG Weitao, et al. Liquid-gas phase change driving and manufacturing of flexible underwater hovering robot [J]. Journal of Xi'an Jiaotong University, 2021, 55(2): 10-17.
- [6] CHEN Haifei, XU Yong. High telepresence control for the relay communication based space teleoperation [J]. Acta Astronautica, 2023, 208: 96-104.
- [7] 程瑞洲,黄攀峰,刘正雄,等.一种面向在轨服务的空间遥操作人机交互方法[J].宇航学报,2021,42(9):1187-1196.
CHENG Ruizhou, HUANG Panfeng, LIU Zhengxiong, et al. A human-robot interaction method of on-

orbit service-oriented space teleoperation [J]. Journal of Astronautics, 2021, 42(9): 1187-1196.

- [8] WHITNEY D, YAMANASHI LEIB A. Ensemble perception [J]. Annual Review of Psychology, 2018, 69(1): 105-129.
- [9] KIM G, JANG J, BAEK S, et al. Visual number sense in untrained deep neural networks [J]. Science advances, 2021, 7(1): 6127.
- [10] GIBBS J K, GILLIES M, PAN Xueni. A comparison of the effects of haptic and visual feedback on presence in virtual reality [J]. International Journal of Human-Computer Studies, 2022, 157: 102717.
- [11] BARUA H B, SAU A, DEV ROYCHOUDHURY R. A perspective on robotic telepresence and teleoperation using cognition: are we there yet? [EB/OL]. (2022-03-06) [2022-05-20]. <https://arxiv.org/abs/2203.02959>.
- [12] 陶建华,巫英才,喻纯,等.多模态人机交互综述[J].中国图象图形学报,2022,27(6):1956-1987.
TAO Jianhua, WU Yingcai, YU Chun, et al. A survey on multi-modal human-computer interaction [J]. Journal of Image and Graphics, 2022, 27(6): 1956-1987.
- [13] AZOFEIFA J D, NOGUEZ J, RUIZ S, et al. Systematic review of multimodal human-computer interaction [J]. Informatics, 2022, 9(1): 13.
- [14] LI Ke, SCHMIDT S, BACHER R, et al. Mixed reality tunneling effects for stereoscopic untethered video-see-through head-mounted displays [C]//2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). Piscataway, NJ, USA: IEEE, 2022: 44-53.
- [15] COELHO L P, FREITAS I, KAMINSKA D U, et al. Virtual and augmented reality awareness tools for universal design: towards active preventive healthcare [M]//Emerging Advancements for Virtual and Augmented Reality in Healthcare. Hershey, PA, USA: IGI Global, 2022: 11-24.
- [16] TORRENCE B, DRESSEL J. Critical review of extended reality applications in aviation [C]//Virtual, Augmented and Mixed Reality: Applications in Education, Aviation and Industry. Cham: Springer International Publishing, 2022: 270-288.
- [17] SU Yunpeng, CHEN Xiaoqi, ZHOU T, et al. Mixed reality-enhanced intuitive teleoperation with hybrid virtual fixtures for intelligent robotic welding [J]. Applied Sciences, 2021, 11(23): 11280.
- [18] AGUAYO C, EAMES C. Using mixed reality (XR)

immersive learning to enhance environmental education [J]. The Journal of Environmental Education, 2023, 54(1): 58-71.

[19] AYYANCHIRA A, MAHFOUD E, WANG Weichao, et al. Toward cross-platform immersive visualization for indoor navigation and collaboration with augmented reality [J]. Journal of Visualization, 2022, 25(6): 1249-1266.

[20] GALLALA A, KUMAR A A, HICHRI B, et al. Digital twin for human-robot interactions by means of industry 4. 0 enabling technologies [J]. Sensors, 2022, 22(13): 4950.

[21] 占宏, 梁聪垣, 杨辰光. 基于混合现实的机器人遥操作实验平台 [J]. 实验技术与管理, 2021, 38(8): 114-117.

ZHAN Hong, LIANG Congyuan, YANG Chenguang. Robot teleoperation experimental platform based on mixed reality [J]. Experimental Technology and Management, 2021, 38(8): 114-117.

[22] LIVATINO S, GUASTELLA D C, MUSCATO G, et al. Intuitive robot teleoperation through multi-sensor informed mixed reality visual aids [J]. IEEE Access, 2021, 9: 25795-25808.

[23] ZAMAN F. [DC] Improving multi-user interaction for mixed reality telecollaboration [C]//2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). Piscataway, NJ, USA; IEEE, 2022: 940-941.

[24] LIN T C, KRISHNAN A U, LI Zhi. Comparison of haptic and augmented reality visual cues for assisting tele-manipulation [C]//2022 International Conference on Robotics and Automation (ICRA). Piscataway, NJ, USA; IEEE, 2022: 9309-9316.

[25] PAMPARĂU C, SCHIPOR O A, DANCU A, et al. SAPIENS in XR: operationalizing interaction-attention in extended reality [J/OL]. Virtual Reality, 2023: 1-17 [2022-08-28]. <https://doi.org/10.1007/s10055-023-00776-1>.

(编辑 李慧敏 刘杨)