UNB Off-site Construction Research Centre
Northumbria University NEWCASTLE

Transforming Construction with Off-site Methods and Technologies (TCOT) Conference: Designing Tomorrow's Construction, Today

August 20-22, 2024, Fredericton, New Brunswick, Canada

Development of an Intuitive Mixed Reality Human Robot Interaction Interface for Construction Applications

Tennakoon, D.[1, *], Jadidi, M.[1, *], and RazaviAlavi S.R.[2]
[1] Lassonde School of Engineering, York University, CA
[2] Northumbria University, New Castle, UK
[*] damitht6@my.yorku.ca, mjadidi@yorku.ca

Abstract: The main challenges faced in the construction industry are low productivity rates and the lack of priority given to worker safety. Robotics solutions have been proven to improve task efficiency, product quality, and workplace safety in many industries, though it has posed a challenge to be integrated into the dynamic and unstructured construction environment. This research aims to provide a mode to adopt robotics into the construction industry through an intuitive Human-Robot Interaction (HRI) system that integrates the precision, strength, and deployability of robots while leveraging the decision-making, experience, and creativity of on-site workers. The methodology consists of mounting the ZEDM stereo vision camera, attached on a 2-axis servo actuated gimbal, to the end effector of a 6-axis robot arm. The stereo images are streamed to the Meta Quest 2 virtual reality (VR) head-mounted display (HMD), providing a First-Person View (FPV), with human-like vision, of the end effector. Hand tracking libraries are used to track and map the operator's hands into the FPV view. Localizing the tracked hand models with the end effector enables omnidirectional motion control and gripper controls, via gesture recognition, of the manipulator. In-situ data is spatially mapped to the stereoscopic view, constructing a mixed-reality (MR) HRI interface. The results demonstrate a higher depth and situational awareness of the robot's workspace, with increased efficiency, relative to traditional keyboard teleoperation methods. This MR HRI interface facilitates the use of robotics in construction, enabling workers to safely complete hazardous tasks remotely using HMDs and hand gestures.

Keywords: Mixed Reality; Human Robot Interaction; Robotics; Stereo Vision; Hand Tracking; Construction Robotics

1 INTRODUCTION

The Architecture, Engineering, and Construction (AEC) sector consists of fast paced and dynamic environments that often present on-site worker safety hazards and low productivity rates. The construction industry faces the highest rate of non-fatal injuries, 29.2% higher than any other (Brown, Brooks and Dong 2020).

The application of robotics has proven to increase task efficiency, product quality, and worker safety in many industries. The implementation of industrial robotic manipulators in the manufacturing industry has shown increased productivity rates and product quality. Robotics deployed in the medical sector have shown an increase in minimally invasive surgical procedures due to their precision, compared to human surgeons. Mobile robotic manipulators in agriculture have demonstrated the effectiveness of uninterrupted work through the management of large crop fields (Campilho and Silva 2023). Despite the many benefits of robotics performing a wide variety complex manipulation tasks, there is a lack of robot applications within the construction industry. This is particularly due to the highly unstructured and dynamic environment of construction sites (Kangari 1985). Existing applications of robotics in construction consist of robotic arms fixed-mounted onto mobile manipulators which perform a single task, within a controlled setting. These types of robots pose a safety hazard to on-site workers as the robots are pre-programmed to complete the task without human robot collaboration (HRC) (Brosque, et al. 2020). Aside from pre-programmed robots, many attempts have been made to train robots, for construction tasks, through machine learning techniques, using real-world datasets. One of the limiting factors for this solution has been a lack of quality available training data for construction applications (Zhang, et al. 2023). A more intuitive method to deploy robotics into the AEC sector is through HRC. An intuitive interface enabling robotic manipulation can enable workers to utilize their decision-making, creativity, and problem-solving skills, while taking full advantage of the precision, strength, and safety offered by robots. This can greatly improve task productivity, safety, and labor shortages (Mukherjee, et al. 2022).

Teleoperated robot arms are most useful in situations where complex manipulation tasks must be completed within hostile and inaccessible environments. Hazardous tasks include construction underwater, maintenance of nuclear power plants, or repairs for orbital spacecraft (Fu, et al. 2014). These types of applications have no effective human access, however, still require human operations. Specifically, a task such as welding requires complex human manipulation in environments that are prone to dust, radiation, as well as explosion hazards (Wang, et al. 2020). Hence, there is a growing need for a generalized intuitive human-robot interaction interface. Conventional teleoperated robot arms are composed of a control input segment and a visual feedback segment. The control input segment consists of joysticks, gamepads, keyboards, and/or a mouse to provide end effector controls, in cartesian space, to move the manipulator. This is coupled with real-time visual feedback systems of the robot’s workspace, conventionally through a fixed-mounted monocular (single-lens) camera near the manipulator. The video feed is displayed on 2D screens, providing line-of-sight (LoS) teleoperation. A limiting factor of this mode of teleoperation is a lack of depth perception and situational awareness, which is inherently due to the 2D visual system that prevents immersion and telepresence into the robot’s workspace (Illing, et al. 2020). This results in a misalignment between the user control space and the rudimentary input system which makes completing intricate manipulation tasks extremely challenging and ineffective (Ellis, Adelstein and Welch 2002). One of the immerging solutions to this challenge in robotics is the integration of mixed-reality (MR) technologies, coupled with stereo vision systems, to provide full immersion to the robot’s workspace (Wonsick and Padir 2020).

To understand the concept of MR, it suffices to introduce augmented reality (AR) and virtual reality (VR). AR is the process in overlaying information, via 3D graphics, onto the physical world, with user input systems to interact and manipulate the 3D models within the hybrid environment (Han, et al. 2022). VR takes this concept a step further by completely immersing the user into a virtual environment, utilizing haptic controllers to navigate and interact within the 3D space (Wonsick and Padir 2020). To achieve this type of immersion, head-mounted displays (HMD) are used, which consist of a left and right display, each with a lens, enabling the perception of human-like vision through the generated stereoscopic images. MR fuses these two technologies by enabling the user to view the physical world and the virtual world depending on the user’s preference (Sievers, et al. 2020). Stereo vision is the use of two cameras, that are separated by a known fixed distance, rendering slightly different views of the target scene, enabling for the use of triangulation, based on the disparities of the images, to extract depth information (DANDIL and ÇEVİK 2019). The combination of stereo camera systems with MR hardware enables full immersion into various workspaces, especially advantageous in robotic applications. The coupled system has shown increased depth cues, user performance, and usability, compared to conventional teleoperation modes, in the application of mobile robotics and is continually being adopted for various manipulation scenarios (Luo, et al. 2021); (Wonsick and Padir 2020).

This research project aims to utilize the advantages of such stereo vision systems and MR HMDs, coupled with hand tracking systems, to develop an intuitive hands-free teleoperation interface for a 6-axis robot arm, in the context of hazardous off-site manipulation applications. The outcomes of this research demonstrated an increased operator awareness of the robot’s workspace, higher situational awareness, an increase in operator safety, and an increase in intuitive teleoperation.

2 SYSTEM OVERVIEW

2.1 Architecture

The overarching MR teleoperation interface includes the implementation of multiple systems, software, and hardware. The design approach is to attach a stereo camera system to the end effector of a robot arm, with a near human-like field-of-view (FOV), streamed to a VR headset, to reconstruct human-like perception of the robot's workspace. Utilizing the 4 built-in cameras of the VR HMD and open-source mixed-reality development libraries, real-time hand tracking of the operator is performed in the Unity engine. Further, hand skeleton and joint rendering systems are developed and fused with the stereo images to visualize the tracked hand data in real-time. Gesture recognition algorithms are developed to translate hand motions into executable commands to move the robot arm's end effector in cartesian space. Further, input systems, such as a keyboard controller, are developed to replicate conventional teleoperation modes for the purpose of comparing with the proposed MR teleoperation system. The aforementioned systems are a part of the User Interface segment. A Middleware segment, developed on a single board computer (SBC), is implemented to communicate the data collected from the User Interface segment into command level data to be executed by the robot arm. This system consists of Robot Operating System (ROS) nodes and topics. To bridge the communication between the User Interface segment and the Middleware segment, a UDP/IP (User Datagram Protocol/Internet Protocol) server/client architecture is implemented, in which a ROS node acts as a host server listening for user input from the Unity client. The received data is then published to a ROS topic in which hand tracking and keyboard teleoperation nodes subscribe to and convert into Universal Asynchronous Receive Transmit (UART) commands. This data is then transmitted to the Robot Arm segment that computes and executes inverse kinematics solutions to move the end effector. A high-level description of the overall architecture is shown in Figure 1.

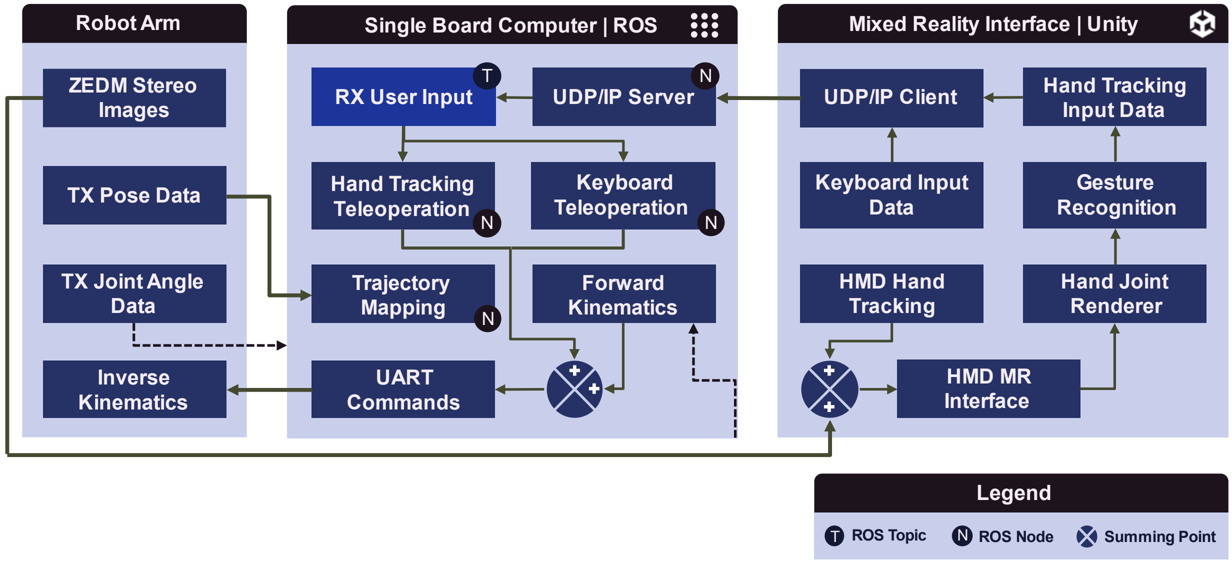
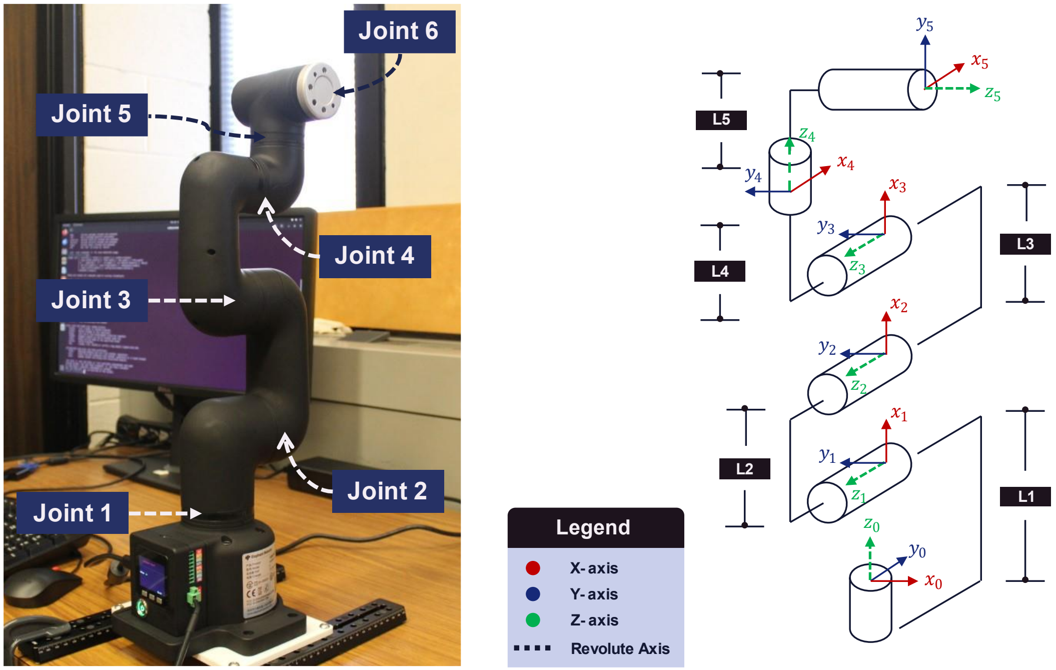


Figure 1: System integration architecture flowchart.

2.2 Hardware Description

The objective of the MR interface is to intuitively teleoperate a robot arm. The manipulator used for this is a 6 degree-of-freedom (DoF) collaborative robot arm from Elephant Robotics called myCobot 320 M5 Stack. This is a highly programmable robot arm widely used for research and development purposes. Some of the features include open-source Python libraries and ROS packages to program the robot in simulation environments, such as Gazebo, and real-world joint space and coordinate space controls. The physical robot arm setup used for this research is described in detail on Figure 2, along with its corresponding robot kinematic chain.



(a) (b)

Figure 2: (a) Description of the robot arm setup. (b) Diagram of the robot kinematic chain.

The MR interface, that is a part to the User Interface segment, is developed in the Unity Universal Render Pipeline (URP) engine on a Lenovo Legion Pro 5i laptop equipped with a 13th Gen Intel Core i5-13500HX CPU, NVDIA GeForce RTX 4050 GPU, and 16 GB of RAM. The HMD used for the MR interface is the Meta Quest 2 VR headset. The SBC used for the Middleware segment is the Raspberry Pi 4b, running ROS2 Humble under the Ubuntu 22.04 operating system. The stereo vision system used for this application is the ZED Mini (ZEDM) by Stereo Labs, operating at a resolution of HD1080, at 30 FPS, with vertical and horizontal FOVs of 40° and 66° respectively.

3 METHODLOGY

3.1 Stereo Vision System

To visualize the robot's workspace in the first-person view (FPV), a bracket was designed to mount the ZEDM camera to the end effector joint of the robot arm. The bracket was designed, considering the weight of the gripper and camera, as well as the 1kg payload capacity of the robot arm. The 3D printed model is suited to fit the camera system and the gripper, then attached to the end effector joint, as shown in Figure 3(b). To visualize the robot's point of view on the VR HMD, the ZED Plugin for unity, developed by StereoLabs, is utilized. The package contains pre-built objects and scripts to stream the left and right images from the camera directly to the Meta Quest 2 HMD, via a USB-C type connection. Further, 3D objects in the Unity editor can be merged with the stereo images to enable passthrough reality, in real-time.

3.2 Hand Tracking

To implement real-time hand tacking, the Oculus Integration package, within Unity, is utilized. The package enables VR rigs to be integrated into the Unity editor, along with options for hand tracking, utilizing the built-in spatial mapping cameras of the HMD. However, the oculus integration package and the ZED Plugin manager are independent of each other and perform two separate tasks; the Oculus Integration package, for the Quest 2, is designed for VR application development while the ZEDM Plugin is designed for AR application development. To achieve the desired MR interface, the hand tracking objects and scripts were fused with the ZEDM rigs, enabling the operator to view the ZEDM stereo images on the HMD along with the tracked hand models, spatially mapped in real-time. In addition to the 3D rendered models of the tracked hands, skeleton and joint rendering scripts were developed to capture a real-time hand skeleton, relative to each of the joints tracked by the built-in cameras. As each joint is represented as a 3D position, a vector relative to the global coordinate system, the relative position between each joint is computed. To detect a simple gesture, such as a pinch, the distance between the tip of the index and thumb joints are computed iteratively. Once a minimum threshold distance has been met, the gesture is recognized. The implemented system, functioning in real-time, is described in Figure 3.


(a) (b) (c)

Figure 3: (a) Image of the operator wearing the HMD and moving their hands. (b) Image of the ZEDM bracket mounted to the end effector of the robot arm. (c) Real-time spatially mapped hand tracking models and joint and skeleton rendering system merged with stereoscopic images, viewed in the HMD displays.

3.3 Mixed-Reality Teleoperation Controller

To utilize the hand tracking system to manipulate the robot arm, a control interface was developed. It consists of a Control Spere, which is a 3D translucent spherical object, spatially mapped a fixed distance away from the centre of the stereo camera system, which is attached to the reference frame of the VR HMD, represented by the position vector $\overrightarrow{\boldsymbol{r_S}}$, and fixed radius R. The objective of the Control Sphere is to provide a reference point in which vector directions can be computed as well as enable the operator to pause/resume control of the robot arm, ascribed in Figure 4. The operator's hand positions and orientations are continuously tracked in the MR application. The position of the palm of the hand is referred to as $\overrightarrow{\boldsymbol{r_H}}$, and is attached to the reference frame of the VR HMD. The relative position vector between the Control Shere and the tracked right hand is computed and is referred to as $\overrightarrow{\boldsymbol{r_{SH}}}$, as described in Equation 1. When the magnitude of the relative position vector, $|\overrightarrow{\boldsymbol{r_{SH}}}|$, is less than or equal to the radius of the Control Sphere, commands are transmitted to pause control of the robot arm. When the magnitude is larger than the radius of the Control Sphere, the relative position vector is normalized, resulting in a unit vector, $\widehat{r_{SH}}$, that is oriented in the direction the operator desires to move the robot arm towards, ascribed by Equation 2. This unit vector is transmitted to the Middleware system in which the current position of the robot arm, $\overrightarrow{\boldsymbol{T_t}}$, at time, t, is incremented by a fixed displacement factor, C, in the desired direction of motion, shown in Equation 3. This updated positional vector, $\overrightarrow{\boldsymbol{T_{t+1}}}$, is converted into executable UART commands and transmitted to the robot arm's processor, in which an inverse kinematic solution is computed to move the robot arm.

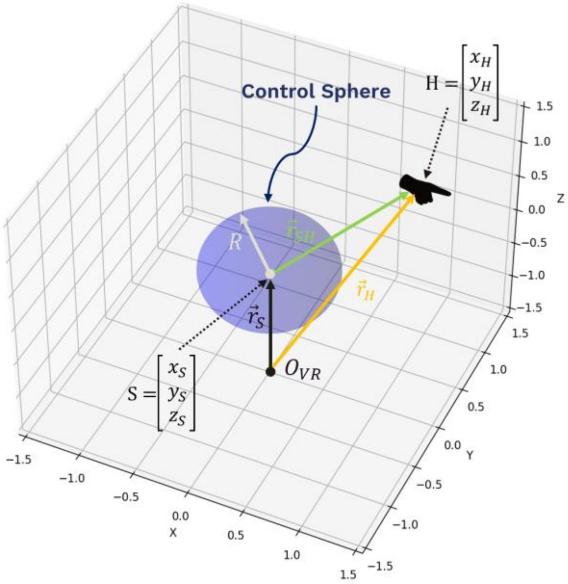


Figure 4: Diagram of control algorithm used in the Mixed Reality interface to convert user hand gesture input into control outputs.

$$\overrightarrow{\boldsymbol{r_{SH}}} = \overrightarrow{\boldsymbol{r_H}} - \overrightarrow{\boldsymbol{r_S}} = \begin{bmatrix} x_H - x_S \\ y_H - y_S \\ z_H - z_S \end{bmatrix} \tag{Eq. 1}$$

$$\widehat{r_{SH}} = \frac{\overrightarrow{\boldsymbol{r_{SH}}}}{|\overrightarrow{r_{SH}}|} = \begin{bmatrix} \dfrac{x_H - x_S}{\sqrt{(x_H - x_S)^2 + (y_H - y_S)^2 + (z_H - z_S)^2}} \\ \dfrac{y_H - y_S}{\sqrt{(x_H - x_S)^2 + (y_H - y_S)^2 + (z_H - z_S)^2}} \\ \dfrac{z_H - z_S}{\sqrt{(x_H - x_S)^2 + (y_H - y_S)^2 + (z_H - z_S)^2}} \end{bmatrix} \tag{Eq. 2}$$

$$\overrightarrow{\boldsymbol{T_{t+1}}} = \overrightarrow{\boldsymbol{T_t}} + C \times \widehat{r_{SH}} = \begin{bmatrix} x_t + C \times \widehat{r_{SH,x}} \\ y_t + C \times \widehat{r_{SH,y}} \\ z_t + C \times \widehat{r_{SH,z}} \end{bmatrix} \tag{Eq. 3}$$

Utilizing the aforementioned control system, the robot arm’s end effector can translate, in real-time, by the operator pointing in the direction they desire the robot arm to move towards. This enables for simplistic and intuitive control of the manipulator’s end effector as the stereo camera feedback system provides a high level of depth awareness while familiar hand gestures enable intuitive motion. The functional MR teleoperation interface is shown in Figure 5, in which an operator controls the robot arm to move to a target cube in the robot’s workspace.

3.4 Experiment Setup

To measure the effectiveness of the developed MR teleoperation system, an experiment was devised in which the teleoperation method was compared with conventional teleoperation methods. The experiment consisted of the robot arm beginning at a fixed stating position with the goal of driving the arm towards a target cube object, fixed at a known location in the robot’s workspace. The experiment was conducted by the same operator in which, first, a keyboard teleoperation method was used to control the robot arm, with the operator sitting near the robot’s workspace, proving LoS (line of sight) control. In the second method, the operator controlled the robot arm using the keyboard controller with a live video feed, from a single lens of the stereo camera mounted to the end effector, displayed on a laptop screen. These two methods represent the most common teleoperation modes for robot arms. Lastly, the operator performed the task using the developed MR HRI interface. For each of the three teleoperation modes, 3 trials were conducted.

In each trial, a ROS node, within the Middleware system, recorded the pose of the robot arm's end effector into a .CSV formatted file. These trajectories were used to analyze the relative depth cues, situational awareness, and intuitiveness of control for each teleoperation methods, as discussed in the next section. The proposed experiment is described in Figure 6, below.

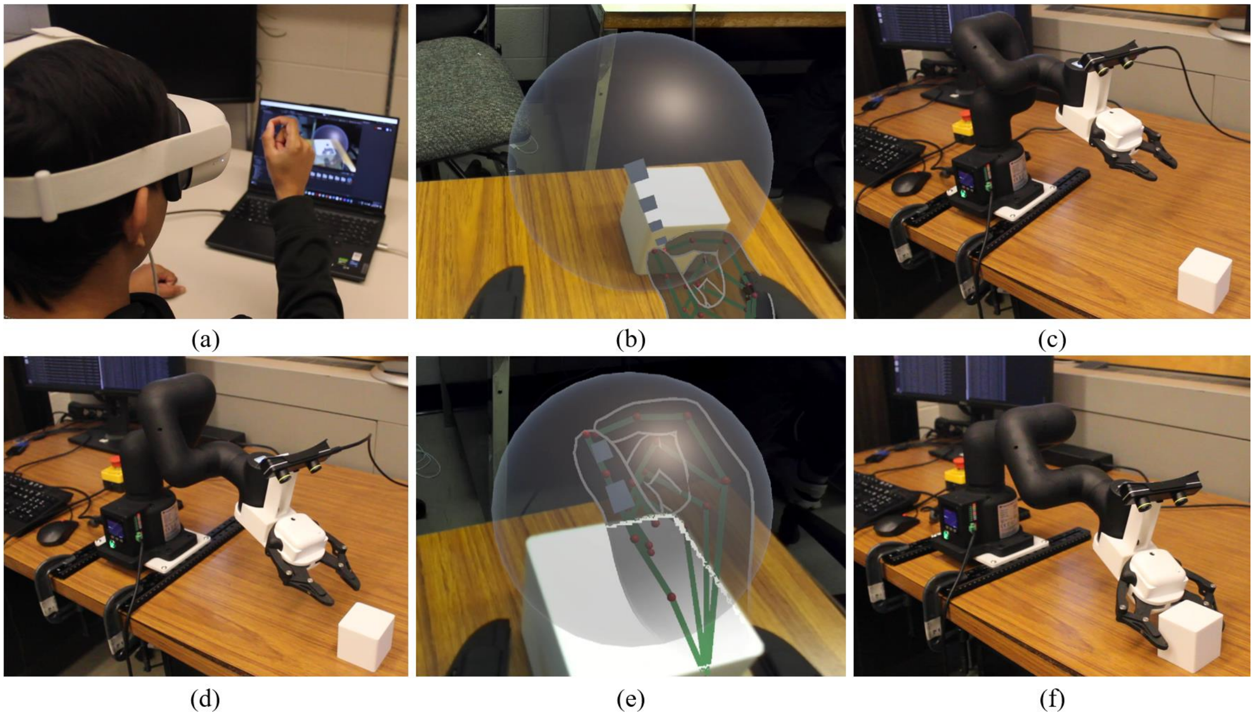


Figure 5: (a) Image of the operator wearing HMD, providing input in the MR interface, using hand gestures. (b) Operator's view as depicted in the VR HMD (FPV of the robot's workspace). (c) Robot arm at starting position. (d) Robot arm moving towards target using the hand tracking data. (e) User placing hand in the Gesture Sphere to pause robot arm control. (f) Robot arm achieving the target position.

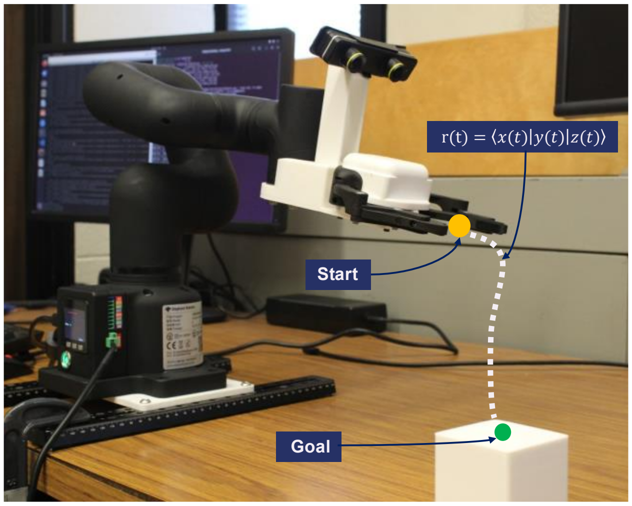
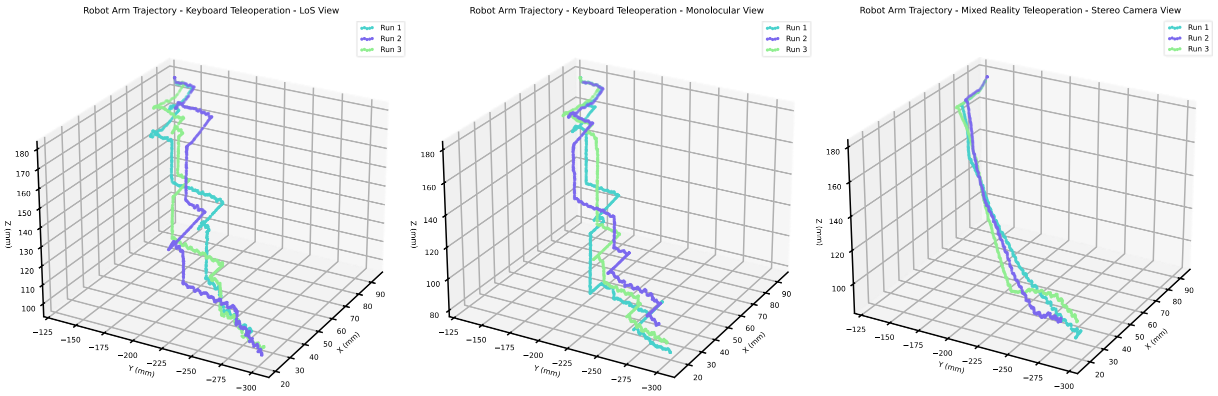


Figure 6: Experiment setup – move in a free path from the start to the goal position. The position and orientation of the end effector is recorded at a fixed interval for analysis.

4 RESULTS

Analyzing the flight paths of the robot arm's end effector using the keyboard LoS teleportation method, ascribed in Figure 7(a), demonstrates irregular trajectories. In the initial stages of the trajectory, the position of the end effector is significantly misaligned with the goal position. Along the way, the operator adjusts the position over larger timesteps. As the end effector converges towards the goal position, more adjustments are made in shorter intervals. This shows that initially, the situational awareness of the operator was low, and as the robot arm minimized the distance, there was an uptick in depth cues. The increased depth cues can be attributed to minimizing the relative distance between the effector and objects in the robot's workspace. Investigating the flight paths of the keyboard teleoperation using the monocular camera system demonstrated similar results to the LoS method. The large corrections that are made over the total flight can be attributed to the lack of depth awareness of the target object due the 2D visual system. As the camera is moved closer to the object, more finer corrections are made. However, the flight paths, for all three runs, are irregular and undesirable. Lastly, analysis of the implemented mixed reality teleoperation method demonstrates regulated and consistent flight paths, over all three runs. Each of the three trajectories correlate to a straight-line origination from the starting position and ending at the target object position. There are minimal corrections occurring over the duration of the runs, relative to the previous teleportation methods. The data suggests that operator's situational awareness is elevated due to an increase in depth awareness of the robot's workspace.



(a) (b) (c)

Figure 7: Trajectory of the robot arm's end effector, for multiple runs, using the (a) line-of-sight keyboard teleoperation method, (b) monocular vision keyboard teleoperation method, and (c) Mixed Reality stereo vision and hand tracking teleoperation method.

Figure 8 illustrates the observations made from the trajectory graphs numerically by plotting the time to close the distance to the goal position. Upon inspection, both keyboard teleoperation methods show intervals in which the distance remains constant, demonstrating abrupt controls during teleoperation. However, the mixed reality teleoperation flight paths all demonstrate a consistent slope, with no abrupt control inputs from the operator. Further, on average, the time to reach the target for the LoS keyboard teleoperation method is 6.23 seconds. The average time for the monocular keyboard teleoperation method is 7.48 seconds. The average time for the mixed reality teleportation method is 4.5 seconds. Hence, the developed mixed reality teleoperation method performs the task, on average, 27.7% faster than the LoS teleoperation method and 39.82% faster that the monocular keyboard teleoperation method.

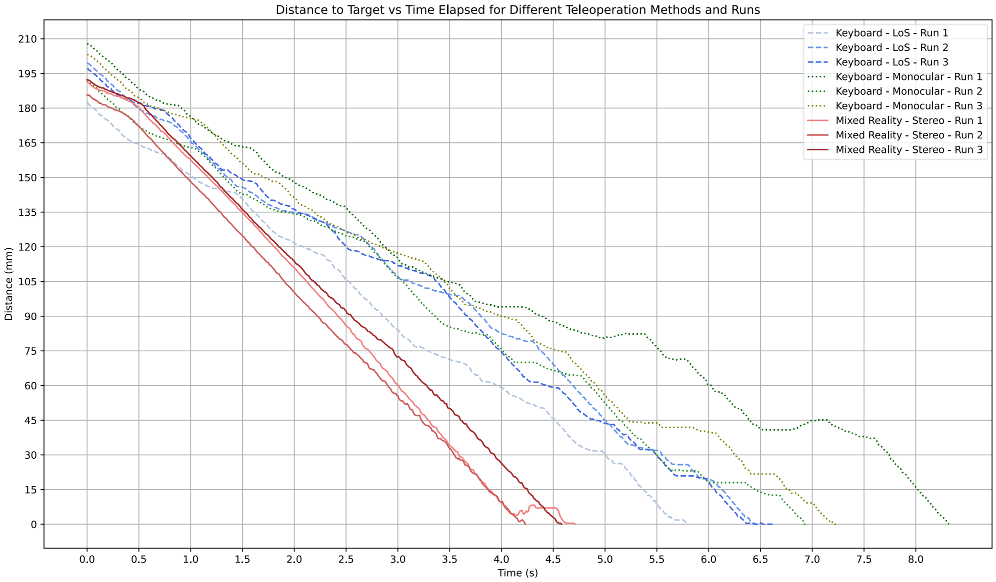


Figure 8: Distance to target vs time elapsed, for multiple runs, for each of the 3 teleportation methods.

5 CONCLUSION & FUTURE WORK

In this paper, we presented and implemented a mixed reality human robot interaction interface, that utilizes stereo vision and hand tracking, to teleoperate a robot arm towards performing intricate tasks in hazardous environments. The developed system was compared to conventional teleoperation methods, namely keyboard teleoperation by line of sight and monocular camera systems. The results of the experiments demonstrated that the MR HRI interface provides operators with a higher situational awareness, increased depth awareness, and ease of teleoperation. Especially, results demonstrated that the MR HRI interface performed 27.7% faster than the LoS teleoperation method and 39.82% faster that the monocular keyboard teleoperation method. The advantage this system provides is ability for operators to intricately control a robot arm from a safe location, while performing strenuous and hazardous tasks with ease. The development of this system enables for robots to be integrated into the dynamic and unstructured construction environment quickly.

At the time of writing, a 2-axis servo actuated gimbal is under development to mount the stereo camera onto. This will enable the camera system to orient itself to the look angles of the VR HMD, allowing the operator to inspect their surrounds during teleoperation. This upgrade is expected to further increase operator situational awareness. Further, software is under development to spatially map holographic models to the stereoscopic images, which will enable pre-modelled Building Information Model (BIM) data to be integrated to the MR interface. This will provide the operator to compare the as-built view with the as-designed assembly of structures in the robot's workspace. This technique will leverage the concepts of Simultaneous Localization and Mapping (SLAM), providing another layer of information visualization for the operator.

6 ACKNOWLEDGMENTS

The authors would like to acknowledge the support received from the Natural Science and Engineering Council of Canada Discovery Grant (NSERC DG 485794) and York University's Academic Innovation Fund (AIF 2023-2026).

7 REFERENCES

Brosque, Cynthia , Elena Galbally, Oussama Khatib, and Martin Fischer. 2020. "Human-Robot Collaboration in Construction: Opportunities and Challenges." In *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 1-8. IEEE.

Brown, Samantha, Raina D Brooks, and Xiuwen Sue Dong. 2020. *Nonfatal injury trends in the construction industry.* CPWR.

Campilho, Raul D. S. G., and Francisco J. G. Silva. 2023. "Industrial Process Improvement by Automation and Robotics." *Machines.*

Castro-Lacouture, Daniel. 2009. "Construction Automation." In *Springer Handbook of Automation*, 1063-1078. Berlin, Heidelberg: Springer Berlin Heidelberg.

DANDIL, Emre , and Kerim Kürşat ÇEVİK. 2019. "Computer Vision Based Distance Measurement System using Stereo Camera View." In *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 1-4. IEEE.

Ellis, Stephen R., Bernard D. Adelstein, and Robert B. Welch. 2002. "Kinesthetic Compensation for Misalignment of Teleoperator Controls through Cross-Modal Transfer of Movement Coordinates." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 1551-1555.

Fu, Bo, Will Seidelman, Yukang Liu, Travis Kent, Melody Carswell, Yuming Zhang, and Ruigang Yang. 2014. "Towards virtualized welding: Visualization and monitoring of remote welding." In *2014 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE.

Han, Xu, Ying Chen, Qinna Feng, and Heng Luo. 2022. "Augmented Reality in Professional Training: A Review of the Literature from 2001 to 202." *Applied Sciences.*

Illing, Boris , Martin Westhoven, Bastian Gaspers, Nanja Smets, Bernd Brüggemann, and Tintu Mathew. 2020. "Evaluation of Immersive Teleoperation Systems using Standardized Tasks and Measurements." In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 278-285. IEEE.

Kangari, Roozbeh. 1985. "Robotics Feasibility in the Construction Industry." In *Proceedings of the 2nd International Symposium on Automation and Robotics in Construction (ISARC)*, 66-103. Pittsburgh: International Association for Automation and Robotics in Construction (IAARC).

Luo, Yiming , Jialin Wang, Hai-Ning Liang, Shan Luo, and Eng Gee Lim. 2021. "Monoscopic vs. Stereoscopic Views and Display Types in the Teleoperation of Unmanned Ground Vehicles for Object Avoidance." In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, 418-425. IEEE.

Mukherjee, Debasmita, Kashish Gupta, Li Hsin Chang, and Homayoun Najjaran. 2022. "A Survey of Robot Learning Strategies for Human-Robot Collaboration in Industrial Settings." *Robotics and Computer-Integrated Manufacturing* 102231.

Sievers, Torsten Sebastian, Bianca Schmitt, Patrick Rückert, Maren Petersen, and Kirsten Tracht. 2020. "Concept of a Mixed-Reality Learning Environment for Collaborative Robotics." *Procedia Manufacturing* 19-24.

Wang, Baicun, S. Jack Hu, Lei Sun, and Theodor Freiheit. 2020. "Intelligent welding system technologies: State-of-the-art review and perspectives." *Journal of Manufacturing Systems* 373-391.

Wonsick, Murphy, and Taskin Padir. 2020. "A Systematic Review of Virtual Reality Interfaces for Controlling and Interacting with Robots." *Applied Sciences.*

Yang, Xiaodong, Qi Yu, Yan Zhang, and Weichun Ma. 2023. "Occupational health risk assessment of construction workers caused by particulate matter exposure on construction sites." *Heliyon.*

Zhang, Ming, Rui Xu, Haitao Wu, Jia Pan, and Xiaowei Luo. 2023. "Human–robot collaboration for on-site construction." *Automation in Construction* 104812.